IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 19 September 2023, accepted 12 October 2023, date of publication 15 December 2023,
date of current version 22 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3343556

RESEARCH ARTICLE

A Comparison of SVD-Augmented Prony Algorithms for Noisy Power System Signals

KRISHNA RAO[1], (Graduate Student Member, IEEE), AND K. N. SHUBHANGA[2], (Senior Member, IEEE)

[1]Department of Electrical and Electronics Engineering, NMAM Institute of Technology (NMAMIT), Nitte (Deemed to be University), Nitte, Karnataka 574110, India
[2]Department of Electrical and Electronics Engineering, National Institute of Technology Karnataka (NITK), Surathkal, Karnataka 575025, India

Corresponding author: Krishna Rao (krishnarao270@gmail.com)

ABSTRACT The conventional Prony algorithm, which is the most prominent power system ring-down mode identification method, fails if the test signal is noisy [with a signal-to-noise ratio (SNR) below 20 dB]. The performance of Prony algorithm can be improved through singular value decomposition (SVD)-based rank reduction of the data matrix. Principal eigenvector (PE)-Prony and total least squares (TLS)-Prony are two known formulations of SVD-augmented Prony algorithms. In both PE-Prony and TLS-Prony algorithms, the Toeplitz structure of the linear prediction data matrix is lost upon SVD-based noise filtering. On the other hand, structured total least squares (STLS)-Prony algorithm retains the Toeplitz structure even after SVD-based filtering and is hence expected to perform better. But a formulation of STLS-Prony algorithm for power systems is not available in the literature. Hence, the same is developed successfully in this paper. As a prelude to the formulation of STLS-Prony algorithm, PE-Prony and TLS-Prony analyses of power system signals are discussed in detail, bringing out their nuances. Further, case studies are carried out on some benchmark power systems to demonstrate that all the three algorithms work successfully even at an SNR of 1 dB when the test signal has only inter-area modes. It is also shown that the performance of STLS-Prony algorithm is superior when the test signal has a highly damped local mode. Further, it is illustrated that by virtue of structure-preserving property, STLS-Prony algorithm is endowed with a unique filtering attribute although it has a longer execution time.

INDEX TERMS Principal eigenvector (PE)-Prony, total least squares (TLS)-Prony, structured total least squares (STLS)-Prony algorithm, power system ring-down mode identification.

I. INTRODUCTION

The ever-increasing demand for electric power has resulted in heavy loading of the tie-lines between different areas of the power grid, and this has led to a rise in occurrence of inter-area oscillations. These oscillations have the potential to disrupt a large portion of the synchronized grid and hence must be detected in the incipient stage. This was not possible with the conventional power system monitoring apparatus, namely, supervisory control and data acquisition (SCADA) system, which was designed basically for monitoring steady-state parameters. The inter-area oscillations have a frequency of up to 1 Hz and hence to visualize these, the data acquisition rate has to be more than 2 samples/s as per Nyquist/Shannon sampling theorem whereas the typical data acquisition rate of the SCADA system is only 1 sample/s [1]. In such cases, to detect the onset of the oscillations, one has to resort to detailed small-signal stability analysis, which is a time-consuming process and hence is not useful in real-time control. Moreover, the efficacy of small-signal stability analysis depends on the accuracy of the parametric data and adequacy of the modeling details of different power system components such as synchronous machines, network elements, controllers and loads. The parameters of the system components change with age, and if these are not updated regularly, then simulation results may not correspond to prevalent power system condition. Further, it is primarily the controllers that produce negative damping [2]. Hence if the

The associate editor coordinating the review of this manuscript and approving it for publication was Nagesh Prabhu.

controllers are not modeled with adequate details, analytical studies may fail to indicate the presence of an unstable mode, as observed during the simulation study carried out for the investigation of western systems coordinating council (WSCC) failure of August 10, 1996 [3].

It is in this context that the recent advent of wide area measurement systems (WAMS) has proved to be a game-changer. The basic building block of the WAMS is phasor measurement unit (PMU), which has satellite-based precise time stamping and reports the phasor frames at a much higher rate of $10-120$ frames/s to the control center [4]. Thus, WAMS is endowed with the unique ability to visualize the power system oscillations directly. Being dependent only on measurements, this does not require the modeling details of the system, and being available online, this is useful for real-time control. However, these oscillatory signals often have multiple modal components, and hence their frequency and decrement factor cannot be determined by mere visual inspection. In fact, if an oscillatory signal has two components – one slightly positively damped and the other slightly negatively damped, it may not be possible to detect the presence of the negatively damped component by visual inspection of a limited-time window. This has necessitated application of measurement-based mode identification algorithms.

Normally, the onset of conspicuous power system oscillations is preceded by a major event such as tripping of a transmission line or a generator. The signals obtained in the immediate aftermath of such major events are termed as *ring-down signals*. The strength of the excited modes in these signals is characteristically high, and therefore mode identification is easy. Hence algorithms suitable to mode identification from the ring-down signals were first applied to power systems. The first and foremost among these is Prony analysis. The basic formulation of this was put forth by Prony way back in 1795 [5] for gas expansion analysis. Subsequently, Hildebrand refined this algorithm by incorporating least-squares principle [6]. The version of the algorithm thus proposed is known as *Least-Squares (LS)-Prony* or *extended Prony algorithm*. However, it was only after the advent of digital computer that Prony analysis found practical application. The first such application was published in the area of radar signal analysis in 1970s [7].

Implementation of Prony algorithm is straightforward whenever the number of component modes in the oscillatory signal is known in advance. However, one characteristic problem with power system oscillatory signals is that the number of component modes is not known. In fact, an integrated power system is a very high order system, but the number of electromechanical modes excited upon occurrence of a major disturbance is far less. There is no universally accepted guideline as to how to determine the number of excited modes. In view of this, Prony analysis is often implemented iteratively [8], [9], [10]. In such an approach, the number of excited modes (also called 'order of linear prediction – OLP' or 'reduced model order') is initially taken very low and increased progressively until the reconstructed signal matches with the original signal.

This iterative Prony formulation suffers from the following two drawbacks: (i) prolonged algorithm execution, particularly in case of signals obtained from large systems (ii) inability to identify the modes if the signal is noisy with a signal-to-noise ratio (SNR) below 20 dB.

As compared to Prony method, the matrix pencil algorithm performs better from the viewpoint of mode identification from noisy signals due to incorporation of singular value decomposition (SVD) of the data matrix [11]. SVD serves as a panacea to both the above problems. On the one hand, SVD enables non-iterative determination of reduced model order. On the other, this also acts as a noise-filtering technique.

This intuits the possibility of improvement in the performance of Prony method by similar integration of SVD. The first successful attempt in this direction was made by Kumaresan and Tufts in the area of acoustic mode identification. They subject the data matrix to SVD and filter out the dyads corresponding to the lowest singular values. This method is known as *Principal Eigenvector (PE)-Prony algorithm* [12]. Further improvement was achieved in acoustic mode identification area by Rahman and Kai-Bor [13], who put forth total least squares (TLS) solution of the linear prediction equation instead of the normal least-squares solution. A comparative case study of application of PE-Prony and TLS-Prony algorithms to power systems has been reported by Zhou et al. [14]. Statistical analysis of TLS-Prony algorithms is discussed recently in [15]. However, a detailed discussion of PE- and TLS-Prony analyses of power system signals that brings out their finer points is not available in the literature, and hence the same is presented in this paper.

In both the above improvisations of Prony, the data matrix or the augmented data matrix is replaced by one of lower rank without retaining its Toeplitz structure. It is reasonable to expect that a lower rank matrix having the same Toeplitz structure as the given data matrix would further improve the performance of Prony algorithm [16]. The same is attempted in structured total least squares (STLS)-Prony algorithm proposed by Park et al. [17] and Lemmerling [18]. While [17] develops the general concept of STLS-Prony algorithm, [18] discusses STLS-Prony algorithm applications specifically in the areas of nuclear magnetic resonance spectroscopy and acoustic mode identification. In fact, structure-preserving rank reduction of a matrix continues to be an active area of research [19], [20]. However, STLS-Prony analysis of power system signals has not been reported anywhere. Although the general formulation of STLS-Prony algorithm is established, its customization to power systems entails considerable effort due to two peculiar attributes of the power system: (i) the number of component modes in a test signal is far less than the order of an integrated system and (ii) some of the modes may be negatively damped. Thus, formulation of STLS-Prony

algorithm for power system signals and examination of its performance relative to other SVD-augmented Prony algorithms are considered as worthwhile research problems and hence are explored in this paper.

The flow of the paper goes as follows: Section II contains a brief discussion on iterative Prony algorithm. This is followed by detailed presentations on PE-Prony algorithm and TLS-Prony algorithm in Sections III and IV respectively. Discussion on iterative Prony, PE-Prony and TLS-Prony algorithms in Sections II-IV serve as a logical prelude to the formulation of STLS-Prony algorithm for power system signals presented in Section V. While the contribution of the paper is primarily in the development of STLS-Prony algorithm for power system signals, the nuances of PE-Prony and TLS-Prony algorithms are also brought out. A thorough comparison of PE-, TLS- and STLS-Prony algorithms through a few case studies on benchmark power systems is taken up in Section VI while the conclusions follow in Section VII.

II. ITERATIVE-PRONY ALGORITHM

Before proceeding to discuss SVD-augmented Prony algorithms, it is in order to make a brief presentation on iterative formulation of Prony algorithm. (In fact, this is the summary of the detailed formulation put forth by the current authors in [9].) This involves the following steps:

Step 1 – linear prediction (LP) solution:

A power system ring-down signal comprises of dc exponentials of the form $A_i\ exp(\sigma_i t)$ and exponentially varying sinusoids of the form $A_i\ exp(\sigma_i t)\cos(\omega_i t+\phi_i)$.

One can express $\cos(\omega_i t+\phi_i)$ in terms of exponentials by applying Euler's theorem. Hence, a ring-down signal can be expressed as a sum of exponentials as:

$$y(t)=\sum_{i=1}^{n} B_i e^{\lambda_i t} \tag{1}$$

where n is the number of exponential components in the signal, B is the residue and complex frequency $\lambda=\sigma+j\omega$. (σ is decrement factor and ω is radian frequency. In case of exponentially varying dc, ω would be zero.)

Equation (1) can be represented in discrete domain as:

$$y[k]=\sum_{i=1}^{n} B_i z_i^k \tag{2}$$

where

$$z_i=e^{(\lambda_i \Delta t)} \tag{3}$$

and Δt is the sampling interval.

Equation (2) is the solution of the following linear prediction (LP) equation of order n:

$$y(k)=a_1\ y(k-1)+a_2\ y(k-2)+\ldots\ldots+a_n\ y(k-n) \tag{4}$$

In other words, a power system that produces a response given by (2) must obey the LP model of (4).

Changing the sampling index k from n to $(N-1)$, (4) can be expressed in matrix form as:

$$\underbrace{\begin{bmatrix} y(n-1) & y(n-2) & \ldots & y(0) \\ y(n-0) & y(n-1) & \ldots & y(1) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-2) & y(N-3) & \ldots & y(N-n-1) \end{bmatrix}}_{\boldsymbol{Y}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}}_{\boldsymbol{a}} = \underbrace{\begin{bmatrix} y(n+0) \\ y(n+1) \\ \vdots \\ y(N-1) \end{bmatrix}}_{\boldsymbol{y}} \tag{5}$$

The linear prediction coefficients $a_1, a_2, \ldots, a_n$ are obtained as the least-squares solution of (5) as signal sample values y are known. However, the number of samples, N has to be greater than twice the number of components, n in order that (5) constitutes an overdetermined system i.e., $N>2n$.

NOTE: The data matrix $\boldsymbol{Y}$ in (5) is Toeplitz as all the elements in a given diagonal represent the same sample. It is important to recognize that this is true regardless of whether the test signal y is noisy or noiseless.

