Parameter estimation of multiple poles by subspace-based methods

Prabhu Vijayan[1,2], Mariya Ishteva[1] and Ivan Markovsky[2]

***Abstract*— This work deals with the problem of estimating model parameters (frequency, damping factor, amplitude and phase) of a signal modelled as a finite sum of complex damped exponentials, that has multiple poles. Various estimation techniques like rotational invariant method and Hankel Total Least Square (HTLS) method estimate the parameters accurately, but they perform poorly, when there are signal poles with multiplicity greater than one. In this paper, a generalized signal model, that considers multiplicity of the poles for the sum of damped exponential, is used, and the parameters of interest are retrieved by various subspace-based and dimensionality reduction methods, and they are compared.**

keywords: parameter estimation, complex damped exponential, harmonic retrieval, multiple poles.

I. Introduction

In fields such as communications and signal processing, the retrieval of information from the measured data is extremely important. These data, corrupted by noise, do not enable to predict the signal deterministically. Thus, various statistical methods of harmonic retrieval are used. In this paper, the harmonic retrieval of the parameters of the sum of complex exponentials is considered, which finds application in electromagnetic pulse situation to define the various pulse wave forms as given in [1]. Numerous techniques to estimate the parameters are available, such as maximum likelihood methods [2], linear prediction methods [3], subspace-based methods [4]-[6] and matrix pencil techniques [7]. This paper focuses on the signal sequence, that has multiple signal poles. The aforementioned methods estimate the parameters poorly, when the poles are multiple [8]. This issue is addressed by the more generalized signal model that considers the multiplicity of poles. Although there aren't any practical application to our knowledge, where poles have multiplicity more than one, it is a challenging mathematical problem leading to interesting results. The discrete-time signal $x[n]$ or x_n is modelled as a sum of K complex damped exponentials, corrupted by white Gaussian noise b_n:

$$\begin{aligned} x_n &= \sum_{k=1}^{K} (a_k e^{j\phi_k})(e^{\{(-\alpha_k + j2\pi f_k)\Delta t\}n}) + b_n, \\ n &= 0, 1, ..., N-1, \end{aligned} \tag{1}$$

where $j = \sqrt{-1}, n$ is the time index, N is the number of samples, K is the model order and Δt is the sampling time interval. The parameters defining the model are the frequency f_k, damping factor α_k, amplitude a_k and phase ϕ_k. The signal model in (1) is generalized further, considering the multiplicity of the signal poles as given in (2).

$$\begin{aligned} x_n &= \sum_{k=1}^{r} \sum_{l=1}^{M_k} (a_k e^{j\phi_k}) n^{l-1} (e^{\{(-\alpha_k + j2\pi f_k)\Delta t\}n}) + b_n, \\ n &= 0, 1, ..., N-1, \end{aligned} \tag{2}$$

Here, r is the number of unique poles whose multiplicity is $M_k, k = 1, 2, ..., r$. Thus, the multiplicities of r signals sums up to K. More compactly, the model (2) can be given as:

$$x_n = \sum_{k=1}^{r} \sum_{l=1}^{M_k} c_k n^{l-1} z_k^n + b_n, \quad n = 0, 1, ..., N-1, \tag{3}$$

where $c_k = a_k e^{j\phi_k}$ and $z_k = e^{\{(-\alpha_k + j2\pi f_k)\Delta t\}}$ are the k^{th} complex amplitude and k^{th} signal pole, respectively.

Various estimation methods as in [8] have been derived to estimate the signal poles z_k from the signal data like matrix pencil techniques [7], forward and backward linear prediction [3] and subspace-based methods [4]-[6]. The problem with these established methods is that, they perform poorly when two or more poles appear closer to each other. The prediction error is significantly high. The main aim of this work is to define a more generalized model and to extend these methods for the multiple poles scenario and compare them.

The paper is structured as follows. Section II explains the estimation techniques, namely Hankel Total Least Square (HTLS), Kung's methods, linear prediction and model order reduction, Section III proposes a method to estimate the parameters. Further, section IV discusses the influence of Hankel matrix size on the estimates, and an optimal size is introduced. Section V presents the results of numerical simulation and compares different methods, leading to the conclusion in Section VI.

II. Signal Poles Estimation By Various Methods

A. Hankel Total Least Square Method

The HTLS method is a subspace-based parameter estimation method, that operates on the data matrix $\boldsymbol{H}$, which is of Hankel structure and constructed from the N data samples of the discrete signal x_n from (2). The size of the Hankel matrix $\boldsymbol{H}$ is given by $L \times M$ and $L > K$, where L is design

[1] Department of Computer Science, Geel Campus, KU Leuven, 2440 Geel, Belgium.
[2] Department ELEC, Vrije Universiteit Brussel, 1050 Brussels Belgium. prabhu.vijayan@kuleuven.be

or hyperparameter and $M = N - L + 1$, $M > K$.

$$\boldsymbol{H} = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_{M-1} \\ x_1 & x_2 & \iddots & \dots & x_M \\ x_2 & \iddots & \iddots & \dots & x_{M+1} \\ \vdots & \iddots & \iddots & \dots & \vdots \\ x_{L-1} & \dots & \dots & x_{N-2} & x_{N-1} \end{bmatrix}$$

The choice of L plays a significant role in the performance of subspace-based methods. The detailed study about the effects of hyperparameter on the performance is provided in Section IV.

