A New Concept of Fractional Order Cumulant and It-based signal processing in α and/or Gaussian Noise

Yi-Ran Shi, Ding-Li Yu, Hong-Yan Shi and Yao-Wu Shi

Abstract—**In this paper, the concept and definitions of the Fractional Order Moment (FOM) and Fractional Order Cumulant (FOC) are proposed, which is based on the fractional derivative of the fractional order Moment-generating function and the fractional order Cumulant-generating function of stochastic processes. The moment and cumulant are defined on an expanded set from positive integer to the whole positive real. This development not only provides a new technology for signal processing, also complements the existing theory in the field. The properties of the FOC have been derived, and their uniformity and particularity with the High Order Cumulant are compared and commented. In addition, the transformation between the FOM and the FOC are derived and discussed in detail. As one of the applications of the new concept to the α and Gaussian processes, a new method of suppressing α and Gaussian noise is proposed. Furthermore, a FOC-based parameter estimation algorithm is developed for the non-minimum phase ARMA processes in α and/or Gaussian noise. Simulation examples are used to demonstrate the effectiveness of the proposed parameter estimation algorithm.**

Index Terms— Fractional order cumulant, fractional order moment, α noise, Gaussian noise, parameter estimation of stochastic processes.

I. Introduction

In traditional research in the area of signal processing, most theories and methods are based on the assumption that the additional noise is of Gaussian distribution. Under this assumption, the second-order moment or High Order Cumulant (HOC) has been widely used as a signal processing tool with excellent performance. Especially, the properties of linear, semi-invariance and the powerful suppression of colored Gaussian noise of HOC, make it an ideal statistical signal processing tool for processes with colored Gauss noise[1]. Therefore, HOC greatly promotes the development of signal processing theory and methods.

Unfortunately, many non-Gaussian noises exist in natural processes in addition to the Gaussian noise [2]. Among them, α stable distribution noise (for short it will be called α noise in the rest of the paper) has the most serious impact on the traditional signal processing methods [3]. Due to the infinite variance of the α noise, the performance of the traditional second-order moment or HOC-based signal processing algorithms deteriorates drastically or even fails in the circumstance where α noise exists [3]. Nevertheless, the Fractional Lower Order Moment (FLOM) exists in α stable distribution process [4]. Therefore, FLOM and various Fractional Lower Order Statistics (FLOS) operators developed from FLOM, such as Negative Order Statistics (ZOS) [5], Covariation [6], Fractional Lower Order Covariance (FLOC) [2], Zero Order Statistics (ZOS) [3, 7, 8], etc. naturally become the basic methods of signal processing for α stable distribution processes and have been widely used [9-13].

However, there are several major defects of the above FLOS operators. First, FLOS operators are nonlinear operators, which make the derivation and theoretical analysis of their signal processing algorithms much more difficult. Second, FLOS operators do not have the property of semi-invariant, which makes it impossible to effectively separate signals and noises. Third, the FLOS of α stable distribution process is not zero, which indicates that the α noise suppression of FLOS operators is poor and will lead to the degradation of signal processing performance.

Due to the problems stated above, the existing FLOS-based signal processing algorithms strictly limit the additional noise to i.i.d α stable distribution process (hereinafter-referred white α noise). When the additional noise is the colored α noise, especially

Yi-Ran Shi is with the College of Communication Engineering, Jilin University, Changchun 130025, P.R.China (e-mail: shiyiran@jlu.edu.cn).
Ding-Li Yu is with the Control Systems Research Group, School of Engineering, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, U.K. (e-mail: d.yu@ljmu.ac.uk).
Hong-Yan Shi is with the College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518060 P.R.China (e-mail: shy-no.1@163.com).
Yao-Wu Shi is with the College of Communication Engineering, Jilin University, Changchun 130025, P.R.China (e-mail: shiyw@jlu.edu.cn).

when the additional noise is a mixture of colored α noise and Gauss colored noise with more practical significance, the performance of them-based signal processing algorithms will decline sharply, or even fail.

In recent years, with the rapid development of fractional calculus theory, a novel mathematical tool is available for solving the signal processing problems in colored α noise. Fractional calculus extends the traditional integral order calculus operation to any non-integral order operation, it is therefore extremely suitable for dealing with the stochastic processes with fractional order exponential form of characteristic functions, such as α noise. Based on the fractional calculus theory, for the problems of the HOC and FLOS, a new concept of the Fractional Order Cumulant (FOC) is proposed in this paper. By analyzing the properties of FOC and the suppression of α noise and Gaussian noise, it is evident that the FOC is very suitable for signal processing in colored α and/or Gaussian noise.

The rest of the paper is organized as follows. Firstly, based on the fractional calculus theory, the definition of the FOC together with the definition of the Fractional Order Moment (FOM) is given in Section Ⅱ. Then, the properties of the FOC and the FOM, the conversion between the two operators, and the estimation methods of the FOC and the FOM are derived and analyzed in Section III. Section IV presents the major properties of the FOC of α and Gaussian processes. In addition, the suppression methods of the α and Gauss noise are proposed in this Section. Furthermore, as an application example of FOC in signal processing, a FOC-based parameter estimation algorithm for the non-minimum phase ARMA processes in colored α and/or Gaussian noise is developed in Section V. Finally, some conclusions are drawn in Section VI.

II. Definitions and Estimations of FOM and FOC

For the convenience of the following description, the Caputo fractional order derivative, the sequence derivative rule and the definition of fractional Fourier transform are given below as mathematical preliminaries.

Definition of Caputo fractional order derivative [14]

Let $f(t)$ be a function defined on the interval (a,b) , $p>0$, $\inf_{n\in\mathbb{N}}\{n: n\geq p\}$, $t>a$. Then, the p^{th} -order left Caputo fractional order derivative is defined as

$$ {}_a^C D_t^p f(t)=\frac{1}{\Gamma(n-p)}\bullet\int_a^t (t-\xi)^{n-p-1} f^{(n)}(\xi)d\xi \tag{1}$$

where, $\Gamma(\bullet)$ is the usual gamma function, as defined in [14],

$$\Gamma(x)=\int_0^\infty t^{x-1}e^{-t}dt \tag{2}$$

In this paper, the fractional order derivative operator in equation (1) is denoted as ${}_a^C D^p f(t)$ or ${}_a^C d^p f(t)/dt^p$. When the lower limit $a=0$, it is denoted as $f^{(p)}(t)$, i.e. $f^{(p)}(t)={}_0^C D^p f(t)$.

Remark 1. There are several definitions of fractional derivative, such as Riemann-Liouville, Grunwald-Letnikov, etc. These definitions each has its own characteristics and can be converted to each other under certain conditions. Take Riemann Liouville fractional derivative as an example. It is defined as [14]:

$$ {}_a^{RL} D_t^p f(t)=\frac{1}{\Gamma(n-p)}\frac{d^n}{dt^n}\int_a^t (t-\xi)^{n-p-1} f(\xi)d\xi \tag{3}$$

As shown in (3), Riemann Liouville fractional derivative is an operation that first performs integration on $(t-\xi)^{n-p-1} f(\xi)$ and then performs n order derivative operation. Although this definition has the advantage of close connection with the traditional integral calculus, there are still some problems such as the fractional derivative of constant is not zero, and the choice of the initial value of fractional derivative is difficult. The Caputo fractional order derivative shown in (1) adopts the definition of first differentiation and then integration to better solve these problems, it is therefore especially suitable for engineering applications. This is also the main reason that the Caputo fractional order derivative is used in this paper.

Definition of sequential fractional order derivative [14]

Let ${}_a^C D^p$ is the p^{th} -order left Caputo fractional order derivative, then

$$ {}_a^C D^{kp} f(t)=\underbrace{{}_a^C D^p\,{}_a^C D^p\cdots{}_a^C D^p}_{k\ \ times} f(t) \tag{4}$$

is the kp^{th} -order left Caputo sequential fractional order derivative，with no possibility of confusion, hereinafter-referred kp^{th} -order fractional order derivative.

Remark 2. For the calculation of Caputo fractional order derivative, ${}_a^C D^{kp} f(t)$， $k>1$, $0<p\leq 1$，there are two different methods: one is to calculate directly by (1) according to the definition, and the other is to calculate by (4) using the sequential fractional order derivative method. It should be noted that the results obtained by these two methods are inconsistent. This is due to

the definition and properties of ${}_a^C D^{kp} f(t)$ itself. Therefore, which method should be used is selected according to the specific application [15]. The reason why the sequential fractional order derivative is used in this paper is that the problem we are dealing with involves ${}_a^C D^p$ as the basic derivative. Secondly, the Taylor series obtained based on the sequential fractional order derivative has the characteristics of power - series of x^{kp}. These greatly facilitate our discussion of the problems involved.

Definition of Fractional Fourier transform [15]

If $f(x)$ is continuous and absolutely integrable in interval $(-\infty,\infty)$, its fractional Fourier transform is defined as

$$F_p\left\{f(x)\right\} = \int_{-\infty}^{\infty} E_p(j^p u^p x^p) f(x)(dx)^p \tag{5}$$

where, $E_p(j^p u^p x^p) = \sum_{k=0}^{\infty} \frac{(jux)^{kp}}{\Gamma(kp+1)}$, $(0 < p \le 1)$, is the Mittag-Leffler function [15].

As the basis of the research on the FOM and FOC, the fractional order Moment-generating function should be defined first.

Definition of fractional order Moment-generating function

Let fractional order probability density function of random variable x is $f(x)$, then define that

$$\Phi_p(u) = \int_{-\infty}^{\infty} E_p(j^p u^p x^p) f(x)dx = E\left\{E_p(j^p u^p x^p)\right\} \tag{6}$$

is the fractional order Moment-generating function (fractional order characteristic function) of variable x.

Definition of fractional order Cumulant-generating function

By $\Phi_p(0)=1$ and the continuity of $\Phi_p(u)$, if $\exists\ \delta > 0$, s.t. $|u| < \delta$, then $\Phi_p(u) > 0$. So the fractional order Cumulant-generating function $\Psi_p(u)$ is defined as the logarithm of the fractional order Moment-generating function $\Phi_p(u)$ of random variable x, i.e.

$$\Psi_p(u) = \ln \Phi_p(u) \tag{7}$$

After defining the above functions, we are now ready to define FOM and FOC.

A. Definitions of FOM and FOC

1. Definitions of FOM and FOC of single random variable

Take the generalized fractional Taylor series expansion [15, 16] for $\Phi_p(u)$ at $u=0$, then

$$\Phi_p(u) = \sum_{k=0}^{\infty} \frac{u^{kp}}{\Gamma\left(kp+1\right)} \Phi_p^{(kp)}(0) \tag{8}$$

where $0 < p \le 1$, k is a positive integer, $\Phi_p^{(kp)}(0) = {}_0^C D^{kp} \Phi_p(u)\big|_{u=0}$, and by equation(6), then

$$\begin{aligned} {}_0^C D^{kp}\Phi_p(u)\big|_{u=0} &= {}_0^C D^{kp} E\left\{E_p(j^p u^p x^p)\right\}\big|_{u=0} \\ &= E\left\{{}_0^C D^{kp}\left\{E_p(j^p u^p x^p)\right\}\right\}\big|_{u=0} \\ &= E\left\{j^{kp} x^{kp} E_p(j^p u^p x^p)\right\}\big|_{u=0} \\ &= j^{kp} E\left[X^{kp}\right] \end{aligned} \tag{9}$$

Substituting (9) into (8), yields

$$\begin{aligned} \Phi_p(u) &= \sum_{k=0}^{\infty} \frac{u^{kp}}{\Gamma\left(kp+1\right)} j^{kp} E[X^{kp}] \\ &= \sum_{k=0}^{\infty} m^{kp} \frac{j^{kp} u^{kp}}{\Gamma\left(kp+1\right)} \end{aligned} \tag{10}$$

where $m^{kp} = E\left[X^{kp}\right] = \Phi_p^{(kp)}(0) j^{-kp}$.

