International Journal Of
Industrial Electronics Control and Optimization
Print ISSN: 3517-2645
Online ISSN: 3568-2645
Homepage: https://ieco.usb.ac.ir/

The Performance and Robustness of One-Parameter Fractional Linear Prediction

Seyed Mostafa Alaviyan Shahri[1] | Hamid Reza Abutalebi[2]

Department of Electrical Engineering, Yazd University, Yazd, Iran.[1,2]
Corresponding author's email: habutalebi@yazd.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 12-January-2024
Received in revised form: 18-March-2024
Accepted: 21-Apri- 2024
Published online: 22-sep-2024

Keywords:
Fractional Calculus,
Sampling Frequency,
Noise,
Reverberation,
Fractional Linear Prediction.

ABSTRACT

This paper deals with the problem of signal modeling using fractional-order linear prediction. In this research, we obtain the closed-form expression of the optimum sampling frequency of the One-Parameter Fractional-order Linear Prediction (OPFLP) and examine the performance when the fractional order (α) is in $0 \le \alpha \le 2$. Our investigation focuses on determining optimum α within the individual ranges of $0 \le \alpha \le 1$ and $1 < \alpha \le 2$ while considering various influential parameters, such as sampling frequency and environmental interferences. We initiate our study by examining the impact of the sampling frequency, a critical parameter that demands meticulous selection, on the optimal value of α. Simulation results demonstrate that if the sampling rate falls within five to six times the maximum frequency of the signal under scrutiny, the optimal range for α resides within $1 < \alpha \le 2$ .Conversely, when the sampling frequency exceeds six times the maximum signal frequency, the optimal α shifts to $0 \le \alpha \le 1$ .This observation underscores the crucial relationship between sampling frequency and the appropriate selection of the fractional order α for effective OPFLP performance. In the next step, we assess the robustness of OPFLP in handling challenging signal processing tasks, particularly in hands-free speech acquisition applications. We evaluate the model's performance and robustness against environmental interferences in three scenarios: noisy environments, reverberant environments, and noisy-reverberant settings. Simulation outcomes highlight OPFLP's superior robustness compared to second-order LP in handling environmental interferences. Furthermore, our investigations elucidate that noise exerts a more detrimental impact on OPFLP performance than reverberation, emphasizing the nuanced effects of these interferences on the model's efficacy.

I. Introduction

In various scientific and engineering domains, significant attention has been directed towards fractional calculus and its applications over recent decades [1]–[6]. Models incorporating fractional derivatives and integrals are referred to as fractional order systems, aiming to introduce memory into signal and system modeling [1], [4]. The fractional order (α) typically falls within the range $0 \le \alpha \le 2$. However, it is crucial to note that the system behavior, concerning stability and robustness, differs considerably between the ranges $0 \le \alpha \le 1$ and $1 < \alpha \le 2$ [7],[8].

Linear prediction (LP) is widely employed in the processing of various signals, including audio and speech [9], image [10], and biomedical signals. In the context of speech processing, LP has been applied in synthesis, coding, and compression [9].

Frequency-domain linear prediction (FDLP) has been a valuable tool for modeling and analyzing signals in the frequency domain, particularly for Automatic Speech

Recognition (ASR) applications [11]. It leverages the concept of linear prediction, where the current sample is predicted using past samples in a linear combination. In the context of FDLP, this prediction is performed in the frequency domain, focusing on the spectral characteristics of the signal.

Recently, a noteworthy advancement was presented in [12] with the introduction of Complex Frequency-Domain Linear Prediction (CFDLP). CFDLP offers a significant improvement over conventional FDLP by enabling straightforward interpretation of the complex cepstrum as temporal modulations within an all-pole model approximation of the speech signal's power spectrum.

Furthermore, CFDLP boasts a remarkable advantage in terms of computational efficiency. Compared to conventional FDLP, CFDLP achieves the same level of approximation with significantly faster processing speeds. This is a crucial advantage for real-time speech processing applications where computational efficiency is paramount. [13].

Recently, fractional-order linear prediction (FLP) has emerged in signal processing, termed as a generalized version of LP. However, its exploration in this field has been limited to a few notable papers [14]–[18].

In [14], the authors introduced One-Parameter FLP (OPFLP) within the range $0 \le \alpha \le 1$ and determined the optimal coefficients for the corresponding model. Their approach involved expanding the fractional Taylor series and truncating it to a single term to obtain the optimal OPFLP coefficients. Subsequently, they compared the performance of OPFLP against the first- and second-order LPs using standard sinusoidal waves and real signals, such as clean speech and Electrocardiogram (ECG) signals. In another study [15], the authors examined the performance of OPFLP ($0 \le \alpha \le 1$) by applying it to clean musical signals, such as piano and Western music. In [16], standard sinusoidal and damped sinusoidal waves were tackled to derive a closed-form solution for the optimal derivative order. Additionally, the authors proved that the optimal derivative order for OPFLP is inversely related to the number of memory samples utilized in the model.

The authors of [17] proposed a novel approach for bearing fault diagnosis in induction motors using OPFLP as a noise elimination technique. The presence of noise in the stator current, unrelated to bearing faults, can hinder accurate fault detection and lead to false alarms. This work proposes FLP as a method to address this challenge. FLP leverages the principles of linear prediction theory to identify and remove predictable noise components, leaving behind the unpredictable fault signature. A key advantage of this approach lies in its computational efficiency. The paper employs a limited memory FLP model, requiring optimization of only the fractional derivative order parameter after determining the optimal number of memory samples. This significantly reduces the computational complexity compared to other methods.

Navish et al. [18] delve into the challenge of noise contamination in biomedical signals acquired from various devices. These signals, often crucial for diagnoses, can be impacted by external factors like the environment, device functionality, and even clinician activity. To address this, the study investigates the effectiveness of two denoising techniques: stationary wavelet transform (SWT) and OPFLP, specifically applied to electromyography (EMG) signals. The primary objective is to compare the performance of these two methods in removing noise from EMG signals and ultimately improve the accuracy of diagnosis. The research employs five different methods to calculate the fractal dimension of the denoised signals: Higuchi, Katz, rescaled-range analysis, relative dispersion, and power spectral density. Based on the analysis, the study concludes that FLP outperforms SWT in terms of achieving a more reliable denoised signal, as evidenced by the fractal dimension analysis.