The data matrix $\boldsymbol{H}$, for the noiseless and simple pole case, admits Vandermonde decomposition (4), which can be used to estimate complex amplitudes c_k and signal poles $z_k, k = 1, 2, ..., K$

$$\boldsymbol{H} = \boldsymbol{S}\boldsymbol{C}\tilde{\boldsymbol{S}}^T, \tag{4}$$

where $\boldsymbol{C} = \text{diag}(c_1, c_2, ..., c_K)$, $\boldsymbol{S}$ and $\tilde{\boldsymbol{S}}$ are the Vandermonde matrices of size $L \times K$ and, $M \times K$ respectively:

$$\boldsymbol{S} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_K \\ z_1^2 & z_2^2 & \dots & z_K^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{L-1} & z_2^{L-1} & \dots & z_K^{L-1} \end{bmatrix}.$$

The above matrix and decomposition in (4) are true only when all the poles are simple. For the general case, where poles can be multiple, the decomposition is generalized with block diagonal matrix $\boldsymbol{C}$ and confluent Vandermonde matrices $\boldsymbol{S}$ and $\tilde{\boldsymbol{S}}$, as given in [13]. The estimated signal poles of (2) differ numerically, that they are considered as distinct. The parameters can be indirectly estimated by exploiting the shift-invariance property of $\boldsymbol{S}$ given in [9], as:

$$\boldsymbol{S}_{(+1)} = \boldsymbol{S}_{(-1)}\boldsymbol{Z}, \tag{5}$$

where $\boldsymbol{S}_{(+1)}$ and $\boldsymbol{S}_{(-1)}$ are matrices created by removing one row from the top and bottom of $\boldsymbol{S}$ respectively. $\boldsymbol{Z}$ is a diagonal matrix consisting of all signal poles $z_k, k = 1, 2, ..., K$. Using the Singular Value Decomposition (SVD) of the data matrix $\boldsymbol{H}$ and equation (5), the parameters can be estimated with better numerical stability. For the data matrix with noise, the total least square method is used [9].

The SVD of $\boldsymbol{H}$ is given by $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$ and truncating it to rank K: $\boldsymbol{H} \approx \boldsymbol{H}_K = \boldsymbol{U}_K\boldsymbol{\Sigma}_K\boldsymbol{V}_K^H$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are matrices containing (ordered) left and right singular vectors, respectively. $\boldsymbol{U}_K$ and $\boldsymbol{V}_K$ are sub-matrices containing first K columns of $\boldsymbol{U}$ and, $\boldsymbol{V}$, respectively. $\boldsymbol{\Sigma}$ is the singular value matrix and $\boldsymbol{\Sigma}_K$ is the leading $K \times K$ submatrix of $\boldsymbol{\Sigma}$. Using the shift-invariance property, the signal poles can be estimated as given in (6):

$$\boldsymbol{U}_{K(+1)} = \boldsymbol{U}_{K(-1)}\boldsymbol{E}, \tag{6}$$

$\boldsymbol{U}_{K(+1)}$ and $\boldsymbol{U}_{K(-1)}$ are formed by deleting a row from the top and bottom of $\boldsymbol{U}_K$ respectively. The eigenvalues of the $K \times K$ matrix $\boldsymbol{E}$ are the estimates of signal poles $\hat{z}_k$, $k = 1, ..., K$. The equation (6) is solved by total least square method. The estimated signal poles are used to retrieve the model parameters, as explained in Section III.

B. Kung's Realization Method

The problem of system identification is solved by another subspace method known as Kung's method [10]. The rank revealing factorization of the data matrix $\boldsymbol{H}$ is given by,

$$\boldsymbol{H} = \boldsymbol{\Gamma}\boldsymbol{\Delta}, \tag{7}$$

where $\boldsymbol{\Gamma}$ and $\boldsymbol{\Delta}$ are the observability and controllability matrices. But this is a nontrivial problem in finite precision arithmetic. Thus, the numerically stable SVD is used to determine the rank of the data matrix,

$$\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H. \tag{8}$$

The model order K being the rank of the matrix, the first K singular values are significant to compute a Hankel matrix equivalent to the original data matrix $\boldsymbol{H}$. Thus, the partition is defined as:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{U}_K & \tilde{\boldsymbol{U}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_K & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{\Sigma}} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_K^H \\ \tilde{\boldsymbol{V}}^H \end{bmatrix}, \tag{9}$$

where $\boldsymbol{U}_K$ and $\boldsymbol{V}_K$ are first K singular vectors and $\boldsymbol{\Sigma}_K$ presents first K singular values. $\tilde{\boldsymbol{U}}$ and $\tilde{\boldsymbol{V}}$ are remaining singular vectors of $\boldsymbol{U}$, and $\boldsymbol{V}$ respectively, and $\tilde{\boldsymbol{\Sigma}}$ presents remaining singular values of $\boldsymbol{\Sigma}$.