Definition: Define $m_x^{kp} = E[X^{kp}] = j^{-kp}\, {}_0^C D^{kp} \Phi_p(u)\big|_{u=0}$, $0 < p \le 1$, k is an integer, as the kp^{th} -order Fractional Order Moment (FOM) of the single random variable X.

Especially, in the above definition, when $k=1$, i.e. $m^p = E\left[X^p\right]$, is called the M^p mean value of the single random variable.

Similar to the above, take the generalized fractional Taylor series expansion for $\Psi_p(u)$ in(7) at $u=0$, then

$$\begin{aligned}\Psi_p(u) &= \sum_{k=0}^{\infty} \frac{\Psi_p^{(kp)}(0)}{\Gamma(kp+1)} u^{kp} \\ &= 0 + \sum_{k=1}^{\infty} j^{-kp} \frac{\Psi_p^{(kp)}(0)}{\Gamma(kp+1)} j^{kp} u^{kp} \\ &= \sum_{k=1}^{\infty} C^{kp} \frac{j^{kp} u^{kp}}{\Gamma(kp+1)}\end{aligned} \tag{11}$$

where, $C^{kp} = j^{-kp}\Psi_p^{(kp)}(0) = j^{-kp}\, {}_0^C D^{kp} \ln \Phi_p(u)\big|_{u=0}$.

Definition Define $C^{kp} = j^{-kp}\, {}_0^C D^{kp} \Psi_p^{(kp)}(0)\big|_{u=0}$, $0 < p \le 1$, k is an integer, as the kp^{th} -order Fractional Order Cumulant (FOC) of the single random variable.

2. Definitions of FOM and FOC of multiple random variables

Let $f(x_1, x_2, \cdots, x_k)$ be the fractional joint probability density function of k real-value random variables $x_1, x_2, \cdots, x_k$. Then, the k dimension fractional joint Moment-generating function can be denoted as:

$$\begin{aligned}&\Phi_p(u_1, u_2, \cdots, u_k) = E\left(E_p((ju_1x_1)^p)E_p((ju_2x_2)^p)\cdots E_p((ju_kx_k)^p)\right) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1, x_2, \cdots, x_k) E_p((ju_1x_1)^p)E_p((ju_2x_2)^p)\cdots E_p((ju_kx_k)^p)(dx_1)^p \cdots (dx_k)^p\end{aligned} \tag{12}$$

The kp^{th} -order fractional partial derivative in (12) respect to $u_1, \cdots, u_k$ is

$$\begin{aligned}\frac{{}_0^C\partial^{kp}\Phi_p(u_1, \cdots, u_k)}{\partial u_1^p \cdots \partial u_k^p} &= E\left\{\frac{{}_0^C\partial^{kp} E_p((ju_1x_1)^p)E_p((ju_2x_2)^p)\cdots E_p((ju_kx_k)^p)}{\partial u_1^p \cdots \partial u_k^p}\right\} \\ &= j^{kp} E\left\{x_1^p x_2^p \cdots x_k^p E_p((ju_1x_1)^p)E_p((ju_2x_2)^p)\cdots E_p((ju_kx_k)^p)\right\}\end{aligned} \tag{13}$$

where $0 < p \le 1$, k is an integer in (12), (13).

By using a similar function to the definition in (10), the kp^{th} -order FOM of the real-value random variables $x_1, \cdots, x_k$ is defined as

$$\begin{aligned}m^{kp} &= mom^{kp}(x_1^p, \cdots, x_k^p) = E\left[x_1^p \cdots x_k^p\right] \\ &= j^{-kp} \frac{\partial^{kp}\Phi_p(u_1, \cdots, u_k)}{\partial u_1^p \cdots \partial u_k^p}\bigg|_{u_1 = \cdots = u_k = 0}\end{aligned} \tag{14}$$

Similarly, the k dimension fractional joint Cumulant-generating function $\Psi_p(u_1, \cdots, u_k)$ is denoted as

$$\Psi_p(u_1, \cdots, u_k) = \ln\left[\Phi_p(u_1, \cdots, u_k)\right] \tag{15}$$

Also, by using a similar function to that in (11), the kp^{th} -order FOC of the real-value random variables $x_1, x_2, \cdots, x_k$ is defined as

$$C^{kp} = cum^{kp}\left[x_1^p, \cdots, x_k^p\right] \triangleq j^{-kp} \frac{{}_0^C\partial^{kp}\Psi_p(u_1, \cdots, u_k)}{\partial u_1^p \cdots \partial u_k^p}\bigg|_{u_1 = \cdots = u_k = 0} \tag{16}$$

3. Definitions of FOM and FOC of continuous stationary stochastic process

Let $\{x(t), t \in T\}$ be the stationary stochastic process. The random vector $X = \left[x(t), x(t+\tau_1), \cdots, x(t+\tau_{k-1})\right]^T$ is obtained by sampling $\{x(t)\}$ at any k sample instant $t, t+\tau_1, \cdots, t+\tau_{k-1}$. Referring to the definition of FOM and FOC of multiple random variables, its kp^{th} -order FOM and kp^{th} -order FOC are respectively defined as

$$m^{kp}(\tau_1, \tau_2, \cdots, \tau_{k-1}) \triangleq mom^{kp}\left[x^p(t), x^p(t+\tau_1), \cdots, x^p(t+\tau_{k-1})\right] \tag{17}$$

$$C^{kp}(\tau_1, \tau_2, \cdots, \tau_{k-1}) \triangleq cum^{kp}\left[x^p(t), x^p(t+\tau_1), \cdots, x^p(t+\tau_{k-1})\right] \tag{18}$$

Remark 3. By the definitions of FOM and FOC above, it is clear that while $p = 1$, the FOM and FOC can be transformed into the traditional High Order Moment (HOM) and HOC respectively. Therefore, the FOM and FOC fill the gaps between the integer orders of the HOM and HOC, and extends the order definitions from positive integer to the complete positive real field. Just because of the introduction of fractional order, the FOM and FOC have the added properties that the HOM and HOC do not possess. These properties enable some new methods (will be described latter) for signal processing.

Furthermore, it is noteworthy that FOM (FOC) changes the integer order of HOM (HOC) into fractional order, which causes a little increase of the computational complexity. However, It is believed that for the great benefit of improved property that brought by the FOC, paying the cost of a little extra computing load is really worth.

B. Conversion between FOM and FOC

By the definitions of the k dimension fractional joint Moment-generating function and the k dimension fractional joint Cumulant-generating function, then

$$\Phi_p\left(u_1,u_2,\cdots,u_k\right)=e^{\Psi_p(u_1,\cdots,u_k)} \tag{19}$$

Take multi-dimensional fractional Taylor series expansion [16] at both sides of equation (19) respectively, and compare the same power coefficients of $u_1^p u_2^p \cdots u_k^p$ at both sides of the equation, the conversion between the FOM and FOC can be obtained as follows,

$$mom^{kp}(I)=\sum_{\bigcup_{m=1}^{q} I_m=I}\prod_{m=1}^{q} cum^{kp}(I_m) \tag{20}$$

$$cum^{kp}(I)=\sum_{\bigcup_{m=1}^{q} I_m=I} \mathrm{e}^{-\mathrm{j}\pi(q-1)}(q-1)!\prod_{m=1}^{q} mom^{kp}(I_m) \tag{21}$$

where, $I=(1,2,\cdots,k)$ is the set of indices of the components of random vector $X=\left[x_1,x_2,\cdots,x_k\right]$, k is the dimension of random vector. $mon^{kp}(I)=mon^{kp}\left[x_1^p,x_2^p,\cdots,x_k^p\right]$, $com^{kp}(I)=com^{kp}\left[x_1^p,x_2^p,\cdots,x_k^p\right]$, $0<p\le 1$. Set I_m as the unordered collection of set I. q is the number of partitions of I_m and takes number $1,2,\cdots,k$. $\sum_{\bigcup_{m=1}^{q} I_m=I}$ denotes the sum of the functions corresponding to all I_m. The derivation of equations (20) and (21) is detailed in **Appendix 1**.

For the convenience of description, equation (20) is called Fractional Cumulant to Moment (FC-M) formula, and equation (21) is called Fractional Moment to Cumulant (FM-C) formula.

Here, we take $com^{4p}\left[x_1^p,x_2^p,x_3^p,x_4^p\right]$ calculation as an example to illustrate the specific application of the above FM-C formula.

If $X=\left[x_1,x_2,x_2,x_4\right]$, then $k=4$, $I=(1,2,3,4)$. I can be divided into the following four classifications:

$I_{m=1}$: $q=1$, I remains as 1 subset. There is only one case, i.e. $\{1,2,3,4\}$.

$I_{m=2}$: $q=2$, I is divided into 2 subsets. There are 7 cases, i.e. $\{1\}\{2,3,4\}$, $\{2\}\{1,3,4\}$, $\{3\}\{1,2,4\}$, $\{4\}\{1,2,3\}$, $\{1,2\}\{3,4\}$, $\{1,3\}\{2,4\}$, $\{1,4\}\{2,3\}$.

$I_{m=3}$: $q=3$, I is divided into 3 subsets. There are 6 cases, i.e. $\{1\}\{2\}\{3,4\}$, $\{1\}\{3\}\{2,4\}$, $\{1\}\{4\}\{2,3\}$, $\{2\}\{3\}\{1,4\}$, $\{2\}\{4\}\{1,3\}$, $\{3\}\{4\}\{1,2\}$.

$I_{m=4}$: $q=4$, I is divided into 4 subsets. There is only one case, i.e. $\{1\}\{2\}\{3\}\{4\}$.

Substituting all the above cases into the FM-C formula, then

$$\begin{aligned}
cum^{4p}(x_1^p,x_2^p,x_3^p,x_4^p) &= mom^{4p}(x_1^p,x_2^p,x_3^p,x_4^p)-mom^p(x_1^p)mom^{3p}(x_2^p,x_3^p,x_4^p)-mom^p(x_2^p)mom^{3p}(x_1^p,x_3^p,x_4^p)\\
&-mom^p(x_3^p)mom^{3p}(x_1^p,x_2^p,x_4^p)-mom^p(x_4^p)mom^{3p}(x_1^p,x_2^p,x_3^p)-mom^{2p}(x_1^p,x_2^p)mom^{2p}(x_3^p,x_4^p)\\
&-mom^{2p}(x_1^p,x_3^p)mom^{2p}(x_2^p,x_4^p)-mom^{2p}(x_1^p,x_4^p)mom^{2p}(x_2^p,x_3^p)+2mom^p(x_1^p)mom^p(x_2^p)mom^{2p}(x_3^p,x_4^p)\\
&+2mom^p(x_1^p)mom^p(x_3^p)mom^{2p}(x_2^p,x_4^p)+2mom^p(x_1^p)mom^p(x_4^p)mom^{2p}(x_2^p,x_3^p)\\
&+2mom^p(x_2^p)mom^p(x_3^p)mom^{2p}(x_1^p,x_4^p)+2mom^p(x_2^p)mom^p(x_4^p)mom^{2p}(x_1^p,x_3^p)\\
&+2mom^p(x_3^p)mom^p(x_4^p)mom^{2p}(x_1^p,x_2^p)-6mom^p(x_1^p)mom^p(x_2^p)mom^p(x_3^p)mom^p(x_4^p)
\end{aligned} \tag{22}$$

As shown in equation(22), in general, the above FM-C formula is relatively complex. While the order of FOC increases further, the computational complexity will increase significantly. And the same is true for FC-M formula. However, for the zero-M^p mean stationary stochastic processes, due to $mom(x_i^p)=0$, $cum^p(x_i^p)=0$, $i=1,\cdots,k$. The above FM-C and FC-M formulas can be simplified as follows:

$$mom^{kp}(I)=\sum_{\bigcup_{m=1}^{q}I_m=I-I_e}\prod_{m=1}^{q}cum^{kp}(I_m) \tag{23}$$

$$cum^{kp}(I)=\sum_{\bigcup_{m=1}^{q}I_m=I-I_e}(-1)^{(q-1)}(q-1)!\prod_{m=1}^{q}mom^{kp}(I_m) \tag{24}$$

where, $I_m=I-I_e$ is a subset with no size-one elements in set I .