Step 2 – polynomial rooting:

The characteristic equation of (4) is:

$$z^n-\left(a_1\ z^{n-1}+a_2\ z^{n-2}+\ldots+a_n\ z^0\right)=0 \tag{6}$$

The roots of (6) give n different values for z.

Step 3 – determination of continuous-time eigenvalues:

The continuous-time eigenvalues λ_i are then obtained from discrete time eigenvalues z_i using (3). The real- and imaginary-part of λ_i yield the decrement factor σ_i and the radian frequency ω_i of component i respectively.

Step 4 – Vandermonde solution:

Again, by varying k from 0 to $(N-1)$, (2) is represented in matrix form as:

$$\underbrace{\begin{bmatrix} z_1^0 & z_2^0 & \ldots & z_n^0 \\ z_1^1 & z_2^1 & \ldots & z_n^1 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{N-1} & z_2^{N-1} & \ldots & z_n^{N-1} \end{bmatrix}}_{\boldsymbol{Z_1}} \underbrace{\begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix}}_{\boldsymbol{b}} = \underbrace{\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix}}_{\boldsymbol{y}} \tag{7}$$

One gets residues B_i by least-squares solution of (7). (Please note that $\boldsymbol{Z_1}$ is a Vandermonde matrix.)

From B_i, amplitude A_i and phase-angle ϕ_i are obtained.

With the knowledge of amplitude, phase-angle, decrement factor and radian frequency of different component modes, the signal can be reconstructed.

Step 5 – choice of order of linear prediction (OLP), n:

Implementation of the Prony algorithm with the above four steps is straightforward if OLP (order of linear prediction) n (which also represents the number of components in the signal) is known. However, in case of power system ring-down signals, n is not known in advance. Hence in the iterative Prony method, the value of OLP, n is initially taken very low, and the reconstructed signal obtained with this value of n is compared against the original signal. If the

reconstructed signal does not match with the original signal, then n is increased in steps until satisfactory matching is attained.

This iterative approach needs a *fitness metric* to decide the extent of matching between the original and the reconstructed signals. Signal-to-estimation-error ratio (SER) [often called 'signal-to-noise Ratio (SNR)'] is the metric normally employed. The definition of SER is [8]:

$$\text{SER} = 20\log_{10}\frac{\text{rms}(y)}{\text{rms}(y-\widehat{y})} \tag{8}$$

where y is the original signal and
$\widehat{y}$ is the reconstructed (or the estimated) signal.

As the matching between the reconstructed signal and the original signal improves, SER increases. An SER of 30 dB normally ensures satisfactory matching between the original and the reconstructed signals [21], [22] provided the signal is subjected to detrending [23] prior to being taken up for mode identification.

NOTE: Detrending is the process of removing the trend of the signal. This is a pre-requisite for successful identification of oscillatory modes, which is crucial for detecting onset of instability. The trend can be any of the following: simple DC offset, an inclined straight line, a quadratic curve or an exponentially varying DC.

An alternative fitness metric for Prony analysis is mean absolute percentage error (MAPE) [9], [24], which is defined as:

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i-\widehat{y}_i}{y_i}\right| \tag{9}$$

where

N number of data points;
y_i sample value of the original signal;
$\widehat{y}_i$ sample value of the reconstructed signal.

As the reconstructed signal matches better with the original signal, MAPE decreases. When MAPE value reaches down to 0.1, matching between the original- and the reconstructed-signal is found to be satisfactory in general.

Interestingly, it is noted in [9] and [10] that in some cases even when the mode identification is satisfactory, either SER has not gone above the threshold of 30 dB or MAPE has not fallen below the threshold of 0.1. This leads to unnecessary iterations if SER *alone* or MAPE *alone* is employed as the fitness metric. Hence 'MAPE combined with SER' is proposed as a superior fitness metric so that iteration can be terminated if either MAPE or SER reaches the threshold [9].

Step 6 – selection of dominant modes:

Prony algorithm throws up a few spurious modes. Therefore it is imperative to deploy a yardstick to sift the dominant component modes. Energy of the signal is generally adopted as such a yardstick [25]. Here energy is the sum of the squares of the instantaneous values of the signal component over the entire analysis window. Further, 10% of the energy of the most dominant mode is fixed as the threshold of dominance i.e., modes with energy above this threshold are considered dominant while those below the threshold are ignored.

III. PRINCIPAL EIGENVECTOR (PE)-PRONY ALGORITHM

The oldest among the SVD-augmented Prony algorithms is PE-Prony algorithm, which is enunciated in the following lines.

As mentioned in Section II, the first step in Prony analysis is the solution of linear prediction equation. (The basic difference among the three SVD-augmented Prony algorithms lies in this step.)

The general approach in SVD-augmented Prony algorithms is to initially assume the number of component modes in the signal as very high and then to determine the actual number of component modes n by SVD of the data matrix. Accordingly, in PE-Prony algorithm, the number of component modes (or the OLP) initially assumed is denoted as r and is taken roughly equal to one third of the samples in the signal analysis window. That is,

$$r = \text{round}\left(\frac{N}{3}\right) \tag{10}$$

Thus, replacing n by r in (5), one gets

$$\underbrace{\begin{bmatrix} y(r-1) & y(r-2) & \dots & y(0) \\ y(r-0) & y(r-1) & \dots & y(1) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-2) & y(N-3) & \dots & y(N-r-1) \end{bmatrix}}_{\boldsymbol{Y}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{bmatrix}}_{\boldsymbol{a}} = \underbrace{\begin{bmatrix} y(r) \\ y(r+1) \\ \vdots \\ y(N-1) \end{bmatrix}}_{\boldsymbol{y}} \tag{11}$$

where

$\boldsymbol{Y}$ data matrix of size $(N-r)\times r$;
$\boldsymbol{a}$ linear prediction coefficient vector of length r;
$\boldsymbol{y}$ observation vector of length $(N-r)$.

The next step in PE-Prony algorithm is to subject the data matrix $\mathbf{Y}$ to SVD as given below [26].

$$[\boldsymbol{Y}]_{(N-r)\times(r)} = [\boldsymbol{P}]_{(N-r)\times(N-r)} \quad [\boldsymbol{D}]_{(N-r)\times r} \quad [\boldsymbol{Q}^T]_{(r\times r)} \tag{12}$$

NOTE:

- In (12), the square matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ are orthogonal and comprise respectively of left- and right-singular-vectors of $\boldsymbol{Y}$. On the other hand, matrix $\boldsymbol{D}$ has the same size as $\boldsymbol{Y}$ i.e., $[(N-r)\times r]$. Further, $\boldsymbol{D}$ is a diagonal matrix and the number of diagonal elements equals r. These diagonal elements are either positive or zero, appear in the decreasing order and represent the singular values of $\boldsymbol{Y}$. The number of nonzero singular values is equal to the rank of the given matrix $\boldsymbol{Y}$.
- It is proved in literature that if the signal is noiseless, then the data matrix $\boldsymbol{Y}$ would have a rank n equal to the actual number of components in the signal [17], [27]. In such an event, out of r singular values of $\boldsymbol{Y}$, only the first n singular values would be nonzero. On the

other hand, when the signal is noisy, all the r singular values would most likely be nonzero. However, in such a case, the first n singular values would be significantly large and the remaining $(r-n)$ singular values would be very small. Hence it is logical to surmise that the large singular values correspond to actual component modes in the given signal whereas the smaller ones are due to noise [28]. Accordingly, by segregating the large singular values one can decide the actual number of component modes in the given noisy signal. This segregation becomes easy if a clear gap can be observed at one place when the singular values are arranged in the descending order; singular values above the gap can be regarded as corresponding to the actual modal components of the signal while those below the gap as corresponding to noise. However, such a clear gap is not visible at times. In such cases, segregation of the singular values is carried out by fixing a threshold in terms of the largest singular value. The most common threshold selected is 10^{-3} times the largest singular value [29]. The number of singular values above the threshold is counted as n, which gives the number of component modes in the given signal. Thus, the model order, n is determined directly here without resorting to any iterative procedure.

- Hence, n is considered as the numerical rank of $\boldsymbol{Y}$ whereas r is its actual rank.
- Alternatively, (12) can be rearranged as:

$$[\boldsymbol{Y}]_{(N-r)\times(r)} = \sum_{i=1}^{r} d_i\, \boldsymbol{p_i} \boldsymbol{q_i}^T \tag{13}$$

Each product $(d_i\, \boldsymbol{p_i} \boldsymbol{q_i}^T)$ in (13) represents a rank-1 matrix of same size as $\boldsymbol{Y}$, which is called a *dyad*. It is the outer product of the i^{th} left-singular vector $\boldsymbol{p_i}$ and the i^{th} right-singular vector $\boldsymbol{q_i}$, scaled by the i^{th} singular value d_i. Thus, (13) represents the 'dyadic decomposition' of matrix $\boldsymbol{Y}$ [29].

- Hence the matrix $\boldsymbol{Y_n}$, which is the rank-n equivalent of matrix $\boldsymbol{Y}$ can be represented as:

$$\boldsymbol{Y_n} = \sum_{i=1}^{n} d_i\, \boldsymbol{p_i} \boldsymbol{q_i}^T \tag{14}$$

That is, $\boldsymbol{Y_n}$ is obtained from $\boldsymbol{Y}$ by retaining the first n dyads of $\boldsymbol{Y}$. The remaining dyads, which are supposed to represent the noise in the signal are filtered out. Effectively, matrix $\boldsymbol{Y_n}$ is obtained from matrix $\boldsymbol{Y}$ by setting its $(r-n)$ singular values below the threshold to zero.

- Although $\boldsymbol{Y}$ is a Toeplitz matrix, the individual dyads do not have the Toeplitz structure in general. As such, the matrix $\boldsymbol{Y_n}$, comprising the selected dyads of $\boldsymbol{Y}$, is not Toeplitz in general.
- The matrix $\boldsymbol{Y_n}$ can be represented in terms of matrices $\boldsymbol{P}, \boldsymbol{D}$ and $\boldsymbol{Q}$ as:

$$[\boldsymbol{Y_n}]_{(N-r)\times r} = [\boldsymbol{P_n}]_{(N-r)\times n}\ \ [\boldsymbol{D_n}]_{(n\times n)}\ \ [\boldsymbol{Q_n}^T]_{(n\times r)}$$

where
$$\begin{aligned} \boldsymbol{P_n} &= \boldsymbol{P}(:, 1:n) \\ \boldsymbol{D_n} &= \boldsymbol{D}(1:n, 1:n) \\ \boldsymbol{Q_n} &= \boldsymbol{Q}(:, 1:n) \end{aligned} \tag{15}$$

That is, $\boldsymbol{P_n}$ contains the first n columns of matrix $\boldsymbol{P}$ and $\boldsymbol{Q_n}$ contains the first n columns of matrix $\boldsymbol{Q}$. On the other hand, $\boldsymbol{D_n}$ contains the first n rows and first n columns of matrix $\boldsymbol{D}$.