In the aforementioned studies, it has been assumed that the fractional derivative order is in the range of $0 \le \alpha \le 1$. While this assumption is satisfied in some cases, the optimum α can differ as the sampling frequency of the signal changes. According to the Nyquist theorem, the sampling frequency must exceed at least twice the maximum signal frequency. We demonstrate that if the sampling frequency equals two, three, or four times the maximum frequency, the optimal derivative order converges to $\alpha = 0$. However, as the sampling frequency reaches five or six times the maximum signal frequency, the optimal α shifts to $1 < \alpha \le 2$. Beyond six times the maximum frequency, the optimal derivative order resides within $0 \le \alpha \le 1$.

Another assumption when applying OPFLP to speech signals in the aforementioned research often relies on a single frame of voiced speech, leading to results that may not adequately represent practical situations such as hands-free speech signal acquisition. In practical situations, the processed speech signal contains voiced, unvoiced, and silent segments, commonly contaminated by noise and reverberation as two major environmental interferences. These environmental factors exert a substantial impact on speech signals, as extensively discussed in existing literature [19]–[23]. Reverberation occurs in enclosed spaces, manifesting as multiple delayed and attenuated speech reflections from enclosing walls and objects [23]. Meanwhile, diverse noise types, notably diffuse noise, pervade most environments [23].

Motivated by previous researches, our primary objective is to determine the optimal derivative order of OPFLP, assuming $0 \le \alpha \le 2$. Subsequently, we evaluate the performance of OPFLP in hands-free speech signal acquisition, assessing how noise and reverberation influence its efficacy. Additionally, we examine the robustness of the OPFLP model under three scenarios: 1) the presence of reverberation, 2) the presence of noise, and 3) the simultaneous presence of noise and

reverberation. To assess comparative performance, we consider the first- and second-order LPs as baseline methods, given their acceptable performance in various signal processing applications.

In summary, although [14] provides a closed-form expression for the optimal coefficient in $0 \le \alpha \le 1$, it is very conservative to investigate only one parameter, lacking theoretical justification. Furthermore, while acknowledging the critical role of sampling frequency in model performance, [14] does not offer a framework to determine its optimal value. Additionally, the [14] assumes the usage of a single frame of voiced speech for analysis might not be suitable for real-world situations that speech signal typically consists of diverse segments like voiced, unvoiced, and silent, and are often further contaminated by environmental noise. In this research, we overcome these limitations by deriving a closed-form expression for the optimal sampling frequency and determining the optimal derivative order in $0 \le \alpha \le 2$. Also the OPFLP performance and robustness are evaluated against environmental interferences.

This paper is organized as follows: Section II explains the fundamentals of the first- and second-order LPs alongside OPFLP. Section III details performance metrics employed for robustness and performance assessment and presents evaluation outcomes. Finally, Section IV concludes this research.

II. Preliminaries And Backgrounds

In this section, we briefly explain the first- and second-order LP as the baseline methods. Then, after presenting the basics of the fractional linear prediction, the OPFLP will be obtained.

Considering stationary signal $y(n)$, the linear prediction model estimates the current signal sample as a linear combination of q previous samples, i.e. [14]–[16]:

$$\hat{y}(n) = \sum_{i=1}^{q} a_i y(n-i) \qquad (1)$$

where $\hat{y}(n)$ is the estimated signal sample and q is called the prediction order. $q=1$ and $q=2$ results in the first- and the second-order LPs, respectively, as follows [14]–[16]:

$$q=1, \quad \hat{y}(n) = a.y(n-1) \qquad (2)$$

and

$$q=2, \hat{y}(n) = a_1.y(n-1) + a_2.y(n-2)a \qquad (3)$$

The MMSE framework can be used to obtain the optimum coefficient(s) in each case. In this regard, the error function is introduced as [14]–[16]:

$$e(n) = \hat{y}(n) - y(n) \qquad (4)$$

Then, the optimal coefficient(s) is(are) obtained as the solution of the following optimization problems:

$$a_{opt} = \arg\min_{a} \; J = E\left[e^2(n)\right] = \frac{R_y(1)}{R_y(0)}, \quad (q=1) \qquad (5)$$

and

$$a_{1_{opt}}, a_{2_{opt}} = \arg\min_{a_1, a_2} \; J = \mathrm{E}\left[e^2(n)\right] = \begin{cases} a_{1_{opt}} = \dfrac{R_y(1)R_y(0) - R_y(2)}{R_y^2(0) - R_y^2(1)} \\ a_{2_{opt}} = \dfrac{R_y(2)R_y(0) - R_y(1)}{R_y^2(0) - R_y^2(1)} \end{cases}, (q=2) \qquad (6)$$

where $R_y(k)$ denotes the autocorrelation function of y at lag k.

In the Fractional-order Linear Prediction (FLP) model, the current signal sample is estimated by the linear combination of q fractional-order terms [14]:

$$\hat{y}(t) = \sum_{i=1}^{q} a_i D_t^{\alpha_i} y(t) \qquad (7)$$

where $\hat{y}(n)$ is the estimated signal, a_i $(i=1,...,q)$ are the FLP coefficients, D^{α} is the fractional-order derivative operator, and α_i $(i=1,...,q)$ are the derivative orders.