Still take the $4p$ -th order FOC calculation as an example. While $mom(x_i^p)=0$, equation (22) can be simplified as

$$\begin{aligned}cum^{4p}(x_1^p,x_2^p,x_3^p,x_4^p)&=mom^{4p}(x_1^p,x_2^p,x_3^p,x_4^p)-mom^{2p}(x_1^p,x_2^p)mom^{2p}(x_3^p,x_4^p)\\&-mom^{2p}(x_1^p,x_3^p)mom^{2p}(x_2^p,x_4^p)-mom^{2p}(x_1^p,x_4^p)mom^{2p}(x_2^p,x_3^p)\end{aligned} \tag{25}$$

Since any non-zero- M^p mean stationary stochastic process can be transformed into zero- M^p mean process simply by de-zero- M^p mean processing, equations (23) and (24) are very useful to simplify the calculation of FM-C and FC-M.

C. Estimation of FOC

When the fractional order Moment-generating function of the stochastic process is known, the FOC of any order can be obtained directly by the definition. However, in practical applications the fractional order Moment-generating function of the stochastic processes is generally unknown. In this case, we can only obtain the FOM and FOC by estimation. Using an estimation method for mathematical expectation and the FM-C conversion formula, we developed an estimation method for FOM and FOC of any order as follows.

Let the stationary stochastic process $x(n)$ has up to $2kp^{th}$ -order FOM. Because the FOM of stochastic processes is defined by the form of mathematical expectation, i.e. $mon^{kp}(x_1^p,x_2^p,\cdots,x_k^p)=E(x_1^p x_2^p\cdots x_k^p)$, the mean squared uniform estimation of kp^{th} -order FOM of $x(n)$ is

$$\begin{aligned}\hat{m}^{kp}(\tau_1,\tau_2,\cdots,\tau_{k-1})&=E\left[x^p(t)x^p(t+\tau_1)\cdots x^p(t+\tau_{k-1})\right]\\&=\frac{1}{N}\sum_{n=1}^{N-\tau}x^p(n)x^p(n+\tau_1)\cdots x^p(n+\tau_{k-1})\end{aligned} \tag{26}$$

where, $\tau=Max(\tau_1,\cdots,\tau_{k-1})$. While the FOM estimation of each order of $x(n)$ is obtained, the corresponding FOC estimation can be obtained by the FM-C conversion formula.

The definitions of the FOM and FOC, relations between the FOC and the HOC, as well as the conversion between the FOM and FOC are presented or derived in this section. Some important properties of both the FOM and FOC will be derived in the next section.

III. PROPERTIES OF FOM AND FOC

Five important properties of the FOM and FOC are derived according to their definitions and the fractional derivative feature, and presented as follows. The proofs of the five properties are given in the **Appendix 2**.

Property 1

If $a_1,a_2,\cdots,a_k$ are arbitrary constants and $X(k)=\left[x_1,x_2,\cdots,x_k\right]$ is a random variable, then

$$mom^{kp}\left[a_1x_1^p,\cdots,a_kx_k^p\right]=a_1\cdots a_k mom^{kp}\left[x_1^p,\cdots,x_k^p\right] \tag{27}$$

$$cum^{kp}\left[a_1x_1^p,\cdots,a_kx_k^{p_m}\right]=a_1\cdots a_k cum\left[x_1^p,\cdots,x_k^p\right] \tag{28}$$

Property 2

The FOM and FOC are symmetric in their arguments, i.e.

$$mom^{kp}\left[x_1^p,\cdots,x_k^p\right]=mom^{kp}\left[x_{i_1}^p,\cdots,x_{i_k}^p\right] \tag{29}$$

$$cum^{kp}\left[x_1^p,\cdots,x_k^p\right]=cum^{kp}\left[x_{i_1}^p,\cdots,x_{i_k}^p\right] \tag{30}$$

where, $i_1,\cdots,i_k$ are arbitrary different permutations of $1,\cdots,k$.

Property 3

If a subset of the set of random variables $\{x_i\}$ is independent of the rest subsets, then

$$cum^{kp}\left[x_1^p,\cdots,x_k^p\right]\equiv 0 \tag{31}$$

But normally,

$$mom^{kp}\left[x_1^p,\cdots,x_k^p\right]\neq 0 \tag{32}$$

Property 4

If the random variables $\left[x_1,\cdots,x_k\right]$ are independent of the random variables $\left[y_1,\cdots,y_k\right]$, then

$$\begin{aligned}&cum^{kp}\left[\left(x_1+y_1\right)^p,\cdots,\left(x_k+y_k\right)^p\right]\\&=cum^{kp}\left[x_1^p,\cdots,x_k^p\right]+cum^{kp}\left[y_1^p,\cdots,y_k^p\right]\end{aligned} \tag{33}$$

But, generally,

$$\begin{aligned}&mom^{kp}\left[\left(x_1+y_1\right)^p,\cdots,\left(x_k+y_k\right)^p\right]\\&\neq mom^{kp}\left[x_1^p,\cdots,x_k^p\right]+mom^{kp}\left[y_1^p,\cdots,y_k^p\right]\end{aligned} \tag{34}$$

This property of the FOC is referred to semi-invariant, which implies that the FOC of the sum of two statistically independent stochastic processes equals the sum of the FOCs of the individual stochastic process. But, this is not true for FOM.

Property 5

If a is an arbitrary constant, $k>1$, then

$$cum^{kp}\left[\left(x_1+a\right)^p,x_2^p,\cdots,x_k^p\right]=cum^{kp}\left[x_1^p,x_2^p,\cdots,x_k^p\right] \tag{35}$$

Equation (35) means that FOC is blind to any additive constants. But this feature does not apply to the FOM, i.e.

$$mom^{kp}\left[\left(x_1+a\right)^p,x_2^p,\cdots,x_k^p\right]\neq mom^{kp}\left[x_1^p,x_2^p\cdots,x_k^p\right] \tag{36}$$

IV. Suppression of FOC for α and Gaussian Distributions

In consideration of the importance of Symmetric α Stable ($S\alpha S$) distribution process in signal processing theory and applications, we focus on the properties of FOC of $S\alpha S$ process. Based on these properties, a method for suppressing $S\alpha S$ distribution process is proposed in this section. Firstly, the definition of the fractional order characteristic function of $S\alpha S$ distribution random variable X is given as below [3].

$$\phi_e(u)=\exp\{jau-\gamma\,|\,u\,|^{\alpha}\} \tag{37}$$

where, $a\left(-\infty<a<\infty\right)$ is the location parameter, $\alpha\left(0<\alpha\leq 2\right)$ is the characteristic exponent, $\gamma\left(\gamma\geq 0\right)$ is the dispersion parameter.

In equation (37), while $\alpha=2$, α stable distribution degenerates to Gaussian distribution, i.e. Gaussian distribution is a special case of α stable distribution.

Theorem 1

Let random variable $X\sim S\alpha S$, $m\in □^{+}$, and $\phi_e(u)=\exp\{jau-\gamma\,|\,u\,|^{\alpha}\}=\exp\left\{jau\right\}\exp\left\{-\gamma\,|\,u\,|^{\alpha}\right\}=\phi_a(u)\phi_\alpha(u)$. Then, the kp^{th} -order FOC of X is

$${}^{C}C^{kp}=\underbrace{\begin{cases}0\,, & kp\neq 1\\ a\,, & kp=1\end{cases}}_{\text{acording to }\phi_a\text{(u) part}}+\underbrace{\begin{cases}0, & \alpha-kp>0\text{ and not an integer, or }1\leq\alpha-kp\leq m\ \textit{and}\text{ integer}\\ -\gamma\Gamma(\alpha+1)\,\mathrm{j}^{-kp}, & kp=\alpha\text{; particularly, while }kp=\alpha=1\text{, it can be simplified to }\gamma j\text{.}\\ \infty, & \alpha-kp<0\text{ and not an integer}\end{cases}}_{\text{acording to }\phi_\alpha\text{(u) part}} \tag{38}$$

The proof is given in **Appendix 3**.

By using the same proof method of Theorem 1, it can be proved that the above conclusions are also applicable to $x(t)\sim S\alpha S$ stochastic processes.

Remark 4. As shown in Theorem 1, for any stochastic process, $x(t)\sim S\alpha S$, whether it is white or colored, its kp^{th} -order FOC equals zero when $kp<\alpha$. This means the FOC can completely cancel the effect of the α noise in theory. This is a great significance for signal processing in α noise, for which details will be discussed in Section VI. Besides, since Gaussian distribution is a special case of α stable distribution for $\alpha{=}2$, this conclusion is still valid for the Gauss distribution process. Interestingly, if we take $\alpha{=}2$, the second item on the right hand side of equation (38) coincides with the property that the k^{th} -order ($k>2$) HOC of Gauss noise is zero. Therefore, this property of HOC is a special case of Theorem 1.

V. The relationship and difference between FOC and HOC and FLOS

From the definition and properties of FOC and the suppression method for α noise, it can be observed that although FOC operators have many similarities with the well-known HOC and FLOS operators, there are still many fundamental differences. Table 1 below concisely shows the characteristics, the main differences and relationships among FOC operators and HOC and FLOS operators.

TABLE I
MAIN SIMILARITIES AND DIFFERENCES AMONG FOC AND HOC AND FLOS

	FOC	HOC	FLOS
Definition	Fractional order derivative of fractional order Cumulant-generating function	Integer order derivative of Cumulant-generating function	Stochastic process theory
Domain of definition	The definition domain is the whole positive real number domain, but for the α stable distribution process, the definition domain is $kp < \alpha$.	The definition domain is only the positive integer domain, and it does not exist for α ($0 < \alpha < 2$) stable distribution process.	For the α stable distribution process, its definition domain is $p < \alpha$ [3]
Application area	Can be used to solve signal processing problems in α and/or Gaussian noise	Cannot be used to solve signal processing problems in α noise	Can be used to solve the signal processing problems in α and/or Gaussian noise
Linear characteristic	Possesses Linear and Semi-invariant	Possesses Linear and Semi-invariant	Does not possess linear and Semi-invariant
Noise suppression	It can suppress both colored or white α and Gauss noise.	Only colored or white Gaussian noise can be suppressed	Neither α nor Gauss noise is zero.
Algorithmic complexity	It-based signal processing method is simple and easy to analyze and calculate.	It-based signal processing method is simple and easy to analyze and calculate.	Signal processing method is complex
Theoretical analysis	Convenient for theoretical analysis	Convenient for theoretical analysis	Theoretical analysis is difficult.

As shown in Table I, the main problem of the famous HOC operator is that it cannot be used in the signal processing in α noise. This is due to the variance of α noise is infinite. Although FLOS operator can be used in the signal processing in α noise, its non-linear and non- semi-invariant properties will make it-based signal processing methods complex, and even hardly deal with the signal processing problems in colored α noise. Compare with HOC and FLOS, the FOC operator proposed in this paper is truly suitable for signal processing in α and/or Gaussian noise because of its properties of linear, semi-invariant and blind of α and Gaussian noise. This can be seen more clearly from the derivation process of the FOC-based parameter estimation algorithm of the ARMA model in the next section.

VI. A FOC-based parameter estimation of non-minimum phase ARMA model

The theoretical research in the literature shows that many stationary stochastic processes can be represented by ARMA model. It is very important to estimate the ARMA model parameters of a real life stationary random signal by using the noisy data from sensing elements. This technology has been widely used in real life applications, such as the signals of sonar, radar, plasma physics, biomedicine, seismic data processing, image reconstruction, harmonic retrieval, time delay estimation, adaptive filtering, array processing, and blind equalization, etc. [17]. Recently, based on HOC operator, some ARMA model parameter estimation methods in Gaussian noise were proposed [17-20]. But this kind of methods cannot solve the problem of model parameter estimation in α noise. In references [4, 5], FLOS operator was used to solve this problem. However, due to the limitations of the properties of FLOS operator, these methods can only be applied to the minimum phase ARMA models in white α noise. For the parameter estimation of non-minimum phase ARMA model in colored α noise, the relevant research has not been reported. In this section, we take this problem as an example to demonstrate the effectiveness of the FOC-based method proposed in this paper.