As $\boldsymbol{Y}$ is replaced by its reduced-rank equivalent $\boldsymbol{Y_n}$, the linear prediction equation changes to:

$$\boldsymbol{Y_n}\ \ \boldsymbol{a} = \boldsymbol{y} \tag{16}$$

From this, linear prediction coefficient vector $\boldsymbol{a}$ is got as:

$$\boldsymbol{a} = (\boldsymbol{Y_n})^+\ \ \boldsymbol{y} \tag{17}$$

where $(\boldsymbol{Y_n})^+$ is the Moore-Penrose pseudoinverse of $\boldsymbol{Y_n}$, and this can be calculated as:

$$(\boldsymbol{Y_n})^+ = \sum_{i=1}^{n} \frac{1}{d_i} \boldsymbol{q_i}\, \boldsymbol{p_i}^T \tag{18}$$

Hence (17) can be written as [12]:

$$\boldsymbol{a} = \sum_{i=1}^{n} \boldsymbol{q_i} \left(\frac{\boldsymbol{p_i}^T \boldsymbol{y}}{d_i} \right) \tag{19}$$

NOTE:

1) The singular values, d_i appear in the denominator on the RHS of (19). Hence the lowest singular values perturb the solution $\boldsymbol{a}$ highest. By filtering out these, PE-Prony algorithm minimizes this perturbation.
2) The rank-reduced data matrix $\boldsymbol{Y_n}$ can be seen as a projection of the original data matrix $\boldsymbol{Y}$ on to the columns of $\boldsymbol{P_n}$ since

$$\boldsymbol{Y_n} = \boldsymbol{P_n} \boldsymbol{P_n}^T \boldsymbol{Y} \tag{20}$$

where $\boldsymbol{P_n}^T$ (transpose of matrix $\boldsymbol{P_n}$) itself is the Moore-Penrose pseudoinverse of $\boldsymbol{P_n}$.
The least-squares approach assumes that the data matrix is error-free and all the error is contained in the observation vector. Hence the observation vector is projected on to the column space of the data matrix. Thus, in iterative-Prony, the observation vector $\boldsymbol{y}$ is projected on to the column space of data matrix $\boldsymbol{Y}$ spanned by first r columns of matrix $\boldsymbol{P}$. On the other hand, in PE-Prony, $\boldsymbol{y}$ is projected on to column space of rank-reduced data matrix $\boldsymbol{Y_n}$ spanned by the columns of matrix $\boldsymbol{P_n}$, which are, in fact, the first n columns of $\boldsymbol{P}$. Hence column space of $\boldsymbol{Y_n}$ is a subspace of the column space of $\boldsymbol{Y}$. Thus, in PE-Prony, the observation vector $\boldsymbol{y}$ is projected on to a smaller subspace as compared to iterative-Prony. Hence perturbation of the observation vector is generally greater in PE-Prony as compared to iterative-Prony. Also, the projection of observation vector is generally shorter in PE-Prony as compared to iterative-Prony.

3) Therefore, the LP solution vector obtained in PE-Prony is shorter than that in iterative-Prony because it involves a shorter projection of the observation vector and it leaves out the smallest singular values that perturb the projected observation vector highest.

Determination of linear prediction coefficient vector $\boldsymbol{a}$ completes the first step in PE-Prony analysis. Other steps are largely similar to those in the iterative Prony algorithm.

Hence PE-Prony algorithm can be summarized as follows:

- **Step 1 – determination of linear prediction coefficients:**
  Signal samples are arranged in the form of LP equation (11), taking the initially assumed number of modes r equal to one third the number of samples i.e., $(N/3)$. The data matrix $\boldsymbol{Y}$ is subjected to SVD and its numerical rank n is determined as the number of singular values of $\boldsymbol{Y}$ above the threshold of 10^{-3} times its largest singular value. Using the first n singular values and the corresponding left-and right-singular vectors, the linear prediction coefficient vector $\boldsymbol{a}$ is obtained according to (19).
- **Steps 2-4:** These steps are identical to those of iterative Prony. These involve determination of discrete eigenvalues z_i (by polynomial rooting) and thence the continuous-time eigenvalues λ_i. From λ_i, the decrement factor σ_i and radian frequency ω_i of different component modes are obtained. Thereafter, residues B_i are decided through Vandermonde solution, which, in turn, yield amplitude A_i and phase-angle ϕ_i of the components.
- **Step 5:** Further, as in the iterative-Prony method, dominant modes are selected based on their energy, taking 10% of energy of the most dominant mode as the threshold.
- **Step 6:** From the details of amplitude, phase-angle, decrement factor and radian frequency, signal can be reconstructed and compared with the original signal. The extent of matching between the original signal and the reconstructed signal can be gauged using fitness metrics SER and MAPE. However, this is an optional step which is not integral to the successful working of the algorithm.

IV. TOTAL LEAST SQUARES (TLS) - PRONY ALGORITHM

In the conventional least-squares approach, one presumes that the observation vector alone is erroneous and the data matrix is error-free. On the other hand, the *total-least-squares* approach consists in treating both the data matrix and the observation vector as erroneous. Such an approach is all the more logical in linear prediction-based algorithms such as Prony analysis as most of the samples that constitute the observation vector figure in the data matrix as well [30].

Reconsider the linear prediction equation given in Section II,

$$\boldsymbol{Y}\boldsymbol{a} = \boldsymbol{y} \tag{21}$$

Since the above equation normally represents an overdetermined system, the equality sign does not hold in a strict sense. In the conventional least-squares approach, one perturbs the observation vector, $\boldsymbol{y}$ by the shortest possible distance $\boldsymbol{g}$ so that the vector $(\boldsymbol{y}+\boldsymbol{g})$ lies in the column space of matrix $\boldsymbol{Y}$. Therefore the solution vector, $\hat{\boldsymbol{a}}$ satisfies the equation:

$$\boldsymbol{Y}\hat{\boldsymbol{a}} = \boldsymbol{y}+\boldsymbol{g} \tag{22}$$

The equality sign holds in a strict sense after optimal perturbation of the observation vector in (22). Further, if this equation is solved using Moore-Penrose pseudoinverse, then the obtained solution vector, $\hat{\boldsymbol{a}}$ will be the shortest possible one.

On the other hand, the total-least-squares approach expects that both the data matrix, $\boldsymbol{Y}$ and the observation vector, $\boldsymbol{y}$ may contain errors and hence it perturbs both of these such that

$$(\boldsymbol{Y}+\boldsymbol{E})\hat{\boldsymbol{a}} = (\boldsymbol{y}+\boldsymbol{g}) \tag{23}$$

In (23), $\boldsymbol{E}$ and $\boldsymbol{g}$ represent the minimum possible perturbations of data matrix, $\boldsymbol{Y}$ and observation vector, $\boldsymbol{y}$ respectively so that the equality sign applies in a strict sense. In other words, the perturbed observation vector, $(\boldsymbol{y}+\boldsymbol{g})$ is now required to lie in the column space of the perturbed data matrix, $(\boldsymbol{Y}+\boldsymbol{E})$.

Equation (23) can be rewritten as a homogeneous equation as:

$$[(\boldsymbol{Y}+\boldsymbol{E}) \mid (\boldsymbol{y}+\boldsymbol{g})] \begin{bmatrix} \hat{\boldsymbol{a}} \\ \hline -1 \end{bmatrix} = \boldsymbol{0} \tag{24}$$

Equation (24) can, in turn, be rearranged as:

$$([\boldsymbol{Y}|\boldsymbol{y}]+[\boldsymbol{E}|\boldsymbol{g}]) \begin{bmatrix} \hat{\boldsymbol{a}} \\ \hline -1 \end{bmatrix} = \boldsymbol{0} \tag{25}$$

or

$$(\boldsymbol{B}+\boldsymbol{F})\boldsymbol{z} = \boldsymbol{0} \tag{26}$$

where

$$[\boldsymbol{B}]_{(N-r)\times(r+1)} = [\boldsymbol{Y}|\boldsymbol{y}] \tag{27}$$

$$[\boldsymbol{F}]_{(N-r)\times(r+1)} = [\boldsymbol{E}|\boldsymbol{g}] \tag{28}$$

$$\boldsymbol{z} = \begin{bmatrix} \hat{\boldsymbol{a}} \\ \hline -1 \end{bmatrix}_{(r+1)\times 1} \tag{29}$$

The TLS approach looks for a solution vector, $\hat{\boldsymbol{a}}$ that minimizes the Frobenius norm of matrix $\mathbf{F}$. The fact that $(\boldsymbol{y}+\boldsymbol{g})$ should belong to the column space of $(\boldsymbol{Y}+\boldsymbol{E})$ implies that matrix $(\boldsymbol{B}+\boldsymbol{F})$ should be rank-deficient. On the other hand, since $\boldsymbol{y}$ does not lie in the column space of $\boldsymbol{Y}$, $\boldsymbol{B}$ is a full-rank matrix provided $\boldsymbol{Y}$ is a full-rank matrix. That is,

$$\text{rank}(\boldsymbol{B}+\boldsymbol{F}) = \text{rank}(\boldsymbol{Y}+\boldsymbol{E}) = \text{rank}(\boldsymbol{Y}) \tag{30}$$

$$\text{rank}(\boldsymbol{B}) = \text{rank}(\boldsymbol{Y})+1 \tag{31}$$

Thus, the perturbation matrix $\boldsymbol{F}$ must be so selected that the rank of $(\boldsymbol{B}+\boldsymbol{F})$ must be one less than that of $\boldsymbol{B}$. This suggests

the following SVD-based approach for the solution of total-least-squares problem:

To begin with, consider the SVD of matrix $\boldsymbol{B}$:

$$[\boldsymbol{B}]_{(N-r)\times(r+1)} = [\boldsymbol{P}]_{(N-r)\times(N-r)} \quad [\boldsymbol{D}]_{(N-r)\times(r+1)} \quad [\boldsymbol{Q}^T]_{(r+1)\times(r+1)} \tag{32}$$

This can be represented equivalently as follows:

$$\boldsymbol{B} = \sum_{i=1}^{r+1} d_i \, \boldsymbol{p_i} \boldsymbol{q_i}^T \tag{33}$$

Equation (33) is the *dyadic decomposition* of matrix $\boldsymbol{B}$. In this, d_i represents the i^{th} singular value of $\boldsymbol{B}$ whereas $\boldsymbol{p_i}$ represents the i^{th} left singular vector and $\boldsymbol{q_i}$ represents i^{th} right singular vector of $\boldsymbol{B}$. The rank-1 matrix obtained from the product $(d_i \ \boldsymbol{p_i} \ \boldsymbol{q_i}^T)$ is the i^{th} dyad. Since $\boldsymbol{p_i}$ and $\boldsymbol{q_i}$ are normal vectors, the Frobenius norm of i^{th} dyad is equal to d_i. It is worth noting that the matrix obtained by deducting anyone dyad from $\boldsymbol{B}$ will have a rank one less than that of $\boldsymbol{B}$. Hence every dyad of $\boldsymbol{B}$ is a candidate for the perturbation matrix $\boldsymbol{F}$. However, $\boldsymbol{F}$ is required to meet the additional criterion that it must have minimum Frobenius norm. Among the dyads, it is the last dyad that has the minimum Frobenius norm – equal to the lowest singular value, d_{r+1} – and hence this qualifies as the right candidate for the perturbation matrix $\boldsymbol{F}$.

$$[\boldsymbol{F}]_{(N-r)\times(r+1)} = -(d_{r+1} \, \boldsymbol{p_{r+1}} \, \boldsymbol{q}_{r+1}^T) \tag{34}$$

With this choice of matrix $\boldsymbol{F}$,

$$\boldsymbol{B}+\boldsymbol{F} = \sum_{i=1}^{r} d_i \, \boldsymbol{p_i} \boldsymbol{q_i}^T \tag{35}$$

Thus, matrix $\boldsymbol{B}+\boldsymbol{F}$ is obtained by retaining only the first r dyads of matrix $\boldsymbol{B}$ while dropping its $(r+1)^{th}$ dyad. In other words, addition of matrix $\boldsymbol{F}$ to the matrix $\boldsymbol{B}$ is tantamount to setting the $(r+1)^{th}$ singular value of matrix $\boldsymbol{B}$, d_{r+1} to zero. So, the matrix $\boldsymbol{B}+\boldsymbol{F}$ has only r nonzero singular values and hence, has a rank r, but it has the same left- and right-singular vector matrices as $\boldsymbol{B}$, namely, $\boldsymbol{P}$ and $\boldsymbol{Q}$. Hence, the first r columns of the right-singular vector matrix, $\boldsymbol{Q}$ serve as an orthonormal basis to the row space of $(\boldsymbol{B}+\boldsymbol{F})$ and the $(r+1)^{th}$ column serves as a basis to the null space of $(\boldsymbol{B}+\boldsymbol{F})$. Therefore, this column can be a solution to the homogeneous equation (25) since such a solution has to lie in the null space of $(\boldsymbol{B}+\boldsymbol{F})$. However, there is an additional criterion in (25) that the last element of the solution vector, z be equal to -1, and in order to meet this criterion, the $(r+1)^{th}$ right singular vector $\boldsymbol{q}_{r+1}$ has to be divided by the negative of its last element.