The fractional derivative can be expanded using the Grundwall-Letnikov (G-L) definition. G-L is one of the basic definitions widely used in the numerical calculation of the fractional derivative of a function, and is expressed as [12], [13]:

$$D_t^{\alpha_i} y(t) = \lim_{h \to \infty} \frac{1}{h^{\alpha}} \sum_{j=0}^{K} (-1)^j \binom{\alpha}{j} y(t - jh) \qquad (8)$$

where K is the integer part of the fraction $\frac{(t-b)}{h}$, t and b are lower and upper limits of differentiation, respectively. h is sampling period ($h = 1/F_s$, F_s : sampling frequency) and $\alpha \in R$ is the arbitrary real order of the fractional derivative. By substituting $t = nh$ into (7) and (8), the discretized version of the equations are obtained as [14], [16]:

$$D^{\alpha} y(n) = \lim_{h \to \infty} \frac{1}{h^{\alpha}} \sum_{j=0}^{K} (-1)^j \binom{\alpha}{j} y(n-j) \qquad (9)$$

and

$$\hat{y}(n) = \sum_{i=1}^{q} a_i D_y^{\alpha_i} y(n) \qquad (10)$$

In (10), the system is known as “commensurate order” if the order of differentiation is integer multiple of a single base

order $\alpha_i = i\alpha$. In this case, (10) is re-written as [14], [16]:

$$\hat{y}(n) = \sum_{i=1}^{q} a_i D_y^{i\alpha} y(n) \quad (11)$$

By considering only the first term of (11), the first-order fractional linear prediction model is obtained as:

$$\hat{y}(n) = a.D^{\alpha} y(n-1) \quad (12)$$

where a and α are determined via the minimization of a cost function (e.g., mean squared error). In this regard, considering $m \in N$ as the upper limit of the summation in (9), we have:

$$D^{\alpha} y(n-1) = \frac{1}{h^{\alpha}} \sum_{j=0}^{m} (-1)^{j} \binom{\alpha}{j} y(n-1-j) \quad (13)$$

If $m = 1, 2,$ or 3 , the following three cases are respectively obtained from (12) and (13):

1. First-order OPFLP model with two samples memory:

$$\hat{y}(n) = \frac{a}{h^{\alpha}} \left(y(n-1) - \alpha y(n-2) \right) a \quad (14)$$

By defining the prediction error as $e(n) = \hat{y}(n) - y(n)$ and minimizing the mean squared error criteria, the optimal coefficient (a_{opt}) is obtained as follows:

$$a_{opt} = \arg\min_{a} \ \mathrm{J}=\mathrm{E}\left[e^2(n) \right]$$
$$a_{opt} = \mathrm{h}^{\alpha} \frac{R_y(1) - \alpha R_y(2)}{R_y(0) - 2\alpha R_y(1) + \alpha^2 R_y(0)} \quad (15)$$

It is called OPFLP-2 hereafter.

2. First-order OPFLP model with three samples memory:

$$\hat{y}(n) = a.h^{\alpha}.y(n-1) - a.h^{\alpha}.\left(\alpha y(n-2) + \frac{\alpha(1-\alpha)}{2} y(n-3) \right) \quad (16)$$

In this case, the optimal coefficient a_{opt} is determined by minimizing the mean squared prediction error:

$$a_{opt} = \arg\min_{a} \ \mathrm{J}=\mathrm{E}\left[e^2(n) \right]$$
$$a_{opt} = h^{\alpha} \frac{\left(R_y(1) - \alpha R_y(2) - \frac{\alpha(1-\alpha)}{2} R_y(3) \right)}{R_y(0) - 2\alpha\left(\frac{1-\alpha}{2} R_y(2) \right) + \alpha^2 (I)} \quad (17)$$
$$I = R_y(0) - (\alpha - 1) R_y(1) + \frac{(\alpha-1)^2}{4} R_y(0),$$

The obtained model is called OPFLP-3 hereafter.

3. First-order OPFLP model with four samples memory:

$$\hat{y}(n) = a.h^{\alpha} y(n-1) - \alpha y(n-2) - \frac{a.h^{\alpha}.\alpha(1-\alpha)}{2} \left(y(n-3) + \frac{2-\alpha}{3} y(n-4) \right) \quad (18)$$

By minimizing the mean squared of the prediction error, the optimal coefficient is obtained as follows:

$$a_{opt} = \arg\min_{a} \ J = E\left[e^2(n) \right] =$$
$$h^{\alpha}.\frac{\left(R_y(1) - \alpha R_y(2) - \frac{\alpha(1-\alpha)}{2}(A) \right)}{\left(B + \frac{\alpha^2(\alpha-1)^2}{4} C + D \right)}$$
$$A = R_y(3) - \frac{\alpha - 2}{3} R_y(4)$$
$$B = R_y(0) - 2\alpha R_y(1) + \alpha^2 R_y(0)$$
$$C = R_y(0) - \frac{(2\alpha - 4)}{3} R_y(1) + \frac{(\alpha-2)^2}{9} R_y(0) \quad (19)$$
$$D = \alpha(\alpha-1) R_y(2) - \alpha R_y(1) - (\alpha - 1)$$
$$\times\alpha\left((\alpha-2) R_y(3) + \alpha(\alpha-2) R_y(4) \right) / 3$$

It is called OPFLP-4 hereafter.

Considering (18), (20) and (22), if the fractional derivative order α tends to zero, in all three cases the coefficient a_{opt} becomes: $a_{opt} = \frac{R_y(1)}{R_y(0)}$ [14], [15], [16].

Therefore, it can be concluded that the OPFLP is the generalization of the classical first-order LP. In other words, the first-order LP is a special case of the first-order OPFLP model when the fractional derivative tends to zero [14], [15], [16].

III. Proposed Methods

In this section, we derive closed-form expressions for the optimal sampling frequency $\left(F_{s-opt} \right)$ parameter across three OPFLP models $(m = 1, 2, 3)$. This provides valuable and comprehensive insights into the theoretical foundations of these models. Additionally, to delve into the practical implications, we introduce a signal model to investigate how real-world environmental interferences such as noise and reverberation influence the performance and robustness of OPFLP systems, offering insights into enhancing the robustness and efficacy of OPFLP techniques amidst real-world challenges.

While the work presented in [14] offers a valuable closed-form expression for the optimal coefficient, its approach to determining the differential order (α_{opt}) relies on a trial-and-error only in $0 \le \alpha \le 1$. This method, while effective, can be inefficient and lacks theoretical justification. Additionally, the

authors rightfully acknowledge the significant impact of sampling frequency on model performance, but [14] does not provide a framework to determine a closed-form solution for the optimal sampling frequency.