As mentioned above, HOC is a special case of FOC while $p=1$, and they have similar properties except that HOC cannot be used for signal processing in α noise. Therefore, the derivation process of signal processing methods based on FOC and HOC must have many similarities. In practice, most HOC-based signal processing methods can be improved to FOC-based methods to enhance the suppression ability to α noise. This is another important advantage of FOC operator. In order to make this more clearly, in this section, referring to the derivation process of the HOC-based SVD-TLS method of AR parameter estimation[17] and the HOC-based $q-$slice method of MA parameter estimation[17], the FOC-based ARMA parameter estimation method is proposed.

A. Model assumptions

Considering a Single Input and Single Output (SISO) Linear and Time-Invariant(LTI) stochastic process that can be described by the following ARMA (p,q) model[18]:

$$\sum_{i=0}^{p} a(i)x(n-i)=\sum_{j=0}^{q} b(j)e(n-j) \tag{39}$$

where, $a(i)$ and $b(i)$ are the model parameters of the AR and MA parts respectively, $x(n)$ is the noiseless system output, $e(n)$ is the system input.

The transfer function of the system is

$$H(z)=\frac{1+b_1 z^{-1}+b_2 z^{-2}+\cdots+b_q z^{-q}}{1+a_1 z^{-1}+a_2 z^{-2}+\cdots+a_p z^{-p}}=\sum_{i=0}^{\infty} h(i)z^{-i} \tag{40}$$

In practice, $x(n)$ is often corrupted by the additive noise $\omega(n)$. Thus, the observed output is

$$y(n)=x(n)+\omega(n) \tag{41}$$

Combining (40) and (41) yields

$$y(n)=\sum_{i=0}^{\infty} h(n-i)e(i)+\omega(n) \tag{42}$$

The following assumptions are made in the modeling:

[AS1] The order (p,q) of the ARMA model is known.

[AS2] The input $e(n)$ is an unobservable, zero- M^p mean, independent and identically distributed (i.i.d.) non-stable distribution process, with at least one finite, non-zero kp^{th} -order FOC ζ_{ke} , $k>2$.

[AS3] The system is causal, linear time invariant, exponentially stable and non-minimum phase, i.e. in the complex Z domain, the poles of the transfer function $H(\text{z})$ lie inside the unit circle and at least one zero lies outside the unit circle.

[AS4] The additive noise $\omega(n)$ is an unobservable colored Gaussian and/or α stable distribution process with known characteristic exponent $\alpha(0<\alpha<2)$. The Gaussian noise, α noise and $e(n)$ are independent of each other.

[AS5] $a(0)=b(0)=1$; this fixes the inherent scale ambiguity.

There are two major differences between the model assumption in this paper and in the reference [18] as follows:

1. In [AS2], this paper changes the input $e(n)$ with at least one finite, non-zero cumulant [18] to with at least one finite, non-zero kp^{th} -order FOC. This is necessary to ensure the establishment of FOC-based model parameter estimation algorithm, where the zero mean changed to zero- M^p mean in order to simplify the calculation of FOC of signals. While $p=1$, this assumption degenerates to the assumption of $e(n)$ in reference [18].

2. Another difference is the assumption of the additional noise $\omega(n)$. In reference[18], $\omega(n)$ is assumed to be colored Gaussian process. But in this paper, it is assumed to be colored Gaussian and/or α stable distribution process with characteristic exponent $\alpha(0<\alpha<2)$. It is precisely because of the existence of colored α noise that HOC method cannot be used. Furthermore, because FLOS method does not have the semi-invariant property and FLOS of colored α noise is not zero, it cannot be used either to solve this problem. At present, there is no solution to this kind of signal processing problem. However, FOC method proposed in this paper is very suitable for solving it.

B. FOC-based Yuel-Walker function of ARMA model

Take the kp^{th} -order ($kp<\alpha, k>2$) FOC of $y(n)$, by Properties 4 and Theorem 1, then

$$\begin{aligned} C_{ky}^{kP}(m_1,m_2,\cdots,m_{k-1}) &= cum_{ky}^{kP}[y(n),y(n+m_1),\cdots,y(n+m_{k-1})] \\ &= cum_{ky}^{kP}[x(n)+\omega(n),x(n+m_1)+\omega(n+m_1),\cdots,x(n+m_{k-1})+\omega(n+m_{k-1})] \\ &= cum_{kx}^{kP}[x(n),x(n+m_1),\cdots,x(n+m_{k-1})]+cum_{k\omega}^{kP}[\omega(n),\omega(n+m_1),\cdots,\omega(n+m_{k-1})] \\ &= cum_{kx}^{kP}[x(n),x(n+m_1),\cdots,x(n+m_{k-1})] \end{aligned} \tag{43}$$

This shows that the additional noise $\omega(n)$ (whether $\omega(n)$ is α and/or Gaussian, colored or white) can be completely suppressed by using kp^{th} -order FOC.

By the stability of $y(n)$, $C_{2y}^{2P}(m)=C_{2y}^{2P}(-m)$. This means that $2p^{th}$ -order FOC $C_{2y}^{2P}(m)$ contains only the amplitude information of $H(z)$ without the phase information. Therefore, it cannot be used for solving the non-minimum phase model parameter estimation problem. For the FOC-based non-minimum phase model parameter estimation problem, we have the following theorem.

Theorem 2

Let the LTI system $H(z)$ be as shown in (40). The input $e(n)$ is an unobservable, zero-M^p mean, i.i.d. non-stable process, FOC of $e(n)$ is finite and non-zero, and $H(1) \neq 0$. Then, the amplitude and phase of the LTI system $H(z)$ can be recovered, up to a sign and linear phase ambiguities, from one of the kp^{th}-order($k > 2$, $kp < \alpha$) FOC of the output.

Theorem 2 can be proven using the similar method in the reference[19].

By using the similar method in reference[17], substituting (42) into (43) and let $k > 2$, yields

$$\begin{aligned} C_{ky}^{kP}(m_1, m_2, \cdots, m_{k-1}) &= cum_{ky}^{kP}[y(n), y(n+m_1), \cdots, y(n+m_{k-1})] \\ &= cum_{kx}^{kP}\left\{\sum_{i_1=0}^{\infty} h(n-i_0)e(i_0), \sum_{i_2=0}^{\infty} h(n+m_1-i_1)e(i_1), \cdots, \sum_{i_k=0}^{\infty} h(n+m_{k-1}-i_{k\text{-}1})e(i_{k\text{-}1})\right\} \\ &= \sum_{i_1=0}^{\infty} h(n-i_0) \sum_{i_2=0}^{\infty} h(n+m_1-i_1) \cdots \sum_{i_k=0}^{\infty} h(n+m_{k-1}-i_{k\text{-}1}) cum_{ke}^{kP}\left[e(i_0), e(i_1), \cdots, e(i_{k\text{-}1})\right] \end{aligned} \tag{44}$$

By assumption [AS2], it is obtained,

$$cum_{ke}^{kP}\left[e(i_0), e(i_1), \cdots, e(i_{k\text{-}1})\right] = \begin{cases} \zeta_{ke} \neq 0, & i_0 = i_1 = \cdots = i_{k-1} = 0 \\ 0 & \text{others} \end{cases} \tag{45}$$

Thus, (44) can be written as,

$$C_{ky}^{kP}(m_1, m_2, \cdots, m_{k-1}) = \zeta_{ke} \sum_{i=0}^{\infty} h(i) h(i+m_1) \cdots h(i+m_{k-1}) \tag{46}$$

Let $m_1 = m$, $m_2 = n$, $m_3 = \cdots = m_{k-1} = 0$ in (46), we have

$$\begin{aligned} C_{ky}^{kP}(m,n) &\equiv C_{ky}^{kP}(m, n, 0, \cdots, 0) \\ &= \zeta_{ke} \sum_{i=0}^{\infty} h^{k-2}(i) h(i+m) h(i+n) \end{aligned} \tag{47}$$

On the other hand, by the definition of the impulse response, we also have

$$\sum_{j=0}^{p} a(j) h(n-j) = \sum_{j=0}^{q} b(j) \delta(n-j) = b(n) \qquad (n = 1, \cdots, q) \tag{48}$$

Substituting (48) into (47), yields

$$\begin{aligned} \sum_{j=0}^{p} a(j) C_{ky}^{kP}(m-j, n) &= \zeta_{ke} \sum_{i=0}^{\infty} h^{k-2}(i) h(i+n) \sum_{j=0}^{p} a(j) h(i+m-j) \\ &= \zeta_{ke} \sum_{i=0}^{\infty} h^{k-2}(i) h(i+n) b(i+m) \end{aligned} \tag{49}$$

Considering $b(i+m) \equiv 0$ for $m > q$ and $i \geq 0$, one has

$$\sum_{i=0}^{p} a(i) C_{ky}^{kP}(m-i, n) = 0 \tag{50}$$

where $m > q$ and n are arbitrary integers. Equation (50) is the kp^{th}-order FOC-based Yuel-Walker function of the ARMA model.

C. Parameter estimation of ARMA model

The formula (46), (47), (48), (49) and (50) are the basic equations for deriving the FOC-based ARMA parameter estimation method. It is noteworthy that the mathematical expression form of the above equations is similar to that of the HOC-based ARMA parameter estimation method. Therefore, we can derivative the FOC-based ARMA parameter estimation method by using the similar method which based on HOC[17, 18].

1. FOC-based TLS-SVD method for AR parameter estimation of ARMA model

Let $m = q+1, \cdots, q+p$, $n = q-p, \cdots, q$ and $k = 3$ in (50), then

$$\begin{bmatrix} C_{ky}^{kp}(q+1-p,q-p) & C_{ky}^{kp}(q+2-p,q-p) & \cdots & C_{ky}^{kp}(q,q-p) \\ \vdots & \vdots & \vdots & \vdots \\ C_{ky}^{kp}(q+1-p,q) & C_{ky}^{kp}(q+2-p,q) & \cdots & C_{ky}^{kp}(q,q) \\ \vdots & \vdots & \vdots & \vdots \\ C_{ky}^{kp}(q,q-p) & C_{ky}^{kp}(q+1,q-p) & \cdots & C_{ky}^{kp}(q+p-1,q-p) \\ \vdots & \vdots & \vdots & \vdots \\ C_{ky}^{kP}(q,q) & C_{ky}^{kP}(q+1,q) & \cdots & C_{ky}^{kp}(q+p-1,q) \end{bmatrix} \quad (51)$$

$$\bullet \begin{bmatrix} a(p) \\ a(p-1) \\ \vdots \\ a(1) \end{bmatrix} = - \begin{bmatrix} C_{ky}^{kp}(q+1,q-p) \\ \vdots \\ C_{ky}^{kp}(q+1,q) \\ \vdots \\ C_{ky}^{kp}(q+p,q-p) \\ \vdots \\ C_{ky}^{kp}(q+p,q) \end{bmatrix}$$

or more compactly as

$$C_{ky}^{kp} A = -B \quad (52)$$

As equation (52) has a similar form to the extended Yuel-walker function in reference[18], with the similar proof method to that in[18], it can be proved that matrix C_{ky}^{kp} is of full rank.

It is noteworthy that in(52), because the fractional order operation is used, the elements in matrix C_{ky}^{kp} and vector b may be plural. Due to that the solution of the equation $A = \begin{bmatrix} a(p) & a(p-1) & \cdots & a(1) \end{bmatrix}^T$ must be real, by operating the real part of both sides of equation(52), that is, taking

$$\mathrm{Re}\left(C_{ky}^{kp}\right) A = -\mathrm{Re}(B) \quad (53)$$

the calculation of the equation will be simplified. Since the real part operation of the matrix does not affect the rank of the matrix, the matrix $\mathrm{Re}\left(C_{ky}^{kp}\right)$ is still full rank.