The factors $\boldsymbol{\Gamma}$ and $\boldsymbol{\Delta}$ in (7) are then given by:

$$\begin{aligned} \boldsymbol{\Gamma} &= \boldsymbol{U}_K\sqrt{\boldsymbol{\Sigma}_K}, \\ \boldsymbol{\Delta} &= \sqrt{\boldsymbol{\Sigma}_K}\boldsymbol{V}_K^H. \end{aligned} \tag{10}$$

Here the following linear system of equations is solved for $\boldsymbol{E}$,

$$\boldsymbol{\Gamma}_{(+1)}\boldsymbol{E} = \boldsymbol{\Gamma}_{(-1)}, \tag{11}$$

where $\boldsymbol{\Gamma}_{(+1)}$ and $\boldsymbol{\Gamma}_{(-1)}$ are sub-matrices of $\boldsymbol{\Gamma}$ with first and last row removed respectively. The eigenvalues of $\boldsymbol{E}$ are the signal pole estimates $\hat{z}_k$, $k = 1, ..., K$. The estimated signal poles are used to retrieve the model parameters, as explained in Section III.

C. Linear Prediction

Linear prediction is an important method in signal processing, to predict the new sample, for a given set of previous samples or to predict past sample, for a given set of next samples. The number of samples provided to the predictor is the order of the model K.

The sum of exponential signal x_n in (2), without noise, satisfies the homogeneous recurrence relation of length K.

$$\begin{aligned} &p_0x_n + p_1x_{n-1} + p_2x_{n-2} + ... + p_Kx_{n-K} = 0, \\ &n = K, K+1, ..., N-1; \\ &p_0 = 1, p_K \neq 0 \end{aligned} \tag{12}$$

With the help of this recurrence relation, the entire signal can be generated, provided the initial K values $x_0, ...x_{K-1}$ are known, as given in [3] and [12]. In the system identification problem, the signal sequence x_n is already known. Thus,

the recurrence relation can be used to solve for the unknown coefficients $p_1, ..., p_K$. With the known signal sequence x_n, the recurrence matrix $\boldsymbol{X}$ of size $(N-K)\times K$ and recurrence vector $\hat{\boldsymbol{x}}$ of size $(N-K)$ can be constructed as below:

$$\boldsymbol{X} = \begin{bmatrix} x_{K-1} & x_{K-2} & x_{K-3} & \dots & x_0 \\ x_K & x_{K-1} & x_{K-2} & \dots & x_1 \\ x_{K+1} & x_K & x_{K-1} & \dots & x_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{N-2} & x_{N-3} & x_{N-4} & \dots & x_{N-K+1} \end{bmatrix},$$

$$\hat{\boldsymbol{x}} = \begin{bmatrix} x_K \\ x_{K+1} \\ x_{K+2} \\ \vdots \\ x_{N-1} \end{bmatrix}.$$

The coefficients $p_1, ..., p_K$ of the monic polynomial (12) are determined by:

$$\boldsymbol{X}\boldsymbol{p} = -\hat{\boldsymbol{x}}, \tag{13}$$

where $\boldsymbol{p}$ is K-length vector containing the coefficients $p_i, i = 1, 2, ..., K$. With the determined coefficients p_i's, the companion matrix $\boldsymbol{A}$ of size $K \times K$ is constructed as below:

$$\boldsymbol{A} = \begin{bmatrix} -p_1 & -p_2 & -p_3 & \dots & -p_{K-1} & -p_K \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

The characteristic equation of the K^{th} order difference equation (12) is given by $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$,

$$\lambda^K + p_1\lambda^{K-1} + p_2\lambda^{K-2} + ... + p_K = 0. \tag{14}$$

Solving (14), the eigenvalues or the signal pole estimates $\hat{z}_k$, $k = 1, ..., K$ are determined. The estimated signal poles are used to retrieve the model parameters, as explained in Section III.

D. Model Order Reduction

The K^{th}-order difference relation given in (12) is K^{th}-order Autoregressive model, where the state equations are described by:

$$\begin{aligned} \boldsymbol{X_n} &= \boldsymbol{A}\boldsymbol{X_{n-1}}, \\ x_n &= \boldsymbol{C}\boldsymbol{X_n}, \end{aligned} \tag{15}$$

where

$$\boldsymbol{X_n} = \begin{bmatrix} x_n \\ x_{n-1} \\ x_{n-2} \\ \vdots \\ x_{n-K+1} \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}_{1\times K}$$

and $\boldsymbol{A}$ is the $K \times K$ companion matrix given in Section II-C. In a compact form, the system in (15) is denoted as:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{C} \end{bmatrix} \in \mathbb{C}^{(K+1)\times K}. \tag{16}$$

Approximating $\boldsymbol{\Sigma}$ with $\hat{\boldsymbol{\Sigma}}$, such that the error between the two systems is small,