Also [14] relies on a single frame of voiced speech, when applying OPFLP to speech. This leads to the results that may not adequately represent practical situations. In practical situations, the processed speech signal contains voiced, unvoiced, and silent segments, commonly contaminated by environmental interferences

In our work, we address these limitations by:

1- Deriving a closed-form expression for the optimal sampling frequency: This eliminates the need for arbitrary choices and provides a data-driven approach to setting this crucial parameter.

2- Determining the optimal derivative order in $0 \le \alpha \le 2$: It is crucial to note that the system behavior, concerning stability and robustness, differs considerably between the cases $0 \le \alpha \le 1$ and $1 < \alpha \le 2$.

3- Evaluating the model's performance and robustness against environmental interferences in diverse scenarios: We analyze the model's performance in noisy, reverberant, and combined noisy-reverberant environments.

By addressing these limitations, our approach offers a more robust and theoretically grounded framework for model development compared to the work presented in [14].

A. Optimal Sampling Frequency

By considering $h = 1/F_s$, the following three cases are respectively obtained from (14), (16) and (18):

1. First-order OPFLP model with two samples memory:

$$\hat{y}(n) = a.F_s^{\alpha}\left(y(n-1) - \alpha y(n-2)\right) \quad (20)$$

To obtain the optimal sampling frequency (F_{s-opt}), defining the prediction error as $e(n) = \hat{y}(n) - y(n)$ and minimizing the mean squared error criteria, we have:

$$F_{s-opt} = \arg\min_{F_s} \ \mathrm{J}=\mathrm{E}\left[e^2(n)\right]$$
$$F_{s-opt} = e^{\frac{\ln\left(\frac{R_y(2)\alpha + R_y(1)}{a\left(R_y(0) + 2\alpha R_y(1) + \alpha^2 R_y(0)\right)}\right)}{\alpha}} \quad (21)$$

2. First-order OPFLP model with three samples memory:

$$\hat{y}(n) = a.F_s^{\alpha} y(n-1) -$$
$$a.F_s^{\alpha}\left(\alpha y(n-2) + \frac{\alpha(1-\alpha)}{2} y(n-3)\right) \quad (22)$$

In this case, F_{s-opt} is determined by minimizing the mean squared prediction error:

$$F_{s-opt} = \arg\min_{F_s} \ \mathrm{J}=\mathrm{E}\left[e^2(n)\right]$$
$$F_{s-opt} = \mathrm{e}^{\frac{\ln\left(\frac{2\left(\alpha^2 R_y(3) - 2R_y(2)\alpha - \alpha R_y(3) + 2R_y(1)\right)}{a(A+B)}\right)}{\alpha}}$$
$$A = R_y(0)\alpha^4 - 2R_y(0)\alpha^3 - 4R_y(1)\alpha^3 + 5R_y(0)\alpha^2$$
$$B = 4R_y(1)\alpha^2 + 4R_y(2)\alpha^2 - 8R_y(1)\alpha - 4R_y(2)\alpha + 4R_y(0) \quad (23)$$

3. First-order OPFLP model with four samples memory:

$$\hat{y}(n) = a.F_s^{\alpha} y(n-1) - \alpha y(n-2) -$$
$$\frac{a.F_s^{\alpha}.\alpha(1-\alpha)}{2}\left(y(n-3) + \frac{2-\alpha}{3} y(n-4)\right) \quad (24)$$

By minimizing the mean squared of the prediction error, the F_{s-opt} is obtained:

$$F_{s-opt} = \arg\min_{F_s} \ \mathrm{J}=\mathrm{E}\left[e^2(n)\right]$$
$$F_{s-opt} = e^{\frac{\ln\left(\frac{6(A+B)}{a(C+D)}\right)}{\alpha}}$$
$$A = -\alpha^3 R_y(4) + 3\alpha^2 R_y(3) + 3\alpha^2 R_y(4)$$
$$B = -6\alpha R_y(2) - 3\alpha R_y(3) - 2\alpha R_y(4) + 6R_y(1)$$
$$C = R_y(0)\alpha^6 - 6R_y(0)\alpha^5 - 6R_y(2)\alpha^5 + 22R_y(0)\alpha^4 + 24R_y(1)\alpha^4 + 12R_y(2)\alpha^4 - 30R_y(0)\alpha^3 - 66R_y(1)\alpha^3 - 36R_y(2)\alpha^3 - 12R_y(3)\alpha^3$$
$$D = 49R_y(0)\alpha^2 + 48R_y(1)\alpha^2 + 60R_y(2)\alpha^2 + 36R_y(3)\alpha^2 - 72R_y(1)\alpha - 36R_y(2)\alpha - 24R_y(3)\alpha + 36R_y(0) \quad (25)$$

B. Signal model

In this study, it is assumed that the microphone receives the speech signal contaminated by noise and reverberation. So, the received microphone signal, $y(n)$, is expressed as follows:

$$y(n) = s(n) * h(n) + v(n) = x(n) + v(n) \quad (26)$$

where $h(n)$ is the impulse response between an unknown speech source and the microphone, $v(n)$ is an additive noise signal, and $*$ stands for linear convolution. The statistical characteristics of $x(n)$ and $v(n)$ are assumed

to be real, broadband, and zero mean. The two signals are also considered to be statistically uncorrelated. Fig.1. shows an arbitrary room where we record the speech signal.

IV. Evaluations And Simulation Results

Our primary objective is to determine the optimal derivative order (α) for OPFLP considering variations in sampling frequency. Then, our investigation extends to the evaluation of OPFLP's effectiveness in hands-free speech signal acquisition.

In this regard, we introduce three performance measures, delving into their relevance to our study in subsection IV.A. Subsection IV.B focuses on the effect of sampling frequency on the optimal derivative order (α). We further explore the impact of environmental interferences, e.g. noise and reverberation, on OPFLP's robustness in subsection IV.C under three distinct scenarios: 1) reverberant environment, 2) Noisy environment, and 3) noisy and reverberant environment. To establish a comparative framework, we leverage the first- and second-order LPs [9] and CFDLP [12] as baseline methods, considering their established performance across diverse signal processing applications.