Let $p_e \ge p$, $q_e \ge q$, $q_e - p_e \ge q - p$, $N_1 \le q - p$ and $N_2 \ge q$, the following set of equations is constructed.

$$\mathrm{Re}\left\{ \begin{bmatrix} C_{ky}^{kp}(q_e+1,N_1) & C_{ky}^{kp}(q_e,N_1) & \cdots & C_{ky}^{kp}(q_e+1-p_e,N_1) \\ \vdots & \vdots & \vdots & \vdots \\ C_{ky}^{kp}(q_e+1,N_2) & C_{ky}^{kp}(q_e,N_2) & \cdots & C_{ky}^{kp}(q_e+1-p_e,N_2) \\ \vdots & \vdots & \vdots & \vdots \\ C_{ky}^{kp}(q_e+p_e,N_1) & C_{ky}^{kp}(q_e+p_e-1,N_1) & \cdots & C_{ky}^{kp}(q_e,N_1) \\ \vdots & \vdots & \vdots & \vdots \\ C_{ky}^{kp}(q_e+p_e,N_2) & C_{ky}^{kp}(q_e+p_e-1,N_2) & \cdots & C_{ky}^{kP}(q_e,N_2) \end{bmatrix} \right\} \begin{bmatrix} 1 \\ a(1) \\ \vdots \\ a(p_e) \end{bmatrix} = 0 \quad (54)$$

or more compactly as

$$\mathrm{Re}\left[\boldsymbol{C}_e^{kp} \right] A_e = 0 \quad (55)$$

Take SVD operation of the augmented matrix $\mathrm{Re}\left\{ \boldsymbol{C}_e^{kp} \right\}$, and arrange the singular values in descending order, i.e. $\sigma_1^2 \ge \sigma_2^2 \ge \cdots \ge \sigma_{p_e}^2$, then

$$\mathrm{Re}\left\{ \boldsymbol{C}_e^{kp} \right\} = U \Sigma V^* \quad (56)$$

where, $\Sigma = diag\left[\sigma_1^2, \sigma_2^2, \cdots, \sigma_{p_e}^2 \right]$, U and V are left and right singular vectors of matrix $\mathrm{Re}\left\{ {}^C\boldsymbol{C}_e^{kp} \right\}$, respectively.

Because the rank of matrix $\mathrm{Re}\left\{ {}^C\boldsymbol{C}_e^{kp} \right\}$ is p , take the optimal linear approximation for matrix $\mathrm{Re}\left\{ {}^C\boldsymbol{C}_e^{kp} \right\}$ by the rank of p , i.e. let

$$S^{(p)}=\sum_{j=1}^{p}\sum_{i=1}^{p_e+1-p}\sigma_j^2 v_j^i (v_j^i)^* \tag{57}$$

where, $v_j^i=\left[v(i,k),v(i+1,k),\cdots,v(i+p,k)\right]^T$, σ_j^2 is the j th singular value of the singular value matrix Σ .

Calculate the inverse of matrix $S^{(p)}$. Considering the first component of vector A_e is 1, let $\hat{A}=\left[\hat{a}(1),\hat{a}(2),\cdots,\hat{a}(p)\right]^T$, the estimation of the unknown parameter vector $\hat{A}$ can be calculate by

$$\hat{a}(i)=S^{-(p)}(i+1,1)/S^{-(p)}(1,1) \qquad (i=1,2,\cdots,p) \tag{58}$$

2. FOC-based Q-slice method for MA parameters estimation of ARMA model

By using the AR parameter estimation $\hat{A}$ obtained previously construct the fitting error function $f_k(m,n)$,

$$f_k(m,n)\triangleq\sum_{j=0}^{p}\hat{a}(j)C_{ky}^{kP}(m-j,n) \tag{59}$$

By equation (49), then

$$f_k(m,n)=\zeta_{ke}\sum_{i=0}^{\infty}h^{k-2}(i)h(i+n)b(i+m) \tag{60}$$

As shown in (60), due to that ζ_{ke} , $\{b(i)\}$, and $\{h(i)\}$ are all real numbers, the fitting error function $f_k(m,n)$ is a real number sequence. Thus, equation (59) can also be simplified equivalently to

$$f_k(m,n)\triangleq\sum_{j=0}^{p}\hat{a}(j)\operatorname{Re}\left[C_{ky}^{kP}(m-j,n)\right] \tag{61}$$

In equation (61), let $m=q$, then

$$f_k(q,n)=\sum_{j=0}^{p}\hat{a}(j)\operatorname{Re}\left[C_{ky}^{kP}(q-j,n)\right]=\zeta_{ke}h(n)b(q) \tag{62}$$

In equation (62), let $n=0$, and notice that $h(0)=1$, then

$$f_k(q,0)=\sum_{j=0}^{p}\hat{a}(j)\operatorname{Re}\left[{}^{C}C_{ky}^{kP}(q-j,0)\right]=\zeta_{ke}b(q) \tag{63}$$

By the model assumptions, $b(q)\neq 0$ and divide (62) by (63), yields

$$h(n)=\frac{f_k(q,n)}{f_k(q,0)}=\frac{\sum_{j=0}^{p}\hat{a}(j)\operatorname{Re}\left[C_{ky}^{kP}(q-j,n)\right]}{\sum_{j=0}^{p}\hat{a}(j)\operatorname{Re}\left[C_{ky}^{kP}(q-j,0)\right]} \qquad n=0,1,\cdots \tag{64}$$

After $h(i)$ is obtained, MA parameters can be calculated by equation (48), i.e.

$$b(n)=\sum_{i=0}^{n}\hat{a}(i)h(n-i) \quad (n=1,\cdots,q) \tag{65}$$

where, $h(\tau)=0$ ($\tau<0$) .

Remark 5

1. By comparing the traditional HOC-based SVD-TLS and Q-slice with the FOC-based method proposed in this paper it is shown that the difference between them is only the difference of the statistical operators used. Because the FOC and HOC have similar properties, the discussion on the performance of SVD-TLS and Q-slice algorithms in [17, 18] is also suitable to this algorithm. And, the modified Q-slice algorithm proposed in [18] can also be transformed into FOC-based Q-slice algorithm.

2. In the above method, the order of ARMA is assumed to be known. For the unknown order of ARMA model, as the same reason above, the order determination method of ARMA model proposed in[20] can also be transformed into FOC-based method.

3. The derivation process of FOC-based parameter estimation method for ARMA model shows that it is feasible to transform HOC-based signal processing method to FOC-based method. This method not only simplifies the research process of FOC-based signal processing algorithms, also effectively improve the robustness of the original algorithms by using this modification.

D. The consistency of ARMA model parameters estimation

In this sub-section, we discuss the consistency of FOC-based ARMA model parameter estimation proposed in this paper. By the parameter estimation equations of ARMA model presented above, it is obvious that the parameter estimates $\hat{a}(i)$ and $\hat{b}(i)$ are measurable functions of $C_X^{3p}(\cdot)$. Therefore, the ARMA parameter estimation of FOC-based method will themselves be consistent, since $\hat{C}_X^{3p}(\square)$ is measurable function of consistent estimators.

Consistency of the 3^{rd} -order HOC estimation has been studied in reference[19], and a similar work can also be seen in reference [21]. In this paper, we will further discuss the consistency of $3p^{th}$ -order FOC estimators on this basis.

Theorem 3

If the input r.p. $\{e(k)\}$ of LTI system satisfies

[AS6] $mon^p\left[e^p(k)\right]=0$,

[AS7] $C_e^{2p}(\square), C_e^{3p}(\square), C_e^{4p}(\square), C_e^{6p}(\square)\in L_1$

[AS8] $\sum_{k=0}^{\infty}|h(k)|<\infty$,

then, the $3p^{th}$ -order FOC estimators $\hat{C}_X^{3p}(\tau_1,\tau_2)$ of the system output r.p. $\{x(k)\}$ converge in probability to $C_X^{3p}(\tau_1,\tau_2)$ as $N\to\infty$.

The proof is given in **Appendix 4**.

In the ARMA parameter estimation method proposed in this paper, by the assumption [AS1] to [AS5], $C_y^{3p}(\tau_1,\tau_2)=C_X^{3p}(\tau_1,\tau_2)$ (see equation (43)), and Theorem 3, $\hat{C}_y^{3p}(\tau_1,\tau_2)$ are the consistent estimators of $C_X^{3p}(\tau_1,\tau_2)$. Further, by the estimation equations of $\hat{a}_N(k)$ and $\hat{b}_N(k)$, both $\hat{a}_N(k)$ and $\hat{b}_N(k)$ are measurable functions of $\hat{C}_y^{3p}(\tau_1,\tau_2)$, $\hat{a}_N(k)$ and $\hat{b}_N(k)$ are the consistent estimators.

VII. Simulation Examples

In this section, we validate the model parameter estimation performance of the new FOC based TLS-SVD and Q-slice method by simulation experiments. In reference [17], various common signal models in practical application were discussed. And it focused on one of the typical non minimum phase ARMA model with all pass factors, simulated and compared the performance of a variety of 3th-order HOC-based model parameter estimation methods in colored Gaussian noise. This model is also used in many literatures such as reference [18]. In order to facilitate comparison, this model is used in this section.

The simulation model is as follow[17]:

$$x(k)-1.30x(k-1)+1.05x(k-2)-0.325x(k-3)=e(k)-2.95e(k-1)+1.90e(k-2) \tag{66}$$

The noisy observation output of system is

$$y(k)=x(k)+\omega(k) \tag{67}$$

In this model, the two zeros are 2.0 and 0.95, and the three poles are $0.4\pm0.7j$ and 0.5 respectively. The model input $e(n)$ is drawn from an i.i.d., zero-M^p mean, single-sided exponential distribution with variance, $\sigma_e^2=1$. Obviously, because $e(n)$ is asymmetric distribution, $C_{3e}^{3p}(m_1,m_2)\neq 0$ 。 $\omega(k)$ is zero-M^p mean colored $S\alpha S$ and/or Gaussian noise. The white $S\alpha S$ noise is generated as that in the reference[22], where the characteristic index of the $S\alpha S$ noise are set as $\alpha=1.6$. The colored $S\alpha S$ noise is generated by the $S\alpha S$ noise through a band-pass filter. And, the Signal to Noise Ratio (SNR) is calculated by the Generalized SNR (GSNR) [2], which is defined as

$$GSNR=10\log\left(\frac{1}{\gamma N}\sum_{k=1}^{N}|s(k)|^2\right) \tag{68}$$

All the noisy observation sequence, colored $S\alpha S$ noise and colored Gaussian noise are processed by M^p mean value subtraction. The additive noise $\omega(k)$ is chosen as a colored $S\alpha S$ noise, a colored Gaussian noise, and a mixed noise of colored $S\alpha S$ and colored Gaussian noise in the following examples respectively. In all examples, assume that the order p , q of the model and the characteristic index α of $S\alpha S$ noise are known.

For the parameter estimation method for AR and MA part, the $3p^{th}$ -order FOC based SVD-TLS and Q-slice method proposed in this paper are used respectively. The simulation sequence is generated by the above method with $N=4096$.

A. *Example 1*

Take 100 times Monte Carlo experiments in colored α noise, colored Gaussian noise, colored α and Gaussian noise respectively. The arithmetic mean and their associated standard deviations (in the bracket) for the estimates of the parameters during GSNR (SNR) changes from 0dB to 20dB in 10dB interval are shown in Table II, Table III and Table IV respectively. Due to that the HOC can only be used in parameter estimation in colored Gaussian noise, for convenience of comparison, the parameter estimation results of HOC-based SVD-TLS and Q-slice method methods are also shown in Table III (shown in shadows).