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} \hat{\boldsymbol{A}} \\ \hat{\boldsymbol{C}} \end{bmatrix} \in \mathbb{C}^{(K+1)\times K}. \tag{17}$$

The approximation is carried out by Lanczos algorithm as explained in [11]. Let $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$ be the SVD decomposition of data matrix $\boldsymbol{H}$. The data matrix can also be decomposed as given in (7). The factors of this decomposition are computed using the SVD factors as:

$$\begin{aligned} \boldsymbol{\Gamma} &= \boldsymbol{U}\sqrt{\boldsymbol{\Sigma}}, \\ \boldsymbol{\Delta} &= \sqrt{\boldsymbol{\Sigma}}\boldsymbol{V^H}. \end{aligned} \tag{18}$$

The key step in this method is determining LU-factorization of the data matrix,

$$\boldsymbol{H} = \hat{\boldsymbol{L}}\hat{\boldsymbol{U}} = \boldsymbol{\Gamma}\boldsymbol{\Delta}, \tag{19}$$

where $\hat{\boldsymbol{L}}$ and $\hat{\boldsymbol{U}}$ are lower and upper triangular matrices. Left-multiplying (19) by $\hat{\boldsymbol{L}}^{-1}$ and right-multiplying by $\hat{\boldsymbol{U}}^{-1}$,

$$\hat{\boldsymbol{L}}^{-1}(\hat{\boldsymbol{L}}\hat{\boldsymbol{U}})\hat{\boldsymbol{U}}^{-1} = \boldsymbol{I} = \hat{\boldsymbol{L}}^{-1}(\boldsymbol{\Gamma}\boldsymbol{\Delta})\hat{\boldsymbol{U}}^{-1}. \tag{20}$$

The projection or transformation matrices defined using this factorization, as given in (21), are used to define the reduced order system $\hat{\boldsymbol{\Sigma}}$.

$$\begin{aligned} \boldsymbol{p_L} &= \hat{\boldsymbol{L}}^{-1}\boldsymbol{\Gamma}, \\ \boldsymbol{p_\Delta} &= \boldsymbol{\Delta}\hat{\boldsymbol{U}}^{-1}, \end{aligned} \tag{21}$$

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} \hat{\boldsymbol{A}} \\ \hat{\boldsymbol{C}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{p_L}\boldsymbol{A}\boldsymbol{p_\Delta} \\ \boldsymbol{C}\boldsymbol{p_\Delta} \end{bmatrix}. \tag{22}$$

The important properties that has to be satisfied here are: (1) $\boldsymbol{p_L}\boldsymbol{p_\Delta} = \boldsymbol{I}$ and (2) $\hat{\boldsymbol{A}}$ is tridiagonal. The eigenvalues of the reduced state matrix $\hat{\boldsymbol{A}}$ are the signal pole estimates $\hat{z}_k$, $k = 1, ..., K$. The estimated signal poles are used to retrieve the model parameters, as explained in Section III.

III. Harmonic Retrieval

The methods explained in Section II estimate the eigenvalue or signal poles $\hat{z}_k$, $k = 1, ..., K$. From that, the estimates of frequencies $\hat{f}_k$ and damping factors $\hat{\alpha}_k$ for the k^{th} exponential are determined as follows:

$$\lambda_k = \hat{z}_k = e^{\{(-\hat{\alpha}_k + j2\pi\hat{f}_k)\Delta t\}}, \tag{23}$$

$$\hat{f}_k = \frac{1}{2\pi\Delta t}\arctan\left(\frac{\Im(\hat{z}_k)}{\Re(\hat{z}_k)}\right), \tag{24}$$

$$\hat{\alpha}_k = \frac{-\log\left(\sqrt{\Re(\hat{z}_k)^2 + \Im(\hat{z}_k)^2}\right)}{\Delta t}, \tag{25}$$

where $\Re(\hat{z}_k)$ and $\Im(\hat{z}_k)$ are the real and imaginary parts of $\hat{z}_k$ respectively. Similarly, the estimates of complex amplitude $\hat{c}_k$, $k = 1, ...K$ are obtained by solving the $N \times K$ matrix, that is constructed by using estimated λ_k or $\hat{z}_k$ and its multiplicity in (3), by least square method. This procedure leads to a modified Vandermonde matrix $\boldsymbol{T}$.

$$\boldsymbol{T}\boldsymbol{c} = \boldsymbol{x_n}, \quad n = 0, 1, ..., N-1 \tag{26}$$

where

$$\boldsymbol{T} = \begin{bmatrix} \begin{bmatrix} \hat{z}_1^n \\ n\hat{z}_1^n \\ n^2\hat{z}_1^n \\ \vdots \\ n^{M_1-1}\hat{z}_1^n \end{bmatrix}^T \cdots \begin{bmatrix} \hat{z}_r^n \\ n\hat{z}_r^n \\ n^2\hat{z}_r^n \\ \vdots \\ n^{M_r-1}\hat{z}_r^n \end{bmatrix}^T \end{bmatrix}.$$

The subscript r is the number of unique poles, and $M_1, ..., M_r$ are the multiplicities of the r unique poles. The modified Vandermonde matrix $\boldsymbol{T}$ introduced in this work is similar to Pascal-Vandermonde matrix given in [12] or Confluent Vandermonde matrix in [13], except that the k^{th} column, $k = 2, ..., M_i, i = 1, ..., r$ is not a scaled $(k-1)^{\text{th}}$ derivative of the first column.