$$z = \begin{bmatrix} \hat{a} \\ \hline -1 \end{bmatrix} = \frac{-1}{(\boldsymbol{q}_{r+1})_{r+1}} \begin{bmatrix} (\boldsymbol{q}_{r+1})_1 \\ (\boldsymbol{q}_{r+1})_2 \\ \vdots \\ (\boldsymbol{q}_{r+1})_r \\ (\boldsymbol{q}_{r+1})_{r+1} \end{bmatrix} \tag{36}$$

Hence

$$[\hat{a}] = \frac{-1}{(\boldsymbol{q}_{r+1})_{r+1}} \begin{bmatrix} (\boldsymbol{q}_{r+1})_1 \\ (\boldsymbol{q}_{r+1})_2 \\ \vdots \\ (\boldsymbol{q}_{r+1})_r \end{bmatrix} \tag{37}$$

As already indicated, the gist of the above approach lies in ensuring that the rank of matrix $(\boldsymbol{B}+\boldsymbol{F})$ equals the rank of matrix $\boldsymbol{Y}$. This works satisfactorily if $\boldsymbol{Y}$ has a low condition number. This implies that the singular values of $\boldsymbol{Y}$ are sufficiently close. This would be true only when the signal considered is noiseless and the actual number of component modes in the signal is equal to the assumed value, r. On the other hand, in case of noisy signals, the lowest singular values pertaining to noise are very low. In regard to mode identification of such noisy signals, the practice followed in Section III is to resort to SVD-based noise filtering wherein all the singular values below a certain threshold (which is selected as 10^{-3} times the largest singular value) are set to zero. It is reasonable to expect that the same approach should be adopted here as well. If n be the number of singular values of $\boldsymbol{Y}$ above the threshold, then a rank-n matrix $(\boldsymbol{B}+\boldsymbol{F})$ has to be obtained by setting the remaining $[(r+1)-n]$ singular values of $\boldsymbol{B}$ to zero. Thus, the rank of $(\boldsymbol{B}+\boldsymbol{F})$ is forced to equal *not the actual rank, but the numerical rank of* $\boldsymbol{Y}$.

The rank-n truncated matrix obtained in such a case will be:

$$(\boldsymbol{B}+\boldsymbol{F})_{\boldsymbol{n}} = \sum_{i=1}^{n} d_i \, \boldsymbol{p_i} \boldsymbol{q_i}^T \tag{38}$$

That is, matrix $(\boldsymbol{B}+\boldsymbol{F})_{\boldsymbol{n}}$ is obtained from matrix $\boldsymbol{B}$ by retaining only the first n dyads and dropping the last $(r+1-n)$ dyads.

With this, the homogeneous equation (26) can be rewritten as:

$$(\boldsymbol{B}+\boldsymbol{F})_{\boldsymbol{n}} \ z = \boldsymbol{0} \tag{39}$$

Now the first n right singular vectors, $\boldsymbol{q_1}, \boldsymbol{q_2}, \ldots, \boldsymbol{q_n}$ serve as a basis for the row space of $(\boldsymbol{B}+\boldsymbol{F})_{\boldsymbol{n}}$ whereas the remaining $[(r+1)-n]$ right singular vectors serve as a basis for the null space of $(\boldsymbol{B}+\boldsymbol{F})_{\boldsymbol{n}}$. Hence each of these last $[(r+1)-n]$ right singular vectors, $\boldsymbol{q_{n+1}}, \boldsymbol{q_{n+2}}, \ldots, \boldsymbol{q_{r+1}}$ becomes a candidate for the solution vector z, upon being suitably scaled. In fact, any vector that lies in the subspace spanned by $\boldsymbol{q_{n+1}}, \boldsymbol{q_{n+2}}, \ldots, \boldsymbol{q_{r+1}}$ is a candidate solution. However, one has to pick z such that its component $\hat{\boldsymbol{a}}$ has minimum length. [It can be noted from (29) that $\hat{\boldsymbol{a}}$ contains all the elements of z except the last one.] To accomplish this, [31] suggests the approach given below.

Arrange the last $[(r+1)-n]$ right singular vectors into a matrix as follows:

$$\boldsymbol{Q_1} = [\boldsymbol{q_{n+1}} \ \ \boldsymbol{q_{n+2}} \cdots \boldsymbol{q_r} \ \ \boldsymbol{q_{r+1}}]_{(r+1)\times(r+1-n)} \tag{40}$$

Multiply $\boldsymbol{Q_1}$ by a Householder matrix $\boldsymbol{H}$ such that in the last row of the resultant matrix $\boldsymbol{Q_2}$, all the elements except the last

one are zeros.

$$\boldsymbol{Q_2} = [\boldsymbol{Q_1}]_{(r+1)\times(r+1-n)}[\boldsymbol{H}]_{(r+1-n)\times(r+1-n)} = \left[\begin{array}{ccc|c} * & * & \dots & * & \\ * & * & \dots & * & \\ \vdots & \vdots & \ddots & \vdots & \boldsymbol{m} \\ * & * & \dots & * & \\ \hline 0 & 0 & \cdots\cdots & 0 & s \end{array}\right]_{(r+1)\times(r+1-n)} \tag{41}$$

The Householder transformation is a reflection which retains the lengths and the relative angles of the column vectors unchanged. Hence, all the columns of $\boldsymbol{Q_2}$ are orthonormal, have $(r+1)$ elements each and span the same column space as $\boldsymbol{Q_1}$. (That is, matrices $\boldsymbol{Q_1}$ and $\boldsymbol{Q_2}$ are right-equivalent [32].) Hence each column of $\boldsymbol{Q_2}$ is a candidate for z. Now if one constitutes a sub-vector by picking the first r elements of the first column of $\boldsymbol{Q_2}$, then it has a length equal to unity. This is because it comprises all the elements of a normal vector except the last one, which is zero. The same is true of the sub-vectors constituted by the first r elements of each column of $\boldsymbol{Q_2}$ except the last column. However, the sub-vector constituted by the first r elements of the last column of $\boldsymbol{Q_2}$ has a length less than unity. (This is because it includes all the elements of a normal vector except the last one, which is nonzero.) Therefore, among the different sub-vectors thus obtained, the sub-vector constituted by the first r elements of the last column of $\boldsymbol{Q_2}$ qualifies as the right candidate for $\hat{\boldsymbol{a}}$ as it has the shortest length. Hence the last column of $\boldsymbol{Q_2}$ can be picked as the solution vector z of the homogeneous equation (39) provided it is scaled suitably so that its last element equals -1.

Therefore

$$z = \begin{bmatrix} \hat{\boldsymbol{a}}_{TLS} \\ \hline -1 \end{bmatrix} = -\frac{1}{s}\begin{bmatrix} \boldsymbol{m} \\ \hline s \end{bmatrix} \tag{42}$$

Hence it follows that

$$\hat{\boldsymbol{a}}_{TLS} = -\frac{1}{s}[\boldsymbol{m}] \tag{43}$$

Note that $\hat{\boldsymbol{a}}_{TLS}$ requires only the last column of $\boldsymbol{Q_2}$, which is the product of $\boldsymbol{Q_1}$ and the last column of matrix $\boldsymbol{H}$. Thus for the TLS solution, it is required to compute only the last column of $\boldsymbol{H}$. This last column, designated as $\boldsymbol{h_{end}}$, is obtained simply by normalizing the last row of $\boldsymbol{Q_1}$ to a unit vector [13].

$$\boldsymbol{h_{end}} = \frac{1}{norm_{(r+1)}}\begin{bmatrix} (\boldsymbol{q_{n+1}})_{r+1} \\ (\boldsymbol{q_{n+2}})_{r+1} \\ \vdots \\ (\boldsymbol{q_r})_{r+1} \\ (\boldsymbol{q_{r+1}})_{r+1} \end{bmatrix} \tag{44}$$

where

$$norm_{(r+1)} = [(\boldsymbol{q_{n+1}})_{r+1}^2 + (\boldsymbol{q_{n+2}})_{r+1}^2 + \ldots + (\boldsymbol{q_r})_{r+1}^2 + (\boldsymbol{q_{r+1}})_{r+1}^2]^{\frac{1}{2}} \tag{45}$$

Hence

$$\begin{bmatrix} \boldsymbol{m} \\ \hline s \end{bmatrix} = \sum_{i=n+1}^{r+1} \frac{(\boldsymbol{q_i})_{r+1}}{norm_{(r+1)}}\boldsymbol{q_i} \tag{46}$$

It is evident that

$$s = \sum_{i=n+1}^{r+1} \frac{(\boldsymbol{q_i})_{r+1}^2}{norm_{(r+1)}} \tag{47}$$

Therefore

$$\hat{\boldsymbol{a}}_{TLS} = -\frac{1}{s}[\boldsymbol{m}] = -\left[\frac{\sum_{i=n+1}^{r+1}(\boldsymbol{q_i})_{r+1}\ \boldsymbol{q_i'}}{\sum_{i=n+1}^{r+1}(\boldsymbol{q_i})_{r+1}^2}\right] \tag{48}$$

where

$$\boldsymbol{q_i} = \begin{bmatrix} \boldsymbol{q_i'} \\ \hline (\boldsymbol{q_i})_{r+1} \end{bmatrix} \tag{49}$$

That is, $\boldsymbol{q_i'}$ is a sub-vector that includes all the elements except the last one of vector $\boldsymbol{q_i}$.

This completes the determination of linear prediction coefficients, which is the first step in Prony analysis. The remaining steps are identical to those given at the end of Section III.

Hence TLS-Prony algorithm can be summarized as follows:

- **Step 1 – determination of linear prediction coefficients:**
  Signal samples are arranged in the form of LP equation (11), taking the initially assumed number of modes r equal to one third the number of samples i.e., $(N/3)$. From (11), augmented data matrix is formed as: $\boldsymbol{B} = [\boldsymbol{Y}|\boldsymbol{y}]$. Matrix $\boldsymbol{B}$ is subjected to SVD and the matrix of right singular vectors $\boldsymbol{Q}$ is got. The LP coefficient vector is hence obtained as:

  $$\hat{\boldsymbol{a}}_{TLS} = -\frac{1}{s}[\boldsymbol{m}] = -\left[\frac{\sum_{i=n+1}^{r+1}(\boldsymbol{q_i})_{r+1}\ \boldsymbol{q_i'}}{\sum_{i=n+1}^{r+1}(\boldsymbol{q_i})_{r+1}^2}\right]$$

  where $\boldsymbol{q_i}$ is the i^{th} right singular vector of $\boldsymbol{B}$ and $\boldsymbol{q_i'}$ is a sub-vector that includes all the elements except the last one of vector $\boldsymbol{q_i}$. Further, n is the *numerical* rank of $\boldsymbol{Y}$ i.e., the number of singular values of $\boldsymbol{Y}$ above the threshold of 10^{-3} times its largest singular value.
- **Steps 2-6:** These steps are identical to the corresponding ones in PE-Prony algorithm elaborated in Section III.

V. STRUCTURED TOTAL LEAST SQUARES (STLS) - PRONY ALGORITHM

The data matrix $\boldsymbol{Y}$ in the linear prediction equation (11) has a Toeplitz structure. In PE-Prony and TLS-Prony algorithms, this data matrix is replaced by a lower-rank equivalent matrix obtained by SVD, which is supposed to represent the filtered ring-down signal. The equivalent matrix does not have a Toeplitz structure as SVD-based dyadic decomposition is

not a structure-preserving decomposition. However, it is evident from (5) in Section II that a noiseless ring-down signal also adheres to the linear prediction model, and hence the corresponding data matrix must also have a Toeplitz structure. Thus, one can expect the results to be more accurate if the rank-reduced data matrix retains the Toeplitz structure [16].