TABLE II

ARITHMETIC MEAN AND STANDARD DEVIATIONS OF ESTIMATES OF ARMA PARAMETERS IN COLORED α NOISE

GSNR \ True value	$a(1)$= -1.30	$a(2)$=1.05	$a(3)$= -0.325	$b(1)$= -2.95	$b(2)$= 1.90
0dB	-1.8475 (0.4697)	1.7651 (0.4071)	0.4887 (0.3103)	1.0122 (3.9895)	4.5241 (4.0247)
10dB	-1.1610 (0.1796)	0.7342 (0.1823)	-0.4225 (0.1592)	-3.8543 (3.6429)	2.6379 (3.4547)
20dB	-1.1738 (0.1813)	0.7871 (0.1593)	-0.4186 (0.1523)	-3.4218 (3.2127)	2.5272 (3.2428)

TABLE III

ARITHMETIC MEAN AND STANDARD DEVIATIONS OF ESTIMATES OF ARMA PARAMETERS IN COLORED GAUSSIAN NOISE

GSNR \ True value	$a(1)$= -1.30	$a(2)$=1.05	$a(3)$= -0.325	$b(1)$= -2.95	$b(2)$= 1.90
0dB	-1.8475 (0.3697)	1.7651 (0.3071)	0.4887 (0.3203)	1.0122 (3.7467)	4.5241 (3.8775)
	-1.7463 (0.3234)	1.4352 (0.2741)	0.1324 (0.3004)	0.9858 (3.5421)	4.2543 (3.7465)
10dB	-1.1720 (0.1707)	0.8025 (0.1564)	-0.4125 (0.1581)	-3.4125 (3.1213)	2.5312 (2.9862)
	-1.1942 (0.1476)	0.8731 (0.1516)	-0.2882 (0.1402)	-3.2475 (2.9167)	2.4547 (2.8293)
20dB	-1.1843 (0.1635)	0.8231 (0.1486)	-0.4053 (0.1469)	-3.2639 (2.8201)	2.3401 (2.8846)
	-1.2065 (0.1392)	0.9113 (0.1372)	-0.3067 (0.1193)	-3.1326 (2.7681)	2.2846 (2.7384)

TABLE IV

ARITHMETIC MEAN AND STANDARD DEVIATIONS OF ESTIMATES OF ARMA PARAMETERS IN COLORED α NOISE AND GAUSSIAN NOISE

GSNR \ True value	$a(1)$= -1.30	$a(2)$=1.05	$a(3)$= -0.325	$b(1)$= -2.95	$b(2)$= 1.90
0dB	-1.9324 (0.4965)	1.8516 (0.4213)	0.5787 (0.3537)	1.1323 (4.0162)	4.7562 (4.2237)
10dB	-1.1471 (0.1863)	0.6972 (0.1842)	-0.4268 (0.1611)	-3.9753 (3.7265)	2.7069 (3.5359)
20dB	-1.1589 (0.1825)	0.7271 (0.1608)	-0.4207 (0.1562)	-3.5468 (3.5253)	2.6842 (3.4329)

As shown from Table II to Table IV, the FOC-based method can well estimate the parameters of non-minimum phase ARMA model whether in colored α noise, colored Gaussian noise, or the mixed noise of colored α and Gaussian noise. This fully reflects the robustness of the proposed FOC method to suppress the colored α noise and Gaussian noise. In addition, with the increase of GSNR, the standard deviations decrease gradually, which is obviously reasonable. Moreover, in the pure Gaussian noise, the accuracy of HOC-based method is slightly higher than FOC-based method. This is because the estimation of HOC is based on integer exponential operation, while the estimation of FOC is based on fractional exponential operation. In this regard, we believe that this is the price that must be paid for the robustness of the algorithm.

B. *Example 2*

In order to better show the estimation performance of the algorithm proposed in this paper under different p and GSNR selections, we change p from 0.1 to 1 in 0.1 interval, i.e. kp from 0.3 to 3, and change GSNR from 0dB to 20dB in 1dB interval, and take 100 times Monte Carlo experiments in colored α noise at every point. The graph plotting the standard deviations of $a(1)$ to $b(2)$ change with GSNR and kp is shown in Figure I.

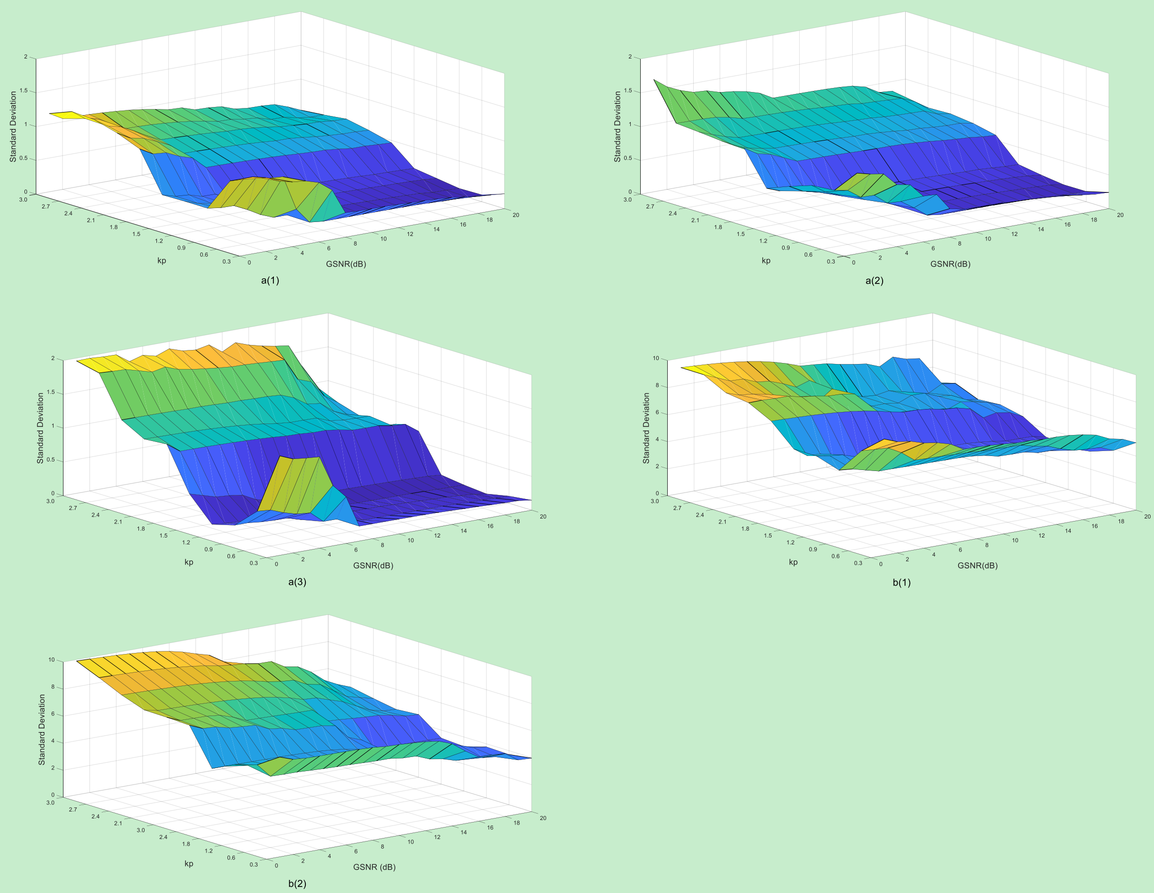

FIGURE I STANDARD DEVIATIONS CURVE WITH P AND GSNR

The results clearly show that when $kp < 0.6$, the standard deviations of each parameter estimation increases rapidly, which shows that kp should not be chosen too small. And when $0.6 \le kp < \alpha$, the standard deviations are small and change gently, so the kp in this range should be chosen. When $kp > \alpha$, the standard deviations increase rapidly, which is consistent with the conclusion in Theorem 1. In addition, it can be seen from Figure 1 that when $GSNR < 5dB$, the standard deviations of each parameter estimation increases significantly. Later, with the GSNR increasing, the standard deviations gradually decrease. When $GSNR > 10dB$, the standard deviations basically stabilize, which is obviously reasonable.

VIII. CONCLUSIONS

In this paper, a new concept of the FOC is proposed and its properties are derived. A parameter estimation algorithm for the non-minimum-phase ARMA processes is developed based on the proposed FOC. The main conclusions from this work are drawn as follows.

1. The proposed FOM and FOC filled the gap between the integer orders of the HOM and HOC, and extended the order definitions from positive integer to the complete positive real field. This further improves the signal processing theory.

2. The most significant difference between FOC and HOC is that FOC can be used for solving the signal processing problems in α noise, but HOC cannot. Comparing with FLOS, FOC possesses the linear and semi-invariant properties, but FLOS does not. Moreover, for α and Gaussian noise, whether they are colored or white, their FOC is zero, but their FLOM is not. Thus, FOC can be effectively used for signal processing in colored α and/or Gaussian noise, but FLOS can hardly be used.

3. In addition to the above main differences with HOC, FOC has similar properties to HOC, which determine that most existing HOC-based signal processing algorithms can be transformed into FOC-based algorithms. This will significantly simplify the research of FOC-based signal processing algorithms. The application example of FOC given in section VI fully demonstrates this.

APPENDICES

Appendix 1: Conversion between the FOM and FOC

By the definition of k -dimension fractional joint Moment-generating function and Cumulant-generating function, then

$$\Phi_p\left(u_1,u_2,\cdots,u_k\right)=e^{\Psi_p\left(u_1,\cdots,u_k\right)} \tag{A1}$$

Take multi-dimensional Taylor series expansion [16] at the left side of equation (A1), then

$$\begin{aligned}\Phi_p\left(u_1,u_2,\cdots,u_k\right)&=\sum_{q=0}^{\infty}\frac{1}{\Gamma(1+qp)}\left\{\left[u_1\frac{\partial}{\partial u_1}+\cdots+u_k\frac{\partial}{\partial u_k}\right]^{qp}\Phi_p\left(u_1,u_2,\cdots,u_k\right)\right\}_{u_1=\cdots=u_k=0}\\&=E_p\left(u_1^p\frac{\partial^p}{\partial u_1^p}\right)E_p\left(u_2^p\frac{\partial^p}{\partial u_2^p}\right)\cdots E_p\left(u_k^p\frac{\partial^p}{\partial u_k^p}\right)\Phi_p\left(u_1,u_2,\cdots,u_k\right)_{u_1=\cdots=u_k=0}\end{aligned} \tag{A2}$$

Similarly, take multi-dimensional Taylor series expansion [16] at the right side of equation (A1), then

$$\begin{aligned}e^{\Psi_p\left(u_1,\cdots,u_k\right)}&=\sum_{q=0}^{\infty}\frac{1}{\Gamma(1+qp)}\left\{\left[u_1\frac{\partial}{\partial u_1}+\cdots+u_k\frac{\partial}{\partial u_k}\right]^{qp}e^{\Psi_p\left(u_1,\cdots,u_k\right)}\right\}_{u_1=\cdots=u_k=0}\\&=E_p\left(u_1^p\frac{\partial^p}{\partial u_1^p}\right)E_p\left(u_2^p\frac{\partial^p}{\partial u_2^p}\right)\cdots E_p\left(u_k^p\frac{\partial^p}{\partial u_k^p}\right)e^{\Psi_p\left(u_1,\cdots,u_k\right)}\Big|_{u_1=\cdots=u_k=0}\end{aligned} \tag{A3}$$

By comparing the coefficients of $u_1^p u_2^p\cdots u_k^p$ in equations (A2) and (A3), then

$$\frac{{}_0^C\partial^{kp}\Phi_p\left(u_1,\cdots,u_k\right)}{\partial u_1^p\cdots\partial u_k^p}\Big|_{u_1=\cdots=u_k=0}=\frac{{}_0^C\partial^{kp}e^{\Psi_p\left(u_1,\cdots,u_k\right)}}{\partial u_1^p\cdots\partial u_k^p}\Big|_{u_1=\cdots=u_k=0}\qquad,k=1,2,\cdots \tag{A4}$$

By equation (A4) and the definitions of the FOM and FOC of multiple random variables, the Fractional Cumulant to Moment (FC-M) conversion formula as shown in equation (20) can be written as,

$$mom^{kp}(I)=\sum_{\bigcup_{m=1}^{q}I_m=I}\prod_{m=1}^{q}cum^{kp}(I_m) \tag{A5}$$

By the above FC-M formula and simple algebraic operation, we can directly obtain the conversion function of Fractional order Moment to Cumulant (FM-C) conversion formula as shown in equation (21),

$$cum^{kp}(I)=\sum_{\bigcup_{m=1}^{q}I_m=I}(-1)^{(q-1)}(q-1)!\prod_{m=1}^{q}mom^{kp}(I_m) \tag{A6}$$

Appendix 2: Proof of Property 1 to Property 5

Proof of Property 1:

By the properties of the moment and mathematical expectation, proof of (27) is straightforward.