CFDLP is a novel approach built upon the conventional FDLP model, offering two key advantages: 1- Enhanced Interpretability by simplifying the interpretation of the complex cepstrum. It facilitates understanding complex cepstral components as temporal modulations within an all-pole model approximation of the signal's power spectrum. 2- Enhancing computational efficiency compared to conventional FDLP. This translates to faster processing times.

A. Performance Measures

Within signal prediction, the evaluation of prediction performance is fundamental, serving as a crucial measure of efficacy. One fundamental metric in this arena is Prediction Gain (PG), extensively utilized for analyzing prediction performance, particularly in relation to alterations in sampling frequency.

PG, measured in decibels (dB), quantifies the ratio between the variance of the input signal and the variance of the error [14], [15], [16]; i.e.:

$$PG\left[dB\right]=10\log_{10}\left(\frac{\sigma_y^2}{\sigma_e^2}\right) \tag{27}$$

where σ_y^2 and σ_e^2 are the input signal variance and the error variance, respectively. Furthermore, our secondary focus lies in the comprehensive assessment of OPFLP performance

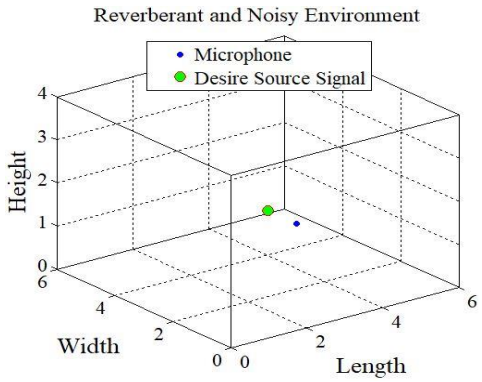

Fig.1: Simulated Environment

within hands-free speech acquisition. Specifically, our scrutiny centers on the received speech quality, a key aspect of analysis.

To gauge and quantify this aspect, we employ the Perceptual Evaluation of Speech Quality (PESQ), an esteemed measure parameter widely acknowledged within the field for its reliability and relevance in evaluating speech quality [24], and Signal To Noise power Ratio (SNR).

B. The Effect of Sampling Frequency

The influence of sampling frequency on linear prediction is fundamental in determining the accuracy and effectiveness of the signal prediction model. Sampling frequency directly affects the resolution and accuracy with which the signal is represented in the digital domain. Higher sampling frequencies capture more details of the original analog signal, allowing for better representation and potentially improving the accuracy of linear prediction models. However, excessively high sampling frequencies can lead to redundant information and increased computational complexity without significant gains in prediction accuracy. Conversely, lower sampling frequencies might result in a loss of crucial signal details, leading to reduced accuracy in prediction models. As per the Nyquist theorem, the sampling frequency must be at least twice the maximum frequency to avoid aliasing. However, beyond merely meeting this criterion, the specific relationship between the sampling frequency and the optimal α becomes critical in fractional linear prediction models. To comprehensively examine the impact of sampling frequency on OPFLP performance, we employ a sinusoidal signal $x\left(t\right)=A\sin\left(2\pi ft\right)$ characterized by frequency $f=1Hz$ and amplitude $A=1$. The sampling frequency is systematically varied from 2 to 70 times the maximum signal frequency to examine its influence on the optimum fractional derivative order. Notably, within the range of α spanning 0 to 2, the α value yielding the maximum PG is selected as the optimal parameter.

Fig. 2 shows the results obtained for the optimal α values.

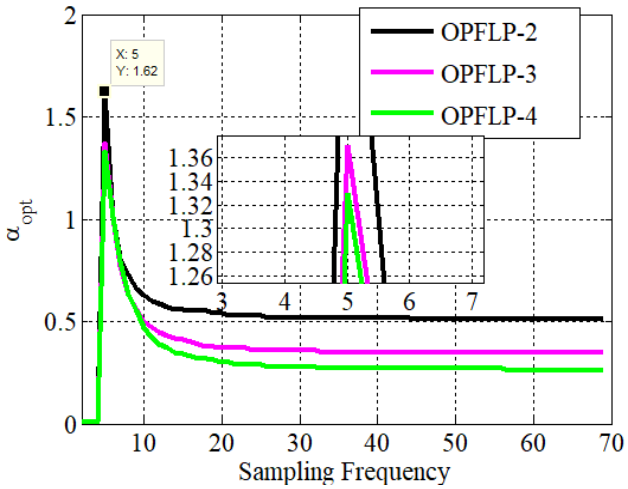

Fig. 2. the optimum derivative order of OPFLP in terms of sampling frequency changes