For the model order $K = 3$ and number of unique poles $r = 1$ with multiplicity $M_1 = 3$, the modified Vandermonde matrix $\boldsymbol{T}$ is given as:

$$\boldsymbol{T} = \begin{bmatrix} \hat{z}_1^n & n\hat{z}_1^n & n^2\hat{z}_1^n \end{bmatrix}$$

$$\boldsymbol{T} = \begin{bmatrix} 1 & 0 & 0 \\ \hat{z}_1 & \hat{z}_1 & \hat{z}_1 \\ \hat{z}_1^2 & 2\hat{z}_1^2 & 4\hat{z}_1^2 \\ \vdots & \vdots & \vdots \\ \hat{z}_1^{N-1} & (N-1)\hat{z}_1^{N-1} & (N-1)^2\hat{z}_1^{N-1} \end{bmatrix}$$

Further, in (26), $\boldsymbol{c}$ is a vector of length K containing complex amplitudes estimates $\hat{c}_k$, $k = 1, ...K$ and $\boldsymbol{x_n}$ is the signal vector of length N. The estimates of amplitude $\hat{a}_k$ and phase $\hat{\phi}_k$ are then determined from complex amplitude $\boldsymbol{c}$, similar to other estimates as given in (28) and (29).

$$\hat{c}_k = \hat{a}_k e^{j\hat{\phi}_k}, \tag{27}$$

$$\hat{a}_k = \sqrt{\Re(\hat{c}_k)^2 + \Im(\hat{c}_k)^2}, \tag{28}$$

$$\hat{\phi}_k = \arctan\left(\frac{\Im(\hat{c}_k)}{\Re(\hat{c}_k)}\right), \tag{29}$$

where $\Re(\hat{c}_k)$ and $\Im(\hat{c}_k)$ are the real and imaginary parts of $\hat{c}_k$ respectively.

IV. Influence Of Design Parameter L On Estimates

The effect of the Hankel matrix row size L on the performance of the HTLS and Kung's realization methods is discussed here. For a discrete-time signal x_n in (2), contaminated by circularly symmetric white Gaussian noise of size $N = 1000$, $K = 2$ damped exponentials, the input parameters $f = [0.02, 0.02]$ Hz, $a = [1, 1], \alpha = [0.009, 0.009], \phi = [0, 0]$ and $\Delta t = 1$. The noise b_n is calculated as below and added to the pure signal, $\tilde{x}_n$,

$$b_n = \frac{RMS(\tilde{x}_n)}{RMS(\tilde{b}_n)} e^{-SNR \cdot \log(10)/20}, \tag{30}$$

where $\tilde{b}_n$ is a random complex number sequence of size N and RMS is the root-mean-square of the quantities. The parameter $L, M \in \mathbb{Z} \in [K+1, N-K]$, since L and M should be greater than K at least by 1. A Monte-Carlo simulation consisting of $N_{runs} = 100$ independent trial runs is performed. The model parameter estimates in Fig. 1 are formulated in terms of averaged RMS error (RMSE) in (31):

$$RMSE(q_k) = \frac{1}{N_{runs}} \sum_{i=1}^{N_{runs}} (\hat{q}_{k,i} - q_k)^2, \tag{31}$$

where q is the model parameter f, α, a or ϕ and q_k is the k^{th} value of a parameter, $\hat{q}_{k,i}$ is the k^{th} estimate for the i^{th} trial run of N_{runs} Monte-Carlo simulation. The averaged RMS error of the parameter q is given as $\overline{RMSE(q)}$:

$$\overline{RMSE(q)} = \frac{1}{K} \sum_{k=1}^{K} \sqrt{RMSE(q_k)}. \tag{32}$$

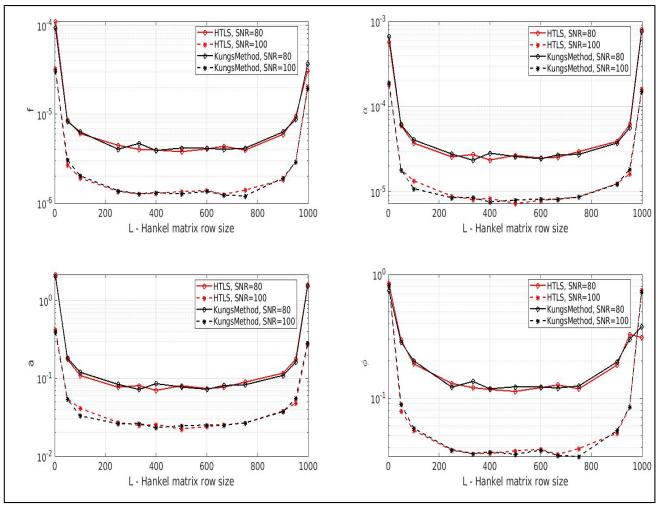

Fig. 1: Averaged RMS error of model parameters estimates vs. Design parameter L for signal with multiple poles by HTLS (red), Kung's Method (black) for $SNR = [80, 100]$. For $L \in [\frac{N}{3}, \frac{N}{2}] \approx [333, 500]$, the estimates are better.