By the definition of the FOC and the operation properties of the sequential fractional order derivative [15], then

$$cum^{kp}\left[a_1x_1^p,\cdots,a_kx_k^p\right]=\frac{{}_0^C\partial^{kp}\ln E\left\{E_p(a_1(ju_1x_1)^p)E_p(a_2(ju_2x_2)^p)\cdots E_p(a_k(ju_kx_k)^p)\right\}}{\partial u_1^p\cdots\partial u_k^p}\Big|_{u_1=\cdots u_k=0}$$
$$=\frac{\dfrac{{}_0^C\partial^{kp}E\left\{E_p(a_1(ju_1x_1)^p)E_p(a_2(ju_2x_2)^p)\cdots E_p(a_k(ju_kx_k)^p)\right\}}{\partial u_1^p\cdots\partial u_k^p}}{E\left\{E_p(a_1(ju_1x_1)^p)E_p(a_2(ju_2x_2)^p)\cdots E_p(a_k(ju_kx_k)^p)\right\}}\Big|_{u_1=\cdots u_k=0} \tag{B1}$$
$$=a_1\cdots a_k\frac{{}_0^C\partial^{kp}\ln E\left\{E_p((ju_1x_1)^p)E_p((ju_2x_2)^p)\cdots E_p((ju_kx_k)^p)\right\}}{\partial u_1^p\cdots\partial u_k^p}\Big|_{u_1=\cdots u_k=0}$$
$$=a_1\cdots a_k cum^{kp}\left[x_1^p,\cdots,x_k^p\right]$$

Equation (28) is then proved.

QED

Proof of Property 2:

1. By the properties of the moment and mathematical expectation, equation (29) can be easily proved.
2. Let $\boldsymbol{X}=\left[x_1,\cdots,x_k\right]$, then the cumulant-generating function of $\boldsymbol{X}$ can be written as

$$\Psi_X(u_1,\cdots u_k)=\ln E\left\{E_p((ju_1x_1)^p)E_p((ju_2x_2)^p)\cdots E_p((ju_kx_k)^p)\right\}$$
$$=\ln E\left\{E_p((ju_{i_1}x_{i_1})^p)E_p((ju_{i_2}x_{i_2})^p)\cdots E_p((ju_{i_k}x_{i_k})^p)\right\} \tag{B2}$$

where, $i_1,\cdots,i_k$ is a permutation of $1,2,\cdots,k$. Thus, by the definition of the FOC, equation (30) is proved.

QED

Proof of Property 3:

By Property 2, the FOC is symmetric in their arguments. So, without loss of generality assume that $(x_1,\cdots,x_i)$ are independent of $(x_{i+1},\cdots,x_k)$, then

$$\Psi_X(u_1,\cdots,u_k)=\ln E\left\{E_p((ju_1x_1)^p)E_p((ju_2x_2)^p)\cdots E_p((ju_kx_k)^p)\right\}$$
$$=\ln\left\{E\left\{E_p((ju_1x_1)^p)\cdots E_p((ju_ix_i)^p)\right\}E\left\{E_p((ju_{i+1}x_1)^p)\cdots E_p((ju_kx_k)^p)\right\}\right\}$$
$$=\ln E\left\{E_p((ju_1x_1)^p)\cdots E_p((ju_ix_i)^p)\right\}+\ln E\left\{E_p((ju_{i+1}x_1)^p)\cdots E_p((ju_kx_k)^p)\right\} \tag{B3}$$
$$=\Psi_X(u_1,\cdots,u_i)+\Psi_X(u_{i+1},\cdots,u_k)$$

Thus,

$$\frac{{}_0^C\partial^{kp}\Psi_X(u_1,\cdots,u_k)}{\partial u_1^p\cdots\partial u_k^p}=\frac{{}_0^C\partial^{kp}\Psi_X(u_1,\cdots,u_i)}{\partial u_1^p\cdots\partial u_k^p}+\frac{{}_0^C\partial^{kp}\Psi_X(u_{i+1},\cdots,u_k)}{\partial u_1^p\cdots\partial u_k^p}=0 \tag{B4}$$

Equation (31) is then proved.

But,

$$\Phi_X(u_1,\cdots,u_k)=\Phi_X(u_1,\cdots,u_i)\Phi_X(u_{i+1},\cdots,u_k) \tag{B5}$$

Then

$$\frac{{}_0^C\partial^{kp}\Phi_X(u_1,\cdots,u_k)}{\partial u_1^p\cdots\partial u_k^p}=\frac{{}_0^C\partial^{kp}\left[\Phi_X(u_1,\cdots,u_i)\Phi_X(u_{i+1},\cdots,u_k)\right]}{\partial u_1^p\cdots\partial u_k^p} \tag{B6}$$

Due to that $\Phi_X(u_1,\cdots,u_i)\Phi_X(u_{i+1},\cdots,u_k)$ contains the variables $u_1,\cdots,u_k$, the fractional partial derivatives of (B6) are then nonzero generally.

QED

Proof of Property 4:

1. Let $\boldsymbol{X}=\left[x_1,\cdots,x_k\right],\boldsymbol{Y}=\left[y_1,\cdots,y_k\right]$, $\boldsymbol{Z}=\boldsymbol{X}+\boldsymbol{Y}=\left[x_1+y_1,\cdots,x_k+y_k\right]$.

By using the independence of $\boldsymbol{X}$ and $\boldsymbol{Y}$, and the properties of Mittag-Leffler function[15], the cumulant-generating function of $\boldsymbol{Z}$ can be written as follows:

$$\begin{aligned}\Psi_Z(u_1,\cdots,u_k) &= \ln E\left\{E_p((ju_1(x_1+y_1))^p)\cdots E_p((ju_k(x_k+y_k))^p)\right\} \\ &= \ln E\left\{E_p((ju_1x_1)^p)E_p((ju_1y_1)^p)\cdots E_p((ju_kx_k)^p)E_p((ju_ky_k)^p)\right\} \\ &= \ln E\left\{E_p((ju_1x_1)^p)\cdots E_p((ju_kx_k)^p)\right\} + \ln E\left\{E_p((ju_1y_1)^p)\cdots E_p((ju_ky_k)^p)\right\} \\ &= \Psi_X(u_1,\cdots,u_k) + \Psi_Y(u_1,\cdots,u_k)\end{aligned} \tag{B7}$$

By the definition of the FOC and (B7), equation (33) can be proved.

2. By the properties of the moment and mathematical expectation, (34) can be proved.

QED

Proof of Property 5:

Let $\boldsymbol{X} = [x_1, x_2, \cdots, x_k]$, $\boldsymbol{Y} = [x_1 + \text{a}, y_2, \cdots, y_k]$, then

$$\begin{aligned}\Psi_Y(u_1,u_2,\cdots,u_k) &= \ln E\left\{E_p((ju_1(x_1+a))^p)\cdots E_p((ju_kx_k)^p)\right\} \\ &= \ln E\left\{E_p((ju_1x_1)^p)\cdots E_p((ju_kx_k)^p)\right\} + \ln E\left\{E_p\left[(ju_1a)^p\right]\right\}\end{aligned} \tag{B8}$$

Furthermore, while $k \geq 2$

$$\begin{aligned}& cum^{kp}\left[(x_1+a)^p, x_2^p, \cdots, x_k^p\right] \\ &= j^{-kp} \frac{{}_0^C\partial^{kp}\Psi_y(u_1,u_2,\cdots,u_k)}{\partial u_1^p\cdots\partial u_k^p}\Bigg|_{u_1=u_2=\cdots=u_k=0} \\ &= j^{-kp}\left\{\frac{{}_0^C\partial^{kp}\left\{\ln E\left\{E_p((ju_1x_1)^p)\cdots E_p((ju_kx_k)^p)\right\}\right\}}{\partial u_1^p\cdots\partial u_k^p} + \frac{{}_0^C\partial^{kp}\left\{\ln E\left\{E_p\left[(ju_1a)^p\right]\right\}\right\}}{\partial u_1^p\cdots\partial u_k^p}\right\}\Bigg|_{u_1=u_2=\cdots=u_k=0} \\ &= cum^{kp}\left[x_1^p, x_2^p, \cdots, x_k^p\right]\end{aligned} \tag{B9}$$

Thus, equation (35) is proved.

But, due to

$$mom^{kp}\left[(x_1+a)^p, x_2^p, \cdots, x_k^p\right] = E\left[(x_1+a)^p x_2^p \cdots x_k^p\right] \tag{B10}$$

It is obvious, because p is fraction, while $a \neq 0$, then

$$mom^{kp}\left[(x_1+a)^p, x_2^p, \cdots, x_k^p\right] \neq mom^{kp}\left[x_1^p, x_2^p \cdots, x_k^p\right] \tag{B11}$$

Thus, the equation (36) is proved.

QED

Appendix 3: Proof of Theorem 1

Proof:

By equation (37), the cumulant-generating function of the $S\alpha S$ -stable distribution can be written as

$$\begin{aligned}\Psi_e(u) &= \ln\left[\Phi_e(u)\right] \\ &= jau - \gamma\,|\,u\,|^\alpha\end{aligned} \tag{C1}$$

Due to the integral domain of the fractional derivative operator ${}_0^CD^p$ used in this paper is $[0,u]$, the kp^{th} -order FOC of random variable X ($X \sim S_\alpha(\gamma,\beta,a)$) is

$$C^{kp} = j^{-kp}\Psi_e^{(kp)}(0) = j^{-kp}\,{}_0^CD^{kp}\left\{jau - \gamma u^\alpha\right\}\Big|_{u=0} = j^{-kp}\,{}_0^CD^{kp}\left\{jau\right\} + j^{-kp}\,{}_0^CD^{kp}\left\{-\gamma u^\alpha\right\} \tag{C2}$$

While $0 < kp < 1$, the first item of equation (C2) is

$$j^{-kp}\,{}_0^CD^{kp}\left\{jau\right\}\Big|_{u=0} = j^{-kp}\frac{ja\Gamma(2)}{\Gamma(2-kp)}(u^{1-kp})\Big|_{u=0} = 0 \tag{C3}$$

While $kp > 1$, by the properties of Caputo fractional derivative, the first item of equation (C2) is

$$j^{-kp}\,{}_0^CD^{kp}\left\{jau\right\}\Big|_{u=0} = 0 \tag{C4}$$

While $kp = 1$, the first item of equation (C2) is

$$j^{-1}D\left\{jau\right\}\Big|_{u=0} = a \tag{C5}$$

When $\alpha - kp$ is not an integer, the second item of equation (C2) can be written as,

$$\begin{aligned} & j^{-kp}\ {}_0^C D^{kp}\left\{-\gamma u^{\alpha}\right\}_{u=0} \\ &= j^{-kp}\frac{-\gamma\Gamma(\alpha+1)}{\Gamma(\alpha+1-kp)}(u)^{\alpha-kp}\Big|_{u=0} \\ &= \begin{cases} 0 & \alpha > kp \\ \infty & \alpha < kp \end{cases} \end{aligned} \tag{C6}$$

While $kp=\alpha$,

$$j^{-kp}D^{kp}\left\{-\gamma u^{\alpha}\right\}\Big|_{u=0} = -\gamma\Gamma(1+\alpha)\, j^{-kp} \tag{C7}$$

When $kp-\alpha$ satisfies $1 < kp-\alpha \le m$ and is an integer, we have

$$j^{-kp}D^{kp}\left\{-\gamma\, |u|^{\alpha}\right\}_{u=0} = \frac{-\gamma\, j^{-kp}\Gamma(\alpha+1)}{\Gamma(m-p+\alpha+1)}D^{m}\left(u^{m-(kp-\alpha)}\right)\Big|_{u=0} = 0 \tag{C8}$$

Thus, consolidate (C3) to (C8), we have equation (38).