One approach suggested in this direction is the *signal enhancement through a composite property mapping algorithm* put forth by Cadzow [27]. In this, a rank-n approximation of the data matrix is first obtained using SVD (where n is the number of singular values above the threshold). This approximation does not preserve the Toeplitz structure. In order to re-establish the Toeplitz structure, the elements along each diagonal are replaced by their average value. Although the matrix thus obtained has a Toeplitz structure, it has a rank greater than n. Hence, SVD-based rank reduction and re-establishment of the Toeplitz structure are repeated iteratively until the $(n+1)^{th}$ singular value of the corrected data matrix falls below a pre-defined tolerance level. The approach based on Cazdows algorithm is reported to be suboptimal [33], [34], and hence the same is not adopted in this work.

Instead, an alternative approach is explored here in which this *structured total least squares* problem is cast as an optimization problem whose objective function is to obtain the rank-reduced data matrix directly as a Toeplitz matrix [10], [17], [18].

The conceptual idea behind rank reduction of a matrix is to add a perturbation matrix so that the resultant has a lower rank. Evidently, if the rank-reduced matrix is to have the same Toeplitz structure as the original matrix, then the perturbation matrix must also have a Toeplitz structure. In fact, it is not just the data matrix $\boldsymbol{Y}$ but also the augmented data matrix $[\boldsymbol{y}|\boldsymbol{Y}]$ that has a Toeplitz structure.

$$[\boldsymbol{y}|\boldsymbol{Y}]=\boldsymbol{T}=\left[\begin{array}{c|ccc} y(r) & y(r-1) & y(r-2) & \dots & y(0) \\ y(r+1) & y(r) & y(r-1) & \dots & y(1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \underbrace{y(N-1)}_{\boldsymbol{y}} & \underbrace{y(N-2)}_{\boldsymbol{Y}} & y(N-3) & \dots & y(N-r-1) \end{array}\right]$$

Thus, STLS consists in finding an augmented perturbation matrix $[\boldsymbol{g}|\boldsymbol{E}]$ of same dimension as $[\boldsymbol{y}|\boldsymbol{Y}]$ with the shortest Frobenius norm and Toeplitz structure.

Let

$$[\boldsymbol{g}|\boldsymbol{E}]=\left[\begin{array}{c|ccc} \eta_{p+1} & \eta_p & \eta_{p-1} & \dots & \eta_1 \\ \eta_{p+2} & \eta_{p+1} & \eta_p & \dots & \eta_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \underbrace{\eta_{p+m}}_{\boldsymbol{g}} & \underbrace{\eta_{p+m-1}}_{\boldsymbol{E}} & \eta_{p+m-2} & \dots & \eta_m \end{array}\right]$$

Number of rows in $\boldsymbol{E}$ = number of rows in $\boldsymbol{g} = m$.

Number of columns in $\boldsymbol{E} = p$.

It is to be noted that in a Toeplitz matrix, the number of *distinct* elements is equal to (number of columns + number of rows -1). In case of matrix $[\boldsymbol{g}|\boldsymbol{E}]$, it is equal to $(p+m)$.

(NOTE: Size of matrix $\boldsymbol{g}$ is same as that of matrix $\boldsymbol{y}$. Similarly, size of matrix $\boldsymbol{E}$ is $(m \times p)$. This is supposed to be equal to size of matrix $\boldsymbol{Y}$, which is $[(N-r)\times r]$. Hence, $m=(N-r)$ and $p=r$. Thus, m and p are variable names used only for convenience, whose values actually depend on N and r.)

Now define

$$\eta = \begin{bmatrix}\eta_1 & \eta_2 & \dots & \eta_{p+m}\end{bmatrix}^T \tag{50}$$

$$\alpha = \begin{bmatrix}\eta_1 & \eta_2 & \dots & \eta_{p+m-1}\end{bmatrix}^T \tag{51}$$

$$\boldsymbol{g} = \begin{bmatrix}\eta_{p+1} & \eta_{p+2} & \dots & \eta_{p+m}\end{bmatrix}^T \tag{52}$$

It may be noted that the vector α has only those η elements that are present in matrix $\boldsymbol{E}$.

Hence one can write

$$\alpha = \boldsymbol{P_0}\eta \tag{53}$$

$$\boldsymbol{g} = \boldsymbol{P_1}\eta \tag{54}$$

where

$$[\boldsymbol{P_0}]_{(p+m-1)\times(p+m)} = \begin{bmatrix}\boldsymbol{I}_{(p+m-1)\times(p+m-1)} & | & \boldsymbol{0}_{(p+m-1)\times 1}\end{bmatrix} \tag{55}$$

$$[\boldsymbol{P_1}]_{m\times(p+m)} = \begin{bmatrix}\boldsymbol{0}_{(m\times p)} & | & \boldsymbol{I}_{(m\times m)}\end{bmatrix} \tag{56}$$

($\boldsymbol{I}$ is an identity matrix and $\boldsymbol{0}$ is a null matrix.)

At this juncture, it is in order to define a *structured residual vector* $\hat{\boldsymbol{r}}$ as:

$$\hat{\boldsymbol{r}}(\eta, \boldsymbol{a}) = (\boldsymbol{y}+\boldsymbol{g}) - (\boldsymbol{Y}+\boldsymbol{E})\boldsymbol{a} \tag{57}$$

Since $(\boldsymbol{y}+\boldsymbol{g})$ lies in the column space of $(\boldsymbol{Y}+\boldsymbol{E})$, $\hat{\boldsymbol{r}}(\eta, a)$ should ideally be zero.

The task at hand is to minimize the Frobenius norm of the matrix $[\boldsymbol{g}|\boldsymbol{E}]$ i.e., $\|\boldsymbol{g}|\boldsymbol{E}\|_F$ subject to $\hat{\boldsymbol{r}} = \boldsymbol{0}$. This is a constrained minimization problem.

This minimization problem can be simplified by representing the Frobenius norm of the matrix $[\boldsymbol{g}|\boldsymbol{E}]$ as a scaled l_2 norm of vector η as follows:

$$\|\boldsymbol{g}|\boldsymbol{E}\|_F = \|\boldsymbol{L}\eta\|_2 \tag{58}$$

where $\boldsymbol{L}$ is a square diagonal matrix that accounts for the repetition of *distinct* elements in $[\boldsymbol{g}|\boldsymbol{E}]$. For example, if η_5 appears 6 times in $[\boldsymbol{g}|\boldsymbol{E}]$, then the fifth diagonal element of $\boldsymbol{L}$ will be $\sqrt{6}$. (The procedure for constructing matrix $\boldsymbol{L}$ is given in Appendix.)

Hence the objective function can be restated as:

to minimize $\|\boldsymbol{L}\eta\|_2$ subject to $\hat{\boldsymbol{r}} = \boldsymbol{0}$.

Since it may not be possible to ensure that $\hat{\boldsymbol{r}}$ is identically a zero vector due to the restriction imposed on the structure of $[\boldsymbol{g}|\boldsymbol{E}]$, one has to minimize $\hat{\boldsymbol{r}}$ as far as possible.

Therefore, the objective function is rephrased as:

$$\min_{\eta,\boldsymbol{a}} \left\| \begin{pmatrix} \hat{\boldsymbol{r}}(\eta,\boldsymbol{a}) \\ \boldsymbol{L}\eta \end{pmatrix} \right\|_2 \tag{59}$$

That is, to find vectors η and $\boldsymbol{a}$ which minimize l_2 norms of vector $\hat{\boldsymbol{r}}$ and the product vector $\boldsymbol{L}\eta$.
It is important to note that the greater emphasis is on the minimization of l_2 norm of $\hat{\boldsymbol{r}}$.

The solution of this problem has to be attempted iteratively starting with suitable initial values. The initial values adopted in this work are:

$$\boldsymbol{a_i} = \boldsymbol{a}_{\hat{TLS}} \tag{60}$$
$$\eta_i = \boldsymbol{0} \tag{61}$$

The initial solution $\boldsymbol{a_i}$ obtained from TLS approach is preferred to that obtained from least-squares as it is naturally more accurate [18]. Beginning with these values, the problem is approached as a linear approximation problem with Newton's method as follows:

$$\hat{\boldsymbol{r}}(\eta+\Delta\eta, \boldsymbol{a}+\Delta\boldsymbol{a}) = \hat{\boldsymbol{r}}(\eta, \boldsymbol{a}) + \frac{\partial \boldsymbol{r}}{\partial \eta}\Delta\eta + \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{a}}\Delta\boldsymbol{a} \tag{62}$$
$$\boldsymbol{L}(\eta+\Delta\eta) = \boldsymbol{L}\eta + \boldsymbol{L}\,\Delta\eta \tag{63}$$

In order to get the partial derivatives, the following alternative representations of $\hat{\boldsymbol{r}}$ are used:

$$\begin{aligned}\hat{\boldsymbol{r}} &= \boldsymbol{y} + \boldsymbol{g} - (\boldsymbol{Y}+\boldsymbol{E})\boldsymbol{a} \\ &= \boldsymbol{y} + \boldsymbol{P_1}\eta - (\boldsymbol{Y}+\boldsymbol{E})\boldsymbol{a} \qquad (64)\\ &= \boldsymbol{y} + \boldsymbol{P_1}\eta - \boldsymbol{Y}\boldsymbol{a} - \boldsymbol{A}\alpha \\ &= \boldsymbol{y} + \boldsymbol{P_1}\eta - \boldsymbol{Y}\boldsymbol{a} - \boldsymbol{A}\boldsymbol{P_0}\eta \qquad (65)\end{aligned}$$

Therefore

$$\frac{\partial \hat{\boldsymbol{r}}}{\partial \boldsymbol{a}} = -(\boldsymbol{Y}+\boldsymbol{E}) \qquad [\text{from}(64)] \tag{66}$$
$$\frac{\partial \hat{\boldsymbol{r}}}{\partial \eta} = \mathbf{P_1} - \mathbf{AP_0} \qquad [\text{from}(65)] \tag{67}$$

In obtaining the above expressions, the following equivalence is used:

$$\boldsymbol{E}\boldsymbol{a} = \boldsymbol{A}\alpha \tag{68}$$

Expanding this, one can get matrix $\boldsymbol{A}$ as follows:

$$\underbrace{\begin{bmatrix} \eta_p & \eta_{p-1} & \dots & \eta_2 & \eta_1 \\ \eta_{p+1} & \eta_p & \dots & \eta_3 & \eta_2 \\ \vdots & \vdots & \ddots & \vdots & \\ \eta_{p+m-1} & \eta_{p+m-2} & \dots & \eta_{m+1} & \eta_m \end{bmatrix}}_{\boldsymbol{E}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}}_{\boldsymbol{a}} = \boldsymbol{A} \underbrace{\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_p \\ \eta_{p+1} \\ \vdots \\ \eta_{p+m-1} \end{bmatrix}}_{\alpha} \tag{69}$$

From (69), it is evident that

$$\boldsymbol{A} = \begin{bmatrix} a_p & a_{p-1} & a_{p-2} & \dots & a_1 & 0 & 0 & \dots & 0 \\ 0 & a_p & a_{p-1} & a_{p-2} & \dots & a_1 & 0 & \dots & 0 \\ 0 & 0 & a_p & a_{p-1} & a_{p-2} & \dots & a_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_p & a_{p-1} & a_{p-2} & \dots & a_1 \end{bmatrix}_{m\times(p+m-1)} \tag{70}$$