From Fig. 1, it is clear that the model parameter estimates are better, when the Hankel matrix is (nearly) square, i.e. $\frac{N}{3} \leq L_{optimal} \leq \frac{N}{2}$. As the data matrix is rectangular, the rank$(\boldsymbol{H}) = \min(L, M = N - L + 1)$ and the (ordered) SVD of $\boldsymbol{H}$ is not sufficient to compute the model parameters efficiently. The symmetric nature of the estimated parameters along the hyperparameter L is because the matrices of size $L > \frac{N}{2}$ are transpose of the matrices of size $L < \frac{N}{2}$, explaining that the SVD of these matrices are also transposed, and the singular values are the same.

Although the (nearly) square Hankel matrices proves to be efficient in terms of performance, it is associated with higher computational cost in comparison to the taller and wider matrices as shown in Fig. 2. The higher computational cost can be explained by the two computationally expensive procedures in the HTLS algorithm: namely SVD of $\boldsymbol{H}$ and the TLS step. Thus, the square matrix with more elements contribute to the higher cost. In cases, where accuracy is more important than simulation time, the design parameter L can be chosen between $\frac{N}{3}$ and $\frac{N}{2}$.

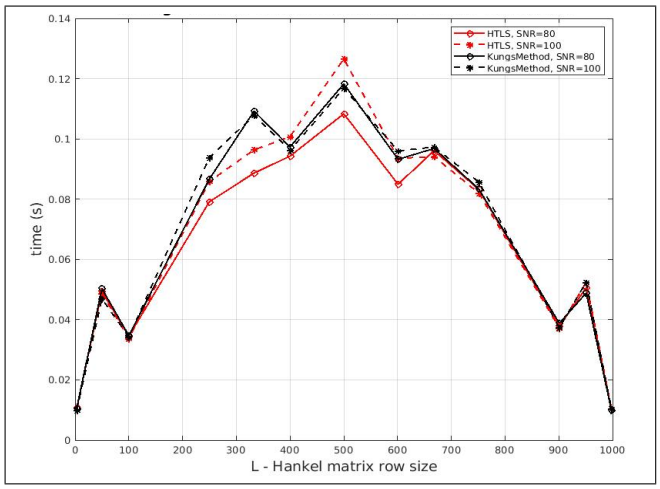

Fig. 2: Average time for 1 Monte-Carlo Simulation vs. Design parameter L for signal with multiple poles by HTLS (red), Kung's Method (black) for $SNR = [80, 100]$. Both of the subspace-based methods perform slower, when $L = \frac{N}{2} = 500$, because of handling large number of elements in the expensive steps: SVD and total least square. Kung's method is considerably slower than HTLS, because of an additional step to determine the factors $\mathbf{\Gamma}$ and $\mathbf{\Delta}$ matrices.

V. Numerical Examples

The algorithms are first tested for the noise free case, where the model parameter estimation is expected to have minimal error. In this example, for $K = 2$, the true parameters of the signal are: $f = [0.02, 0.02]$ Hz, $a = [1, 1], \alpha = [0.009, 0.009], \phi = [0, 0]$. One Monte-Carlo simulation trial with $N = 1000$ samples is performed. The size of the data matrix $\boldsymbol{H}$ for HTLS and Kung's realization method is fixed as $L = \frac{N}{2}$, in order to obtain the better estimates, as explained in IV. The averaged RMS error of the estimates by all four methods is given in the Tab. I. The results from

TABLE I: Averaged RMS Error of estimates for noise less signal

Method	$K = 2$			
	$\hat{f}$	$\hat{\alpha}$	$\hat{a}$	$\hat{\phi}$
HTLS	1e-11	1e-10	1e-15	1e-16
Kung	1e-11	1e-09	1e-15	1e-17
Linear Prediction	1e-09	1e-08	1e-15	1e-16
Model Order Reduction	1e-09	1e-08	1e-15	1e-15

the noise free simulation confirm that the parameters are estimated with good accuracy. Further, the implementation is simulated with noisy data for $N_{runs} = 100$ Monte-Carlo trials, with the same simulation setting as above. The results for $K = 1, 2, 3$ are shown in Fig. 3 - Fig. 5, where the averaged RMS error of estimated model parameters are plotted against the SNR, by all four different methods discussed in Section II.