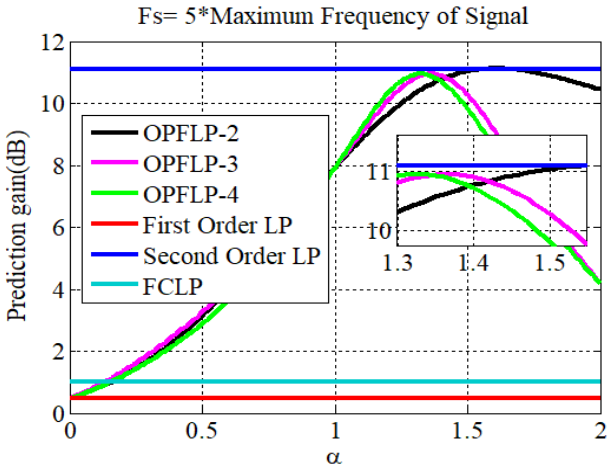

Fig. 3. PG in term of α changes and $F_s = 5\times$ maximum frequency

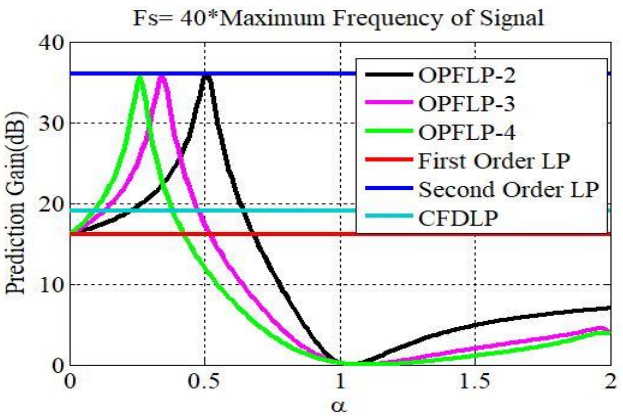

Fig. 4. PG in term of α changes and $F_s = 40\times$ maximum frequency

Remarkably, across all three scenarios, employing OPFLP with memory settings of 2, 3, and 4 at sampling frequencies of 2, 3, and 4 times the maximum signal frequency consistently yields $\alpha_{opt} = 0$. However, when the sampling frequency escalates to 5 and 6 times the maximum signal frequency, the α_{opt} resides within $1 < \alpha \le 2$. Intriguingly, as the sampling frequency surpasses 7 times the maximum signal frequency, the α_{opt} shifts within the $0 \le \alpha \le 1$. Figs. 3 and 4 further elucidate the dynamics, showcasing the impact of α variations on PG at sampling frequencies 5 and 40 times greater than the maximum signal frequency, respectively.

Further analysis reveals important insights. For instance, at a memory setting of 2, α_{opt} at 5 times the frequency settles at 1.6, while at 40 times, it reduces to 0.51. These findings, spanning Figs. 2 to 4, unequivocally highlight the substantial influence of sampling frequency on the performance dynamics of OPFLP.

This sensitivity to sampling frequency highlights the intricate relationship between the sampling rate and the selection of the fractional derivative order, emphasizing the need for meticulous consideration and adjustment of sampling parameters to optimize the performance of fractional linear prediction techniques.

C. Performance and Robustness in Challenging Environmental Conditions

Here, we evaluate and compare the performance and robustness of the OPFLP with the first-, second-order LP, and CFDLP in a simulated noisy and reverberant environment of size $6m \times 6m \times 4m$ (length × width × height). The source-microphone distance is $1m$.

The room impulse response is computed using an effective implementation of the Image method [25]. At the sampling frequency of 16 kHz, the length of room impulse response is set to 4096 samples.

The performance and robustness of the OPFLP are examined in the following three scenarios of the presence of noise and reverberation:

Scenario 1: Reverberant environment

Scenario 2: Noisy environment

Scenario 3: Noisy and reverberant environment

The additive noise contains a mixture of spatially white noise and spherically isotropic (diffuse) noise generated using the method explained in [26]. Experiments are performed on speech signals from the TIMIT dataset [27] including 700 sentences uttered by 26 female and 44 male speakers in 8 different dialects of the American English language. The processing was performed at $16kHz$ sampling frequency on $30ms$ frames of the speech signal. These conditions are called Common Conditions (CC) in the rest of this paper.

Scenario 1:

To evaluate the performance in the presence of reverberation, the OPFLP models are compared to the first-, second-order LPs, and CFDLP at CC at three reverberation times $RT_{60} = 150ms$ (low reverberation), $RT_{60} = 350ms$ (moderate reverberation), and $RT_{60} = 700ms$ (high reverberation). The results of this experiment are presented in Tables 1 and 2 in terms of PG and PESQ, respectively.

Table 1 shows the PG for different models. Comparing the performance of the OPFLP at different reverberation times, it is observed that the prediction performance decreases as the reverberation time increases. For example, the OPFLP-2 can obtain $PG = 9.74dB$ in $RT_{60} = 150ms$ and when the reverberation is increased ($RT_{60} = 700ms$) in the environment, its performance decreases to $PG = 7.12dB$. Therefore $\Delta PG = 9.74 - 7.12 = 2.62dB$ which shows the influence of the reverberation on the performance.

On the other hand, the performance of the OPFLP-2 is better than the OPFLP-3 and OPFLP-4; however, the rate of performance decay of the OPFLP-4 is less than OPFLP-2 and OPFLP-3. Thus, the robustness of OPFLP-4 is higher than the other fractional models. The last column of Table 1 shows the difference in prediction gain (ΔPG) between the cases of low and high reverberation times as the PG loss. Comparing the PG loss of each model, it can be seen that the PG loss of OPFLPs is higher than the first-order LP and CFDLP models, but lower than the second-order LP. In other words, it can be argued that in the presence of the reverberation, the robustness of the OPFLP is lower than the first-order LP and CFDLP, but it is higher than the second-order LP.

Table 2 presents the quality of the estimated speech by the OPFLP and the baseline methods. Comparing the first three columns of Table 2, it is observed that speech quality decreases when reverberation time increases. As can be seen in the fourth column, the robustness of the OPFLP-4 is higher than those with two and three memories. Although the yield loss of the OPFLP is higher than first-order LP and CFDLP, it has been less affected than second-order LP by the reverberation.

The results in this table are verified the result of Table 1. Thus, we can conclude that the robustness of OPFLP is higher than the second LP model in reverberant environments.

Scenario 2:

In this scenario, the performance of the OPFLP is evaluated in a noisy environment. We consider the CC along with three simulated noisy environments at signal-to-noise ratios (SNRs) $5dB$, $15dB$, and $30dB$. These SNR values represent high, medium, and low levels of noise in the environment, respectively. According to the results shown in Tables 3 and 4, similar to the previous scenario, the performance of OPFLP decreases by reducing SNR. In other words, the higher the noise level in the environment, the lower the performance.