NOTE: MATLAB^{TM} [35] has a built-in function 'toeplitz', which generates a Toeplitz matrix when the first column and the first row of the matrix are specified as the arguments. Following the syntax of this function, matrix $\boldsymbol{A}$ may be written as:

$$\boldsymbol{A} = \text{toeplitz}(\boldsymbol{d_1}, \boldsymbol{e}) \tag{71}$$

where

$$\boldsymbol{d_1} = \begin{bmatrix} a_p & \underbrace{0\,0\dots 0}_{(m-1)\text{ zeros}} \end{bmatrix}^T \tag{72}$$
$$\boldsymbol{e} = \begin{bmatrix} a_p\ a_{p-1}\ a_{p-2}\dots\ a_2\ a_1 & \underbrace{0\,0\dots 0}_{(m-1)\text{ zeros}} \end{bmatrix}^T \tag{73}$$

With the above, (62) and (63) can be rewritten as:

$$\hat{\boldsymbol{r}}(\eta+\Delta\eta, \boldsymbol{a}+\Delta\boldsymbol{a}) = \hat{\boldsymbol{r}}(\eta, \boldsymbol{a}) - (\boldsymbol{Y}+\boldsymbol{E})\Delta\eta + (\boldsymbol{P_1} - \boldsymbol{AP_0})\Delta\boldsymbol{a} \tag{74}$$
$$\boldsymbol{L}(\eta+\Delta\eta) = \boldsymbol{L}\eta + \boldsymbol{L}\,\Delta\eta \tag{75}$$

Considering RHS of (74) and (75), (59) can be reformulated as:

$$\min_{\Delta\eta, \Delta\boldsymbol{a}} \left\| \begin{bmatrix} (\boldsymbol{P_1} - \boldsymbol{AP_0}) & -(\boldsymbol{Y}+\boldsymbol{E}) \\ \boldsymbol{L} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \Delta\eta \\ \Delta\boldsymbol{a} \end{bmatrix} + \begin{bmatrix} \hat{\boldsymbol{r}} \\ \boldsymbol{L}\eta \end{bmatrix} \right\|_2 \tag{76}$$

Such a minimization problem is normally handled as a least squares problem as follows:

$$\begin{bmatrix} (\boldsymbol{AP_0} - \boldsymbol{P_1}) & (\boldsymbol{Y}+\boldsymbol{E}) \\ -\boldsymbol{L} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \Delta\eta \\ \Delta\boldsymbol{a} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{r}} \\ \boldsymbol{L}\eta \end{bmatrix} \tag{77}$$

NOTE: As already mentioned, the basic premise behind the least squares approach is that the equations cannot be satisfied exactly. The RHS is not exactly equal to LHS. Therefore, the effort is directed towards *minimizing* the sum of the squares of the errors, which is same as minimizing the l_2 norm of the error vector. For the least-squares formulation of (77), the error vector is the function in (76). Thus, (76) and (77) are equivalent.

However, in applying least-squares principle to an overdetermined set of linear equations, if a few equations have to be satisfied more accurately, such equations are weighted by a factor W, and this leads to weighted least squares [36]. As already indicated, in case of (76), the part corresponding to (74) has to be satisfied more accurately than the part corresponding to (75). Hence the weighted least squares formulation in this case is represented as:

$$\begin{bmatrix} W(\boldsymbol{AP_0} - \boldsymbol{P_1}) & W(\boldsymbol{Y}+\boldsymbol{E}) \\ -\boldsymbol{L} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \Delta\eta \\ \Delta\boldsymbol{a} \end{bmatrix} = \begin{bmatrix} W\hat{\boldsymbol{r}} \\ \boldsymbol{L}\eta \end{bmatrix} \tag{78}$$

Greater the value of the weighting factor, W is, the better [47]. In this work, W is taken as 10^8.

Using the values of $\Delta\eta$ and $\Delta\boldsymbol{a}$ obtained from the solution of (78), the vectors η and $\boldsymbol{a}$ are updated as:

$$\eta_{new} = \eta + \Delta\eta \tag{79}$$
$$\boldsymbol{a_{new}} = \boldsymbol{a} + \Delta\boldsymbol{a} \tag{80}$$

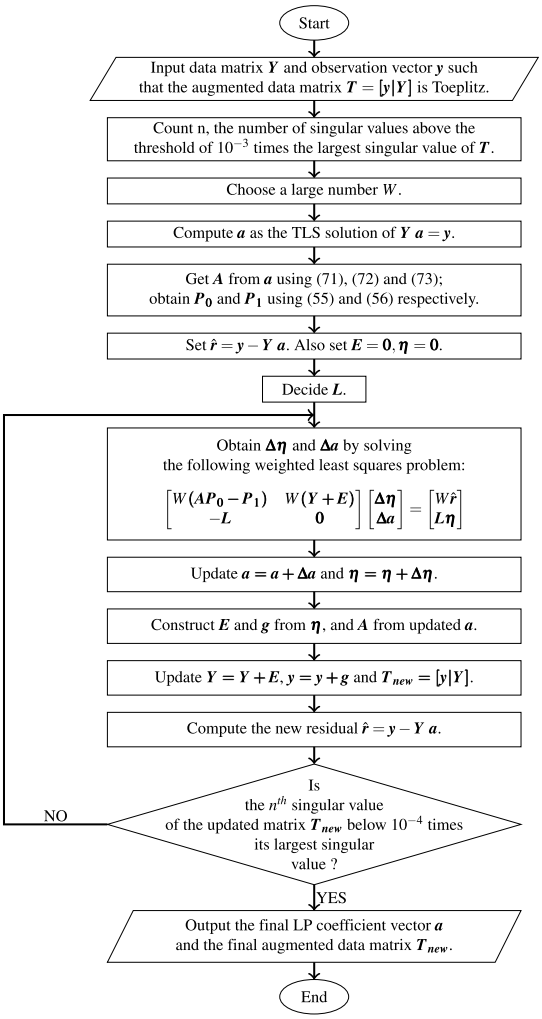

FIGURE 1. Flowchart for linear prediction solution of STLS-Prony algorithm.

Using the new values of η_i, matrix $\boldsymbol{E}$ and vector $\boldsymbol{g}$ are also updated appropriately as $\boldsymbol{E_{new}}$ and $\boldsymbol{g_{new}}$. With these, the augmented data matrix $\boldsymbol{T} = [\boldsymbol{y}|\boldsymbol{Y}]$ is updated as:

$$\boldsymbol{T_{new}} = \boldsymbol{T} + [\boldsymbol{g_{new}}|\boldsymbol{E_{new}}] \tag{81}$$

Conventionally, iteration is terminated when the Eucledian norms (or l_2 norms) of the correction vectors $\Delta\eta$ and $\Delta\boldsymbol{a}$ go below a pre-specified tolerance [17]. The termination criterion adopted in this work is slightly different. Here, the number of singular values of the initial augmented data matrix $\boldsymbol{T}$ whose value is above 10^{-3} times the largest singular value is noted as n. This is the *numerical* rank of the augmented data matrix. The leitmotif of the linear prediction solutions presented in this paper is rank reduction. Hence, in this work, iteration is terminated when the n^{th} singular value of the updated augmented data matrix $\boldsymbol{T_{new}}$ is reduced to 10^{-4} times its largest singular value. This is tantamount to reducing the *numerical* rank of $\boldsymbol{T}$ to at least $(n-1)$.

This completes the determination of LP coefficient vector $\boldsymbol{a}$, which is the first step in Prony analysis.

Hence the STLS-Prony algorithm can be summarized as given below. *It may be noted that among the steps 2 – 6, some deviate slightly from those given at the end of Section III.*

- **Step 1 – determination of LP coefficients:** The flowchart in Fig. 1 encapsulates the determination of LP coefficients in STLS-Prony algorithm.
- **Step 2:** Polynomial rooting for getting discrete modes z_i.
- **Step 3:** Determination of continuous-time eigenvalues λ_i from discrete modes z_i. Decrement factor σ_i and radian frequency ω_i are obtained from λ_i. The steps 2 and 3 are identical to those in PE-Prony algorithm discussed in Section III.
- **Step 4:** Determination of residues B_i by Vandermonde solution. Amplitude A_i and phase-angle ϕ_i are thence calculated.
  NOTE: The unique feature of STLS is that the noise-filtered data matrix $\boldsymbol{T_{new}}$ obtained at the end of step 1 retains the Toeplitz structure of the original augmented data matrix $\boldsymbol{T} = [\boldsymbol{y}|\boldsymbol{Y}]$. Hence it is possible to obtain a unique noise-filtered value for each sample from the data matrix $\boldsymbol{T_{new}}$. This noise-filtered version of the original signal is used in Vandermonde solution.
- **Step 5:** Thereafter, dominant modes are selected based on their energy content employing 10% of the energy of the most dominant mode as the threshold.
- **Step 6:** As an optional step, the signal can be reconstructed from the details of amplitude, phase-angle, decrement factor and radian frequency of different components, and the fitness metrics SER and MAPE can be computed. *However, in computing these, the reconstructed signal is compared with the noise-filtered version of the original signal got from the matrix* $\boldsymbol{T_{new}}$ *here, unlike in other versions of the Prony algorithm, where the reconstructed signal is compared with the original noisy signal itself.*

VI. CASE STUDIES USING SVD-AUGMENTED PRONY ALGORITHMS

The motivation for the work reported in this paper is the improvement of Prony algorithm performance, particularly in case of noisy signals. So long as the SNR level of the ring-down signal is 20 dB or above, the iterative Prony algorithm works effectively, and the need for a better algorithm does not arise. However, as SNR goes below 20 dB (or as the noise content in the ring-down signal increases), proliferation of the number of components n takes place, and then the iterative Prony method fails. This is due to the fact that the linear prediction solution requires that the number of components n be less than or equal to $(N/2)$ i.e., half the number of samples in the window, and this condition is not met when SNR of the test signal goes below 20 dB.

On the other hand, the mode identification results of the SVD-augmented Prony algorithms match quite closely with those of the iterative Prony method up to an SNR level of

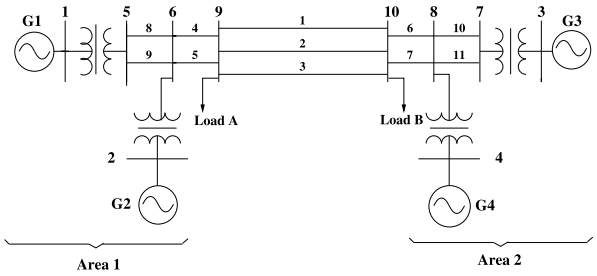

FIGURE 2. Two-area, four-machine power system.

20 dB, and hence there is no need to discuss these in detail. Therefore, the focus here is narrowed down to signals with lower SNR (higher noise content). Accordingly, case studies employing SVD-augmented Prony algorithms are presented for the following four SNR values: (a) 10 dB, (b) 5 dB, (c) 2 dB and (d) 1 dB.

The well-known four- , ten- and sixteen-machine power systems are selected for carrying out case studies as these are among the IEEE-recommended benchmark models [37]. Noisy signals are obtained by adding white Gaussian noise to noiseless signals obtained through time-domain simulation of power system models. Further, the test signal window length is standardized at 20 s in this work. This is based on the recommendation in [38], which is a comprehensive document on Prony analysis of power system signals. In this reference, it is suggested that the minimum window length be at least two cycles of the lowest frequency component expected. The lowest frequency of electromechanical oscillations is 0.1 Hz [39], and two cycles of this frequency correspond to 20 s. Further, for the sake of confirmation, a shorter window length of 10 s is tried, and then all the three SVD-augmented Prony algorithms are found to fail.

A. MODE IDENTIFICATION OF SIGNALS FROM FOUR-MACHINE SYSTEM

The one-line diagram of this system is given in Fig. 2.

Four-machine, two-area system is a small synthesized system that is ideal for studying electromechanical oscillations [39]. There is a clear demarcation between the two areas, with generators 1 and 2 belonging to area-1, and generators 3 and 4 belonging to area-2. The number of electromechanical modes is three (which is equal to number of generators minus one). Out of the three electromechanical modes, one is an inter-area mode while the remaining two are local modes – local mode of area-1 and local mode of area-2.