QED

Appendix 4: Proof of Theorem 3

Proof:

By the assumptions [AS6]to [AS8], the system output sequence $x(k)$ satisfies that

1. $C_x^{p}(\tau)=0$, i.e. the system output sequence $x(k)$ is zero M^{p} mean.
2. $C_x^{2p}(\cdot), C_x^{3p}(\cdot), C_x^{4p}(\cdot), C_x^{6p}(\cdot) \in L_1$, and they are absolutely sum-able.

So $C_x^{3p}(\tau_1,\tau_2)$ can be estimated by the following equation.

$$\hat{C}_x^{3p}(\tau_1,\tau_2) = \frac{1}{N}\sum_{k=0}^{N-\tau} x^{p}(k)x^{p}(k+\tau_1)x^{p}(k+\tau_2) \tag{D1}$$

where, $\tau = \mathrm{Max}(\tau_1,\tau_2)$

Due to that $x(k)$ is weak- stationary with arbitrary fractional order, take expectation for both sides of equation (D1), then

$$E\left\{\hat{C}_x^{3p}(\tau_1,\tau_2)\right\} = \frac{1}{N}\sum_{k=0}^{N-\tau} E\left[x^{p}(k)x^{p}(k+\tau_1)x^{p}(k+\tau_2)\right] = \frac{N-\tau+1}{N}C_x^{3p}(\tau_1,\tau_2) \tag{D2}$$

So that

$$E\left\{\hat{C}_x^{3p}(\tau_1,\tau_2)\right\} = \frac{N-\tau+1}{N}C_x^{3p}(\tau_1,\tau_2) \underset{N\to\infty}{\longrightarrow} C_x^{3p}(\tau_1,\tau_2) \tag{D3}$$

Thus, $\hat{C}_x^{3p}(\tau_1,\tau_2)$ is an asymptotically unbiased estimator of $C_x^{3p}(\tau_1,\tau_2)$.

Next, we calculate the estimated variance of $\hat{C}_x^{3p}(\tau_1,\tau_2)$.

$$\begin{aligned} Var\left\{\hat{C}_x^{3p}(\tau_1,\tau_2)\right\} &= E\left\{\left[\hat{C}_x^{3p}(\tau_1,\tau_2) - E\left[\hat{C}_x^{3p}(\tau_1,\tau_2)\right]\right]^2\right\} \\ &= E\left\{\left[\hat{C}_x^{3p}(\tau_1,\tau_2)\right]^2\right\} - \left\{E\left[\hat{C}_x^{3p}(\tau_1,\tau_2)\right]\right\}^2 \\ &= \frac{1}{N^2}\sum_{k_1=0}^{N-\tau}\sum_{k_2=0}^{N-\tau}\{E\left[x^{p}(k_1)x^{p}(k_1+\tau_1)x^{p}(k_1+\tau_2)x^{p}(k_2)x^{p}(k_2+\tau_1)x^{p}(k_2+\tau_2)\right] - \\ &\quad E\left[x^{p}(k_1)x^{p}(k_1+\tau_1)x^{p}(k_1+\tau_2)\right]E\left[x^{p}(k_2)x^{p}(k_2+\tau_1)x^{p}(k_2+\tau_2)\right]\} \\ &= \frac{1}{N^2}\sum_{k_1=0}^{N-\tau}\sum_{k_2=0}^{N-\tau} mon^{6p}\left[x^{p}(k_1),\ x^{p}(k_1+\tau_1),\ x^{p}(k_1+\tau_2),\ x^{p}(k_2),\ x^{p}(k_2+\tau_1),\ x^{p}(k_2+\tau_2)\right] - \\ &\quad mon^{3p}[x^{p}(k_1),\ x^{p}(k_1+\tau_1),\ x^{p}(k_1+\tau_2)]mon^{3p}[x^{p}(k_2),\ x^{p}(k_2+\tau_1),\ x^{p}(k_2+\tau_2)] \end{aligned} \tag{D4}$$

In equation (D4), due to that $x(k)$ is zero M^{p} mean, by the FC-M function (20) and let $k = k_1 - k_2$, change the double sum of equation (D4) into single sum, then

$$Var\left\{\hat{C}_x^{3p}(\tau_1,\tau_2)\right\}=\frac{1}{N^2}\sum_{|k|=0}^{N-\tau}(N-\tau-k)\left\{C_x^{6p}(\cdot)+\left[C_x^{3p}(\cdot)C_x^{3p}(\cdot)\right]_{10-1}+\left[C_x^{2p}(\cdot)C_x^{4p}(\cdot)\right]_{15}+\left[C_x^{2p}(\cdot)C_x^{2p}(\cdot)C_x^{2p}(\cdot)\right]_{15}\right\} \quad \text{(D5)}$$

where, the subscript j in $[\cdot]_j$ indicates that the item obtained by FC-M function has j terms, which are the same order of FOC but different variables in $[\cdot]$. $[\cdot]_j$ indicates the sum of these j terms.

Clearly, by $C_x^{3p}(\cdot)$, $C_x^{4p}(\cdot)$, $C_x^{6p}(\cdot)\in L_1$, and they are absolutely summable, the expressions in brackets of equation (D5) are absolutely summable. Therefore, the operation order of $\lim_{N\to\infty}$ and $\sum_{-\infty}^{\infty}\square$ can be interchanged. Further, because $\frac{N-\tau-k}{N^2}\underset{N\to\infty}{\to}0$, $\text{Var}\left\{\hat{C}_x^{3p}(\tau_1,\tau_2)\right\}\underset{N\to\infty}{\to}0$ in equation (D5).

Combining (D3) with (D5), and the fact that mean square convergence must be convergence in probability, implies that $\hat{C}_x^{3p}(\tau_1,\tau_2)$ converges in probability to $C_x^{3p}(\tau_1,\tau_2)$ as $N\to\infty$.

QED

References

[1] J. M. Mendel, "Use of higher-order statistics in signal processing and system theory: An update," in *Advanced Algorithms and Architectures for Signal Processing III*, 1988, pp. 126-144.

[2] X. Ma and C. L. Nikias, "Joint estimation of time delay and frequency delay in impulsive noise using fractional lower order statistics," *IEEE Transactions on Signal Processing*, vol. 44, pp. 2669-2687, 1996.

[3] C. L. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*: Wiley-Interscience, 1995.

[4] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: stable processes and their applications," *Proceedings of the IEEE*, vol. 81, pp. 986-1010, 1993.

[5] X. Ma and C. L. Nikias, "Parameter estimation and blind channel identification in impulsive signal environments," *IEEE Transactions on signal processing*, vol. 43, pp. 2884-2897, 1995.

[6] S. Cambnnis and G. Miller, "Linear problems in pth order and stable processes. SIAM 1," *Appl. Math*, vol. 41, pp. 43-69, 1981.

[7] J. G. Gonzalez, "Robust techniques for wireless communications in non-Gaussian environments," 1998.

[8] J. G. Gonzalez, J. L. Paredes, and G. R. Arce, "Zero-order statistics: A mathematical framework for the processing and characterization of very impulsive signals," *IEEE Transactions on Signal Processing*, vol. 54, pp. 3839-3851, 2006.

[9] K. Hari and V. Lalitha, "Subspace-based DOA estimation using fractional lower order statistics," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 2580-2583.

[10] G.-h. You, T.-s. Qiu, and A.-m. Song, "Novel direction findings for cyclostationary signals in impulsive noise environments," *Circuits, Systems, and Signal Processing*, vol. 32, pp. 2939-2956, 2013.

[11] J. Zhang and T. Qiu, "The Fractional Lower Order Moments Based ESPRIT Algorithm for Noncircular Signals in Impulsive Noise Environments," *Wireless Personal Communications*, vol. 96, pp. 1673-1690, 2017.

[12] X. zhou, "Study on Methods of DOA and Polarization Estimation Based on Electromagnetic Vector Sensor Array," Jilin University, 2006.

[13] X. Chang, "Study on Methods of Parameter Estimation of Polarized Array Signal in Alpha Stable Noise," Jilin University, 2011.

[14] I. Podlubny, *Fractional differential equations:an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*: Academic press, 1998.

[15] G. Jumarie, "Probability calculus of fractional order and fractional Taylor's series application to Fokker–Planck equation and information of non-random functions," *Chaos, Solitons & Fractals*, vol. 40, pp. 1428-1448, 2009.

[16] D. Usero, "Fractional Taylor Series for Caputo Fractional Derivatives. Construction of Numerical Schemes," 2008.

[17] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, pp. 278-305, 1991.

[18] A. Swami and J. M. Mendel, "ARMA parameter estimation using only output cumulants," in *Spectrum Estimation and Modeling, 1988., Fourth Annual ASSP Workshop on*, 1988, pp. 193-198.

[19] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using higher order statistics," *IEEE Transactions on Acoustics Speech & Signal Processing*, vol. 37, pp. 360-377, 1989.

[20] G. B. Giannakis and J. M. Mendel, "Cumulant-based order determination of non-Gaussian ARMA models," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 38, pp. 1411-1423, 1990.

[21] M. Rosenblatt and J. W. V. Ness, "Estimation of the Bispectrum," *Annals of Mathematical Statistics*, vol. 36, pp. 1120-1136, 1965.

[22] J. M. Chambers, C. L. Mallows, and B. W. Stuck, "A Method for Simulating Stable Random Variables," *Journal of the American Statistical Association*, vol. 71, pp. 340-344.

Yi-Ran Shi received the B.Eng in control engineering from Changchun University of Science and Technology, China, in 2007, M.Sc in control engineering from Lanzhou University of Technology, China, in 2010, and PhD in control engineering from Jilin University, China, in 2014.

He is currently an associate professor at Jilin University. His current research interests include signal processing and model predictive control. He is the author or co-author of more than 30 papers.

Ding-Li Yu received the B.Eng in electrical engineering from Harbin Civil Engineering College, China, in 1982, M.Sc in electrical engineering from Jilin University of Technology (JUT), China, in 1986, and PhD in control engineering from Coventry University, U.K., in 1995.

He was a lecturer at JUT from 1986 to 1990, a visiting researcher at University of Salford, U.K. from 1991 to 1992, a postdoc research fellow at Liverpool John Moores University (LJMU) from 1995 to 1998. Dr. Yu became a senior lecturer at LJMU in 1998 and was promoted to a reader in 2003, a professor in 2006. He is the author or co-author of one book, several book chapters and more than 200 papers. His current research interests include fault detection and fault tolerant control of bilinear and nonlinear systems, adaptive neural networks and their control applications, model predictive control for chemical processes and automotive engines, and real-time evaluations.

Dr. Yu serves as an associate editor for three international journals and the committee member of IFAC SAFEPROCESS. He has served as organizing committee member and IPC member for many international conferences.

Hong-Yan Shi was born in Jilin, China in 1971. She received her B.Eng in mechanical engineering from Shanghai Jiaotong University, China, M.Sc from Jilin University of Technology, China, and Ph.D from Jilin University, China, in 2002.

Now she is an associate professor of Shenzhen University. Her research interests include condition monitoring of microelectromechanical system. She is an author of 2 books and nearly 40 papers.

Yao-Wu Shi received the B.Eng, M.Sc and PhD degrees from Jilin University of Technology, China, in 1982, 1985 and 1994 respectively.

He was a lecturer at Jilin University of Technology from 1982 to 1987, from 1987 to 1993, and was promoted to an associate professor in 1998, a professor of Jilin University in 2003. He is the author or co-author of one book and several book chapters, more than 200 papers. His current research interests include modern cross spectrum analysis, system identification, DOA estimation, harmonic parameter estimation, and chaotic signal processing.