TABLE 1
PG IN REVERBERANT ENVIRONMENT

Reverberation time / model	RT_{60} = 150ms	RT_{60} = 350ms	RT_{60} = 700ms	ΔPG
First Order LP	7.36	6.72	5.63	**1.73**
Second Order LP	10.34	9.40	7.71	**2.63**
CFDLP [12]	8.06	7.34	6.07	**1.99**
OPFLP-2	9.74	8.86	7.12	**2.62**
OPFLP-3	9.54	8.64	6.94	**2.60**
OPFLP-4	9.52	8.59	6.94	**2.58**

TABLE 2
PESQ IN REVERBERANT ENVIRONMENT

Reverberation time / model	RT_{60} = 150ms	RT_{60} = 350ms	RT_{60} = 700ms	$\Delta PESQ$
First Order LP	3.33	2.29	1.80	**1.53**
Second Order LP	3.70	2.26	1.81	**1.89**
CFDLP [12]	3.41	2.26	1.78	**1.63**
OPFLP-2	3.48	2.19	1.78	**1.7**
OPFLP-3	3.43	2.20	1.76	**1.67**
OPFLP-4	3.40	2.20	1.75	**1.65**

According to the ΔPG , the OPFLP is more robust to noise than the second-order LP, although its robustness is less than the first-order LP and CFDLP. Comparing the results obtained in the first and second scenarios, it can be concluded that the effect of the noise on the OPFLP performance is more severe than the reverberation.

Scenario 3:

Finally, the simulation conditions are considered similar to real-life environments, where the noise and the reverberation exist simultaneously. The results of this experiment are shown in Tables 5 and 6 in terms of the PG and the PESQ, respectively. These results show that the first-order LP and CFDLP are more robust to the environmental interferences rather than the OPFLP; but the OPFLP shows higher robustness compared to the second-order LP. For example, the proposed OPFLP methods (OPFLP-2, OPFLP-3, and OPFLP-4) exhibit a pronounced sensitivity to noise and reverberation. Their performance degradation, ranging from 1.68 to 1.74 units, is significantly lower in low noise and reverberation situations compared to high noisy and reverberant ones.

TABLE 3
PG IN NOISY ENVIRONMENT

Signal-to-Noise Ratio / model	SNR = 30dB	SNR = 15dB	SNR = 5dB	ΔPG
First Order LP	6.78	6.72	1.88	**4.90**
Second Order LP	9.17	9.40	2.09	**7.08**
CFDLP [12]	8.15	7.47	1.47	**6.68**
OPFLP-2	8.22	8.86	1.20	**7.02**
OPFLP-3	8.38	8.64	1.39	**6.99**
OPFLP-4	8.44	8.59	1.54	**6.90**

TABLE 4
PESQ IN NOISY ENVIRONMENT

Signal-to-Noise Ratio / model	SNR = 30dB	SNR = 15dB	SNR = 5dB	$\Delta PESQ$
First Order LP	3.16	2.44	1.66	**1.50**
Second Order LP	3.48	2.66	1.78	**1.70**
CFDLP [12]	3.19	2.41	1.66	**1.53**
OPFLP-2	3.25	2.43	1.72	**1.53**
OPFLP-3	3.21	2.41	1.66	**1.55**
OPFLP-4	3.19	2.40	1.63	**1.56**

In contrast, the second-order LP method suffers from a more substantial performance drop of approximately 2.01 units, highlighting the potential advantage of the OPFLP methods in terms of noise and reverberation robustness. The results of this scenario confirm the outcomes of the previous ones which indicate that if the environmental interferences increase, the OPFLP performance is reduced. Finally, it should be noted that although the second-order LP outperforms the OPFLP, it is less robust to environmental interferences.

Furthermore, the OPFLP requires one parameter to estimate the signal which is very attractive speech coding applications, while the second-order LP requires two parameters.

V. Conclusions

In this paper, the closed-form of optimum sampling frequency was obtained for the One Parameter Fractional Order Linear Prediction (OPFLP). Then, the performance and robustness of the predictor were investigated in the range of $0 \le \alpha \le 2$. Our investigation meticulously probes the optimal α within $0 \le \alpha \le 1$ and $1 < \alpha \le 2$, considering fundamental parameters like sampling frequency and environmental interferences.

TABLE 5
PG IN A NOISY AND REVERBERANT ENVIRONMENT

Signal to Noise Ratio, Reverberation time / model	SNR = 30dB RT_{60} = 150ms	SNR = 15dB, RT_{60} = 350ms	SNR = 5dB, RT_{60} = 700ms	ΔPG
First Order LP	6.74	4.44	1.65	**5.09**
Second Order LP	6.11	5.02	1.83	**4.28**
CFDLP [12]	6.59	4.80	1.56	**5.03**
OPFLP-2	8.35	3.99	1.07	**7.28**
OPFLP-3	8.33	4.29	1.23	**7.10**
OPFLP-4	8.39	4.51	1.37	**7.02**

TABLE 6
PESQ IN NOISY AND REVERBERANT ENVIRONMENT

Signal to Noise Ratio, Reverberation time / model	SNR = 30dB, RT_{60} = 150ms	SNR = 15dB, RT_{60} = 350ms	SNR = 5dB, RT_{60} = 700ms	$\Delta PESQ$
First Order LP	3.16	2.08	1.49	**1.67**
Second Order LP	3.52	2.13	1.51	**2.01**
CFDLP [12]	3.21	2.09	1.51	**1.70**
OPFLP-2	3.22	2.03	1.48	**1.74**
OPFLP-3	3.18	2.01	1.47	**1.71**
OPFLP-4	3.16	2.01	1.48	**1.68**

Simulation results underscore the intrinsic relationship between sampling frequency and the choice of α for optimal OPFLP performance. Notably, when the sampling rate aligns within five to six times the maximum signal frequency, the optimal α range settles within $1 < \alpha \le 2$. Conversely, exceeding a sixfold sampling frequency compared to the signal's maximum frequency shifts the optimal α to $0 \le \alpha \le 1$. Furthermore, in exploring OPFLP's resilience in challenging signal processing tasks, particularly in hands-free speech acquisition applications, our investigations highlight its superior robustness compared to second-order LP in handling environmental interferences. Notably, noise exerts a more pronounced detrimental impact on OPFLP performance than reverberation, illuminating the nuanced effects of these interferences on the model's efficacy. These observations underscore the criticality of considering sampling frequency and environmental interferences in optimizing OPFLP performance. Additionally, the discernment between the impacts of noise and reverberation on OPFLP performance is essential for enhancing its efficacy in practical applications.

REFERENCES

[1] D. Baleanu and A. M. Lopes, *Handbook of Fractional Calculus with Applications*. De Gruyter, 2019.