System data used are obtained from [40]. The model employed for generators is operational impedance 2.2 [41].

NOTE: 2.2 model implies that two coils – one field coil and one damper coil – are considered along the d-axis of the rotor and two damper coils are considered along the q-axis of the rotor. The electrical state variables are flux linkages of these four coils. Added with two mechanical state variables, namely, power angle δ and slip (or per unit speed deviation), this results in a sixth order model.

Static exciters are enabled on all the generators. Loads considered are of constant-impedance type. Time-domain simulation window is 20 s long, and Dormand-Prince (*ODE*5) numerical integration method is employed with a fixed step size of 0.025 s. Note that this step size corresponds to 40 samples/s. This matches well with the reporting rate of WAMS worldwide, which varies between 10 – 120 frames/s [4]. Further, as electromechanical oscillations have a maximum frequency of 4 Hz [1], this sampling rate is quite adequate for reconstructing these oscillations.

For obtaining oscillatory signals, reference voltage of generator 4 is perturbed. The slip (or per unit speed deviation) signal of generator 4 referenced to its center-of-inertia (COI) [42] and the detrended electrical torque signal of generator 1 are taken up for mode identification.

The normal practice for validating these results is to compare them with the benchmark eigenvalues obtained through detailed small-signal stability analysis [43]. The relevant swing-mode values got from small-signal stability analysis in this case are [44], [45]:

$-1.0488 \pm j6.7981$ (highly damped local mode of area-2) and

$-0.0372 \pm j4.4583$ (inter-area mode).

1) MODE IDENTIFICATION OF SLIP SIGNAL OF GENERATOR 4:

The results obtained in this sub-case are presented in Table 1.

From the results in Table 1, one can draw the following inferences:

- As expected, both the highly-damped local mode of area-2 and the inter-area mode can be observed in this signal as it is drawn from the area-2 where the perturbation has originated.
- When SNR is 10 dB, both TLS-Prony and STLS-Prony algorithms successfully identify the highly damped local mode whereas PE-Prony fails to do so. Further, local mode value of $-1.0702 \pm j6.5863$ estimated with STLS-Prony algorithm is closer to the benchmark eigenvalue of $-1.0488 \pm j6.7981$ than that estimated with TLS-Prony algorithm. Thus STLS-Prony algorithm performs better than the other two algorithms. *Hence STLS-Prony algorithm has practical relevance in system identification from noisy signals containing a well-damped mode.*

  NOTE: Mode identification is generally considered successful when the decrement factor matches up to first decimal place and radian frequency deviation is within 4% in comparison to those of the benchmark eigenvalues obtained through detailed small-signal stability analysis [10].
- Further, when SNR is 10 dB, both TLS-Prony and STLS-Prony do show a spurious mode with purely real value. This represents an exponentially varying DC. Throwing up spurious DC modes is a known issue with Prony algorithms [8]. Since the value of this purely real mode

TABLE 1. Mode identification results of highly noisy slip-COI signal of generator 4 from stable 4-machine system: V_{Ref} of G4 is perturbed.

Algorithm	Dominant modes λ_i	Energy $\times 10^{-6}$	MAPE	SER (dB)	Run-time (s)
slip_COI signal of Generator 4: SNR = 10 dB					
PE-Prony	-0.0337 ± j4.4693 -0.9789 ± j6.5402 -0.8996 ± j52.4477 -0.6187 ± j52.2237	0.7878 0.5627 0.2427 0.1026	3.5506	14.6520	0.4926
TLS-Prony	-4.4310 -0.0324 ± j4.4553 -1.0113 ± j7.0377	0.7879 0.7670 0.5612	4.6190	14.6743	0.5027
STLS-Prony	-1.0702 ± j6.5863 -0.0239 ± j4.4639 -12.2056	0.7766 0.7601 0.2239	0.0998	37.5032	4.8413
slip_COI signal of Generator 4: SNR = 5 dB					
PE-Prony	-0.0327 ± j4.4670 -0.9063 ± j6.4369 -0.9712 ± j52.5821 -9.3862 ± j108.8543 -0.9752 ± j50.8669 -0.5481 ± j52.2091	0.8171 0.5085 0.4017 0.2196 0.1322 0.1159	1.1100	10.3165	0.4950
TLS-Prony	-0.0380 ± j4.4629 -0.7419 ± j6.1847	0.8478 0.3814	1.2490	10.3399	0.5207
STLS-Prony	-0.0181 ± j4.4848 -0.8828 ± j6.4984	0.9092 0.5379	0.0637	39.7309	5.3859
slip_COI signal of Generator 4: SNR = 2 dB					
PE-Prony	-0.0307 ± j4.4639 -0.9958 ± j52.7379 -0.8174 ± j6.3760 -8.5614 ± j108.5474 -0.8479 ± j50.8841 -1.5664 ± j114.3195 -1.7655 ± j32.6098 -0.4950 ± j52.2101 -1.2508 ± j56.9680	0.8378 0.4763 0.4483 0.3986 0.1831 0.1461 0.1396 0.1189 0.0942	7.7428	8.2124	0.4893
TLS-Prony	-0.0327 ± j4.4638 -0.7902 ± j6.2229	0.8731 0.4693	17.0778	8.2649	0.5214
STLS-Prony	-0.0177 ± j4.4669 -0.7912 ± j6.5062	0.8343 0.4971	0.1752	31.8606	5.9587
slip_COI signal of Generator 4: SNR = 1 dB					
PE-Prony	-0.0297 ± j4.4624 -0.9908 ± j52.7857 -8.3186 ± j108.4561 -0.7800 ± j6.3568 -0.8090 ± j50.8843 -1.5601 ± j114.3494 -1.7340 ± j32.7057 -0.4809 ± j52.2109 -1.2837 ± j56.9993	0.8451 0.5118 0.4845 0.4208 0.2061 0.1819 0.1787 0.1267 0.1223	1.9847	7.6244	0.4890
TLS-Prony	-0.0305 ± j4.4630 -0.7723 ± j6.2215	0.8800 0.4690	2.1256	7.7642	0.5160
STLS-Prony	-0.0154 ± j4.4661 -0.7386 ± j6.4966	0.8444 0.4627	0.0914	35.6405	5.9459

is negative, it represents an exponentially *decaying* DC component, and hence is not worrisome.

- When SNR goes below 10 dB, all the three algorithms fail to identify the highly damped local mode.
- However, all the three algorithms succeed in identifying the inter-area mode right down to an SNR of 1 dB.
- As already mentioned, the rank-reduced augmented data matrix obtained in STLS retains the Toeplitz structure. Hence the diagonals of this Toeplitz matrix offer unique values for the individual samples. These values can be regarded as the filtered sample values. It is this *filtered input signal* that is compared against the reconstructed signal for the computation of the fitness metrics in the STLS-Prony algorithm. Hence, it can be seen that even when the SNR level of the input signal is 1 dB, SER between the filtered input signal and the reconstructed signal is around 30 dB, which indicates a very good matching. This validates the efficacy of the STLS approach as a filtering technique.

  On the other hand, in case of PE-Prony and TLS-Prony algorithms, the rank-reduced data matrix does not retain the Toeplitz structure. As such, it is not possible to assign unique values to individual samples from the rank-reduced matrix. Hence, a filtered version of the input signal cannot be obtained. Therefore, in case of these algorithms, the reconstructed signal is compared against the original unfiltered input signal. Naturally, as the SNR level of the input signal decreases, the SER between the reconstructed signal and the original unfiltered input signal turns out to be lower and lower, indicating a poor matching.
- However, the run-time of STLS-Prony method is much longer than those of the other two algorithms due to the iterative procedure involved in obtaining an optimally filtered data matrix.

2) MODE IDENTIFICATION OF DETRENDED ELECTRICAL TORQUE SIGNAL OF GENERATOR 1:

The results of this sub-case are listed in Table 2. As the perturbation has taken place in area-2, the local swing mode of area-1 is not excited. Since the test signal considered here is drawn from area-1, one gets to observe only the inter-area mode in this signal.

Given below are the inferences drawn from the results in Table 2.

- The poorly-damped inter-area mode is the only electromechanical mode present in the signal. Hence all the three algorithms succeed in identifying it even when the SNR is reduced to 1 dB.
- However, when the SNR is 2 dB or lower, PE-Prony throws up some spurious high frequency modes as dominant in addition to genuine electromechanical modes. Among these, there are two modes, whose frequency is around 8 rad/s. This is clearly within the frequency range of electromechanical modes, and hence causes confusion. (The swing mode frequency range goes normally up to 2 Hz or 12.57 rad/s and extends occasionally to 4 Hz or 25.13 rad/s [40].)
- STLS-Prony method also shows a spurious high frequency mode of 125.66 rad/s as dominant when the SNR level of the signal is 1 dB. However, the frequency range of this spurious mode is outside that of electromechanical modes, and hence this does not cause confusion. In fact, 125.66 rad/s equates to 20 Hz, which is half the sampling frequency. Throwing up a spurious mode with frequency equal to half the

TABLE 2. Mode identification results of highly noisy detrended electrical torque of generator 1 of stable 4-machine system: V_{Ref} of G4 is perturbed.

Algorithm	Dominant modes λ_i	Energy	MAPE	SER (dB)	Run-time (s)
Detrended electrical torque *Te* signal of Generator 1: SNR = 10 dB					
PE-Prony	-0.0349 ± j4.4684	0.4020	0.6349	14.7472	0.4932
TLS-Prony	-0.0360 ± j4.4654	0.3981	0.6558	14.8360	0.5203
STLS-Prony	-0.0359 ± j4.4641	0.4051	0.0222	41.6939	3.8341
Detrended electrical torque *Te* signal of Generator 1: SNR = 5 dB					
PE-Prony	-0.0358 ± j4.4644	0.4028	1.7416	10.5310	0.4912
TLS-Prony	-0.0366 ± j4.4598	0.3936	1.3605	10.7547	0.5195
STLS-Prony	-0.0354 ± j4.4576	0.4038	0.0688	38.4500	5.9275
Detrended electrical torque *Te* signal of Generator 1: SNR = 2 dB					
PE-Prony	-0.0367 ± j4.4610 -3.7621 ± j8.2093 -1.6334 ± j8.1154 -7.9218 ± j108.2464 -0.9127 ± j52.9328	0.4033 0.1044 0.0900 0.0700 0.0471	1.5313	8.4546	0.4928
TLS-Prony	-0.0372 ± j4.4548	0.3885	1.5043	8.7331	0.5227
STLS-Prony	-0.0347 ± j4.4519	0.4042	0.0555	36.8363	4.8834
Detrended electrical torque *Te* signal of Generator 1 : SNR = 1 dB					
PE-Prony	-0.0370 ± j4.4596 -3.6813 ± j8.2038 -1.6723 ± j8.1571 -7.9109 ± j108.2434 -0.9128 ± j52.9328	0.4035 0.1472 0.1287 0.0884 0.0594	2.6645	7.8671	0.4965
TLS-Prony	-0.0375 ± j4.4526	0.3866	3.0620	8.1580	0.5183
STLS-Prony	-0.0344 ± j4.4496 -4.0379 ± j125.6637	0.4043 0.0485	0.0754	34.4258	4.0435

sampling frequency is a well-known drawback of Prony methods [8].

B. MODE IDENTIFICATION OF A SIGNAL FROM TEN-MACHINE SYSTEM

The one-line diagram of the ten-machine, four-area power system is depicted in Fig. 3. This system is taken up for a case study as it represents a practical system, namely, New England power system.

In this case also, power system data are borrowed from [40]. The model employed for generators is 1.1. This implies that only one coil, namely, field coil is considered along rotor d-axis and one damper coil is considered along rotor q-axis. The two electrical state variables selected are the flux linkages of these two coils. Along with the two mechanical state variables – power angle δ and slip, this results in a fourth order model.