The following observations can be made:

- For the single pole case, the estimates are good, and get better with increasing SNR. The estimates by HTLS and Kung's realization methods are approximately the same, because the computation of signal poles involve similar procedure, by determining the left singular vector and singular values, and then solving the linear system of equations with shift-invariance property.
- The estimates by linear prediction are poorer than the former two methods, because the computation of state matrix or companion matrix $\boldsymbol{A}$ using the polynomial

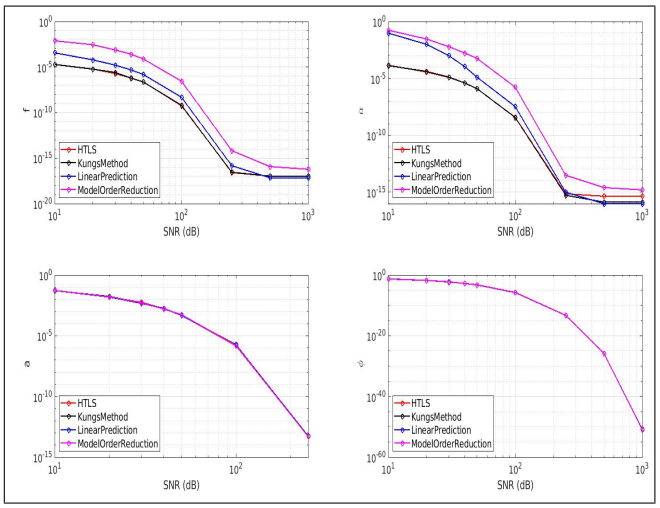

Fig. 3: RMS error of model parameter estimates vs. SNR for simple pole $K = 1, f = 0.02$ Hz, $a = 1, \alpha = 0.009, \phi = 0$ by HTLS (red), Kung's method (black), Linear prediction (blue), model order reduction (magenta). HTLS and Kung's method perform better than linear prediction and model order reduction, because they are based on stable SVD.

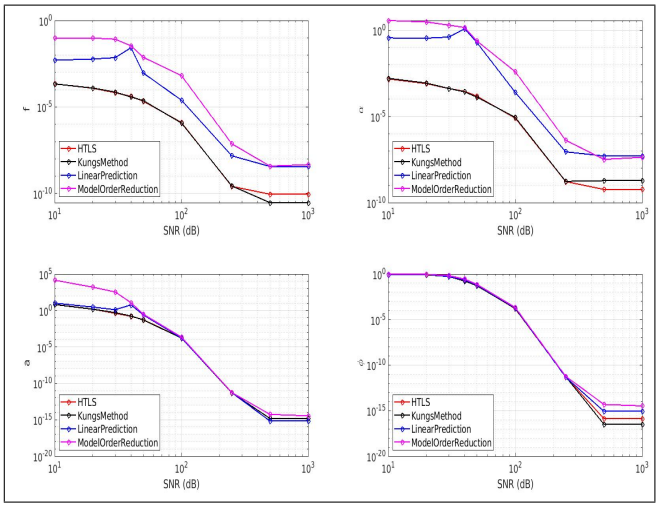

Fig. 4: Averaged RMS error of model parameter estimates vs. SNR for multiple poles $K = 2, f = [0.02, 0.02]$ Hz, $a = [1, 1], \alpha = [0.009, 0.009], \phi = [0, 0]$ by HTLS (red), Kung's method (black), Linear prediction (blue), model order reduction (magenta). Like simple pole case, HTLS and Kung's method perform better. But the estimation error is, in general, higher than the simple pole case, is because of the polynomial times exponential signal model being unstable.

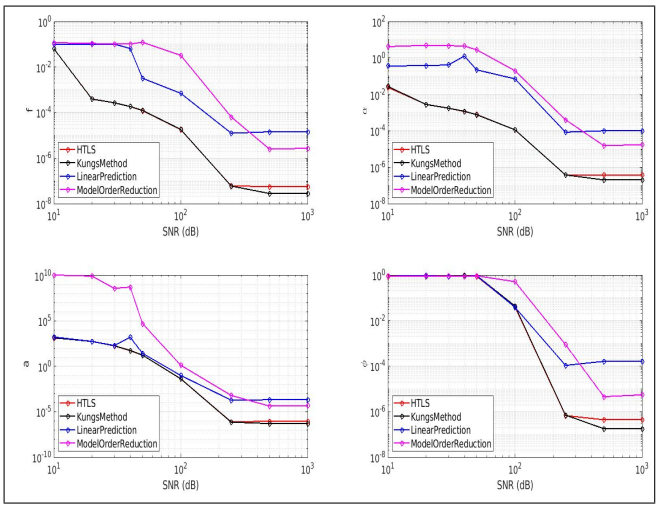

Fig. 5: Averaged RMS error of model parameter estimates vs. SNR for multiple poles $K = 3, f = [0.02, 0.02, 0.02]$ Hz, $a = [1, 1, 1], \alpha = \phi = [0, 0, 0]$ by HTLS (red), Kung's method (black), Linear prediction (blue), model order reduction (magenta). The estimation error increases further because of the higher multiplicity, resulting in unstable signal.

coefficients $p1, ..., p_K$ is unstable, and thus the signal pole estimates $\hat{z}_k$, $k = 1, ..., K$ computed thereafter.