[2] D. Baleanu, Z. B. Güvenç, J. A. T. Machado, and others, *New trends in nanotechnology and fractional calculus applications*. Springer, 2010.

[3] S. Das and I. Pan, *Fractional order signal processing: introductory concepts and applications*. Springer Science & Business Media, 2011.

[4] H. Sheng, Y. Chen, and T. Qiu, *Fractional processes and fractional-order signal processing: techniques and applications*. Springer Science & Business Media, 2011.

[5] E. Abbaszadeh-Soorami and M. Haddad-Zarif, "Fractional-Order Variable Structure Equations In Robust Control," Authorea Prepr., 2023.

[6] M. Ghamgosar, S. M. Mirhosseini-Alizamini, and M. Dadkhah, "Design of optimal sliding mode control based on linear matrix inequality for fractional time-varying delay systems," Int. J. Ind. Electron. Control Optim., vol. 5, no. 4, pp. 317–325, 2022.

[7] Y.-N. Li, H.-R. Sun, and Z. Feng, "Fractional abstract Cauchy problem with order α in (1, 2)," Dyn. Partial Differ. Equations, vol. 13, no. 2, pp. 155–177, 2016.

[8] C. Li and M. Li, "Hölder regularity for abstract fractional Cauchy problems with order in (0, 1)," J. Appl. Math. Phys, vol. 6, pp. 310–319, 2018.

[9] J. D. Markel and A. H. J. Gray, *Linear prediction of speech*, vol. 12. Springer Science & Business Media, 2013.

[10] H. Yang, Y. Ye, D. Wang, and B. Jiang, "A novel fractional-order signal processing based edge detection method," in 11th International Conference on Control Automation Robotics & Vision, 2010, pp. 1122–1127.

[11] M. Athineos and D. P. W. Ellis, "Frequency-domain linear prediction for temporal features," in IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat. No. 03EX721), 2003, pp. 261–266.

[12] S. Sadhu and H. Hermansky, "Complex Frequency Domain Linear Prediction: A Tool to Compute Modulation Spectrum of Speech," arXiv Prepr. arXiv2203.13216, 2022.

[13] B. Jo and S. Beack, "Representations of the complex-valued frequency-domain LPC for audio coding," IEEE Signal Process. Lett., 2024.

[14] V. Despotovic, T. Skovranek, and Z. Peric, "One-parameter fractional linear prediction," Comput. & Electr. Eng., vol. 69, pp. 158–170, 2018.

[15] T. Skovranek and V. Despotovic, "Audio signal processing using fractional linear prediction," Mathematics, vol. 7, no. 7, p. 580, 2019.

[16] T. Skovranek, V. Despotovic, and Z. Peric, "Optimal fractional linear prediction with restricted memory," IEEE Signal Process. Lett., vol. 26, no. 5, pp. 760–764, 2019.

[17] K. Xu and X. Song, "A Current Noise Cancellation Method Based on Fractional Linear Prediction for Bearing Fault Detection," Sensors, vol. 24, no. 1, p. 52, 2023.

[18] A. A. Navish, M. Priya, and R. Uthayakumar, "A comparative study on estimation of fractal dimension of EMG signal using SWT and FLP," Comput. Methods Biomech. Biomed. Eng. Imaging ¥& Vis., vol. 11, no. 3, pp. 586–597, 2023.

[19] E. A. P. Habets and J. Benesty, "A two-stage beamforming approach for noise reduction and dereverberation," IEEE Trans. Audio. Speech. Lang. Process, vol. 21, no. 5, pp. 945–958, 2013.

[20] V. W. Neo, C. Evers, and P. A. Naylor, "Enhancement of noisy reverberant speech using polynomial matrix eigenvalue decomposition," IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 29, pp. 3255–3266, 2021.

[21] N. Yazdi and K. Todros, "Measure-transformed MVDR beamforming," IEEE Signal Process. Lett., vol. 27, pp. 1959–1963, 2020.

[22] S. Zhang and X. Li, "Microphone array generalization for multichannel narrowband deep speech enhancement," arXiv Prepr. arXiv2107.12601, 2021.

[23] J. Benesty, J. Chen, and E. A. P. Habets, *Speech enhancement in the STFT domain*. Springer Science & Business Media, 2011.

[24] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP),, 200, vol. 2, pp. 749–752, May 2001.

[25] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am., vol. 65, no. 4, pp. 943–950, 1979.

[26] E. A. P. Habets, I. Cohen, and S. Gannot, "Generating nonstationary multisensor signals under a spatial coherence constraint," J. Acoust. Soc. Am., vol. 124, no. 5, pp. 2911–2917, 2008.

[27] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallet, and N. L. Dahlgren, "TIMIT Acoustic Phonetic Continuous Speech Corpus," Linguist. Data Consortium, 1993, Accessed: May 24, 2019. [Online].Available:http://ci.nii.ac.jp/naid/20000921365/en/.

Seyed Mostafa Alaviyan Shahri received the B.Sc. degree in Telecommunication Engineering from Birjand University and the M.Sc. degree in Communication System Engineering from Yazd University, Yazd, Iran, in 2012 and 2015, respectively. He is currently Ph.D candidate at Yazd University, Yazd, Iran. His research interests include speech processing, fractional calculus, and Artificial intelligence. Email: mostafa.alavian@yahoo.com.

Hamid Reza Abutalebi received the B.Sc. and M.Sc. degrees from Sharif University of Technology, Tehran, Iran in 1996 and 1998, respectively, and the Ph.D. degree from Amirkabir University of Technology, Tehran, Iran in 2003, all in electrical engineering (signal processing). Also, he was with University of Waterloo, Ontario, Canada as a visiting scholar during Mar. 2002 to Feb. 2003. Since 2003, Dr. Abutalebi has been with Electrical Engineering Department of Yazd University, Yazd, Iran, where he is now a full professor. During fall 2010 to summer 2011, he was on sabbatical at Idiap Research Institute, Martigny, Switzerland. His research interests are microphone arrays, speech enhancement, sound source localization, and time-frequency analysis.