- In the reduced model order, the estimates have much higher deviation, because the two important properties of the transformation matrix and reduced state matrix mentioned in II-D are not satisfied. That is, (1) $\boldsymbol{p_L p_\Delta} \neq \boldsymbol{I}$ and (2) $\hat{\boldsymbol{A}}$ is not tridiagonal. The LU-factorization of the data matrix $\boldsymbol{H}$ in the Lanczos process leads to $\hat{\boldsymbol{L}}$, that is not lower triangular. In cases where LU factors are not lower and upper triangular, rows can be interchanged to attain the desired decomposition. But doing so here, Hankel structure of the data matrix $\boldsymbol{H}$ will be lost. Thus, the method results in increased error.
- In the multiple poles case, the estimates are poorer, especially when the multiplicity of a pole is greater than 3. This inconsistency can be explained by the fact that the signal computed from true parameters is numerically unstable with increasing multiplicity, as the polynomial times exponential function leads to signal overshoot. Thus, the inverse modeling of a system with multiplicity of poles 3 or less, can be successful with the model presented in (2).
- Another important remark is that the model parameter estimates are not necessarily estimated in the same order as they were input in the system. This might lead to an increased error. To avoid such unforeseen error, it is recommended to input the true parameters in an ordered fashion (ascending or descending).

VI. Conclusion

In this paper, a generalized signal model for sum of damped complex exponential is presented, that considers the multiplicities of signal poles. Although this kind of signal is not found in any application to the best of our knowledge, it is an interesting mathematical problem, that lead to study about the stability of sum of damped complex exponential signal with multiple poles. The model parameters were estimated and compared by four different methods: HTLS, Kung's realization, linear prediction and model order reduction. The algorithms were tested by several independent Monte-Carlo simulations with white Gaussian noise. It was also shown that, the data matrix size in case of subspace-based methods, play a significant role in computing the statistical accuracy of parameter estimates, leading to a conclusion that the optimal size of the data matrix should be between one-third and one-half the sample size. It was also demonstrated that the subspace-based methods: HTLS and Kung's realization resulted in better estimates, as they involve SVD, which is numerically stable. On the other hand, the linear prediction and model order reduction methods determine the eigenvalues or signal poles through state matrix of the system, that is constructed by the coefficients determined from the recurrence relation of the polynomial function, which leads to numerically unstable results. Although the algorithms estimate model parameters well, the multiplicities of poles are limited to 3. Higher than that, the signal model becomes numerically unstable, and the prediction error will eventually be high.

References

[1] Sircar, P.: Mathematical Aspects of Signal Processing, 2016, Delhi.

[2] Stoica, P., Sharmen, K.C.: Maximum likelihood methods for direction-of-arrival estimation. IEEE Trans. Acous., 1990, p. 1132–1143.

[3] Kumaresan, R., Tufts, D.W.: Estimating the Parameters of Exponentially Damped Sinusoids and Pole-Zero Modeling in Noise, IEEE Trans. of Acous., Vol. ASSP-30, No.6, 1982, p. 833-839.

[4] Bhaskar Rao, D.V., Arun, K.S.: Model-based processing of signals: a state-space approach. Proc. IEEE 1992, 80, p. 283–309.

[5] Kung, S.Y., Arun, K.S., Bhaskar Rao, D.V.: State-space and singular value decomposition-based approximation methods for the harmonic retrieval problem. J. Opt. Soc. Am. 198, 73(12), p. 1799–1811.

[6] Markovsky, I., Vaccaro, R.J., Van Huffel, S.: System identification by optimal subspace estimation, Dept. EE, K.U.Leuven, 2006.

[7] Hua, Y., Sarkar, T.K.: Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. IEEE Trans. Acoust. 1990, 38, p. 814–824.

[8] Papy, J.-M., De Lathauwer, L., Van Huffel, S.: Exponential data fitting using multilinear algebra: the decimative case, Journal of Chemometrics, DOI:10.1002/cem.1212, p. 341–351.

[9] Morren, G., Lemmerling, P., Van Huffel, S.: Decimative subspace-based parameter estimation techniques, Elsevier, Signal Processing 83 (2003), p. 1025 – 1033

[10] Markovsky, I.: Low-Rank Approximation, Algorithms, Implementation, Applications, 2018, ch.3, p. 72-75.

[11] Antoulas, A.C., Sorensen, D.C.: Approximation of Large-Scale Dynamical Systems: An Overview, International Journal of Applied Mathematics and Computer Science, 2001, Vol.11, No.5, p. 1093–1121.

[12] Roland, B., Bertrand, D., and Gaël, R.: High resolution spectral analysis of mixtures of complex exponentials modulated by polynomials, IEEE Transactions on Signal Processing, Vol. 54, No. 4, April 2006.

[13] Franklin T.L., Sanzheng, Q., David, V.: Exponential Decomposition and Hankel Matrix, Mathematics in Signal Processing V: Based on the Proceedings of a Conference, 2000, p. 275-285.