IEEE $AC4A$ exciters [42] are enabled on all the ten generators. As in the previous case, time-domain simulation window is 20 s long, and Dormand-Prince ($ODE5$) numerical integration method is employed with a fixed step size of 0.025 s.

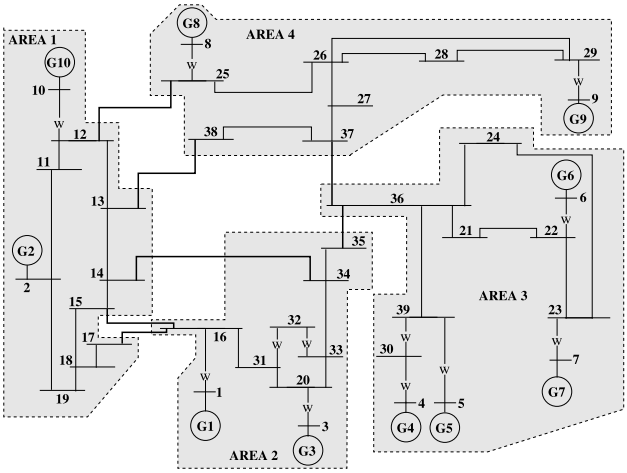

FIGURE 3. Four-area, ten-machine New England power system.

For obtaining oscillatory signals, reference voltage of generator 7 is perturbed. The active power signal in the tie-line between buses 12 and 25 is then taken up for mode identification. This case is interesting because the system goes unstable. In fact, small-signal stability analysis shows that there are two negatively-damped inter-area swing modes, whose values are [44]:

$0.2111 \pm j6.0077$ and $0.0624 \pm j3.9632$.

The corresponding mode identification results are contained in Table 3.

The inferences drawn from the results in Table 3 are as follows:

- All the three algorithms succeed in identifying both the negatively-damped inter-area modes right down to an SNR of 1 dB. This is very significant from the viewpoint of online control.
- Below an SNR of 5 dB, PE-Prony and STLS-Prony algorithms throw up some spurious high frequency modes. However, the frequency of such modes is outside the frequency range of swing modes.

C. MODE IDENTIFICATION OF A SIGNAL FROM SIXTEEN-MACHINE SYSTEM

This system is selected as a representative of large power systems. The one-line diagram of the sixteen-machine power system is depicted in Fig. 4.

The data of this power system are adopted from [46]. The model employed for generators is operational impedance 2.2 (sixth order model). $DC1A$ exciters [42] are enabled on generators $1-8$ and static exciter with slip PSS is enabled on generator 9. However, no exciter is enabled on generators $10-16$. To initiate oscillations, a symmetrical fault lasting 0.05 s is simulated at bus 66. The signal considered for mode identification is the detrended active power in the line between buses 9 and 30, and the analysis window lasts for 20 s

TABLE 3. Mode identification results of highly noisy detrended power in line 12 – 25 of 10-machine system: V_{Ref} of G7 is perturbed.

Algorithm	Dominant modes λ_i	Energy	MAPE	SER (dB)	Run-time (s)
*P*12_25*d* signal: SNR = 10 dB.					
PE-Prony	0.2176 ± j5.9664 0.0746 ± j3.9556	2.1143 0.4646	3.2996	14.7385	0.4964
TLS-Prony	0.2198 ± j5.9615 0.0839 ± j3.9666	2.1632 0.4835	3.5786	15.0338	0.5153
STLS-Prony	0.2149 ± j5.9580 0.0930 ± j3.9550	2.1002 0.5310	0.2412	41.3024	5.8575
*P*12_25*d* signal: SNR = 5 dB.					
PE-Prony	0.2135 ± j5.9733 0.0704 ± j3.9571	2.0924 0.4910	1.3559	10.5343	0.5005
TLS-Prony	0.2184 ± j5.9658 0.0886 ± j3.9826	2.1633 0.5123	1.5563	10.9483	0.5138
STLS-Prony	0.2070 ± j5.9602 0.1033 ± j3.9541	2.0171 0.6037	0.0849	36.9103	5.9182
*P*12_25*d* signal: SNR = 2 dB.					
PE-Prony	0.2091 ± j5.9799 0.0641 ± j3.9575 -7.2275 ± j108.0019 -0.8627 ± j52.9712	2.0677 0.5167 0.3973 0.2120	1.9069	8.4560	0.4983
TLS-Prony	0.2170 ± j5.9709 0.0914 ± j4.0014	2.1461 0.5315	1.6926	8.8679	0.5193
STLS-Prony	0.2020 ± j5.9600 0.1127 ± j3.9556 -1.0738 ± j110.4934 -0.5194 ± j110.8542	1.9887 0.6870 0.5079 0.4275	0.1037	35.5796	5.9534
*P*12_25*d* signal: SNR = 1 dB.					
PE-Prony	0.2072 ± j5.9825 0.0613 ± j3.9575 -7.2104 ± j107.9962 -0.8618 ± j52.9729 -1.5084 ± j114.4895	2.0568 0.5273 0.4983 0.2645 0.2168	1.2965	7.8636	0.4993
TLS-Prony	0.2165 ± j5.9729 0.0923 ± j4.0088	2.1400 0.5411	2.1371	8.2769	0.5194
STLS-Prony	0.2003 ± j5.9599 0.1162 ± j3.9566 -1.3858 ± j110.4332 -0.4251 ± j110.8654	1.9842 0.7231 0.2846 0.2052	0.3235	29.0866	6.0300

starting 1s after the initiation of the fault. (This 1 s delay in the starting point of the analysis window is meant to exclude the nonlinear characteristics triggered due to fault occurrence.) The numerical integration method used is Dormand-Prince (*ODE*5) with a fixed step size of 0.02 s. Note that the step size is selected here deliberately to be slightly different from that in the previous case studies. This is meant to confirm that there is no great variation in the performance of the algorithms when the reporting rate is varied moderately. However, even in this case, the sampling rate of 50 samples/s is within the band of $10 - 120$ frames/s employed in the practical WAMS [4].

The results of mode identification are listed in Table 4. The relevant eigenvalues indicated by detailed small-signal stability analysis are:

$-0.0888 \pm j2.4678$ and $-0.0762 \pm j3.1489$.

Following are the observations made from the results in Table 4:

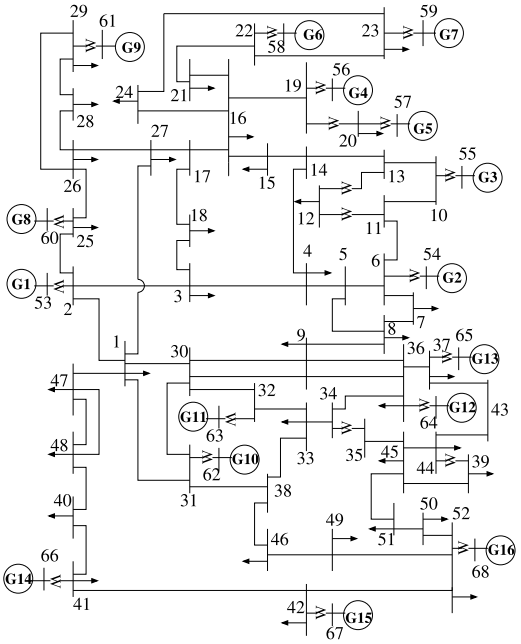

FIGURE 4. Five-area, sixteen-machine power system.

- Although the signal has two electromechanical modes, these are lightly damped. Hence all the three algorithms are effective in identifying these right down to an SNR of 1 dB.
- When SNR is 10 dB, PE-Prony algorithm throws up a spurious mode with a purely real positive value of 0.0211. Again, when SNR is 5 dB, PE-Prony throws up a similar mode of value 0.0091. This can result in a false alarm as a purely real positive mode represents an exponentially *growing* DC which implies instability.
- The run-time of TLS-Prony method is more than 9 s in this case whereas it is around 0.5 s in all the previous cases. (A system with i7 processor, 3.4 GHz frequency and 4 GB system memory is used. Operating system employed is Ubuntu 14.04, and MATLAB^{TM} version deployed is 8.3 or *R*2014*A*.) This is because the Vandermonde solution is obtained with QR decomposition *via Givens rotation* [47] here due to rank-defeciency issue encountered when the solution is attempted with QR decomposition *via Householder transformation*. This problem of rank defeciency is encountered, and QR decomposition has to be approached via Givens rotation for Vandermonde solution in case of STLS-Prony algorithm as well. The 'mldivide' operator in MATLAB^{TM} resorts to 'QR decomposition via Householder transformation' when it is used for the least-squares solution of an overdetermined system of linear equations. Thus, MATLAB^{TM} can be said to have a built-in function for 'QR decomposition via Householder transformation' whereas 'QR decomposition via Givens rotation' has to be coded by the user. A user-defined function takes longer to execute as compared to a built-in function.

TABLE 4. Mode identification results of highly noisy detrended power in line 9-30 of 16-machine system for a fault at bus 66: 1 – 21 s after fault.

Algorithm	Dominant modes λ_i	Energy	MAPE	SER (dB)	Run-time (s)
*P*9_30*d* signal: SNR = 10 dB.					
PE-Prony	-0.0904 ± j2.4513 -0.0660 ± j3.1270 -0.2077 0.0211	3.7159 2.3426 1.9366 0.7854	0.7232	16.5639	0.6115
TLS-Prony	-0.0789 ± j2.4642 -0.0925 ± j3.1374 -0.3127 -0.0287	3.8215 2.6829 2.5381 0.5541	0.8785	15.6219	12.0218
STLS-Prony	-0.0925 ± j2.4574 -0.0692 ± j3.1329 -1.2498 -3.1948	3.9062 2.4644 0.9533 0.8690	0.0896	32.8349	18.4106
*P*9_30*d* signal: SNR = 5 dB.					
PE-Prony	-0.0927 ± j2.4477 -0.1933 -0.0638 ± j3.1277 0.0091 -9.4651 ± j98.0394 -4.1233 ± j157.0796	3.7083 2.3881 2.3842 1.0318 0.5743 0.4079	0.9877	12.1549	0.6085
TLS-Prony	-0.0734 ± j2.4826 -0.1147 ± j3.1220 -0.3429 -0.0404	4.0260 3.0510 2.6128 0.4786	1.1817	10.9508	9.3753
STLS-Prony	-0.0929 ± j2.4565 -0.0663 ± j3.1341 -0.4601 -9.5767	3.8894 2.5579 0.5858 0.4032	0.0466	41.3665	18.1014
*P*9_30*d* signal: SNR = 2 dB.					
PE-Prony	-0.0954 ± j2.4444 -0.0633 ± j3.1283 -9.7539 ± j98.1154 -4.2299 ± j157.0796 -2.1159 ± j85.3205 -6.5429 ± j27.6858	3.7243 2.4514 1.1772 0.8290 0.6895 0.4824	1.6453	9.7912	0.6050
TLS-Prony	-0.0721 ± j2.5073 -0.1243 ± j3.0979 -0.3689 -12.7864 ± j28.9815 -2.6130 ± j137.0950	4.2034 3.2998 2.5946 0.4573 0.4474	2.0357	8.4250	9.3712
STLS-Prony	-0.1022 ± j2.4546 -0.0674 ± j3.1420	4.2360 2.7359	0.0253	42.8247	16.2829
*P*9_30*d* signal: SNR = 1 dB.					
PE-Prony	-0.0966 ± j2.4432 -0.0633 ± j3.1286 -9.8773 ± j98.1333 -4.2708 ± j157.0796 -2.1207 ± j85.3280 -6.5876 ± j27.6652 -0.9623 ± j41.5397	3.7378 2.4868 1.5010 1.0543 0.8736 0.6062 0.4530	1.7473	9.0746	0.6005
TLS-Prony	-0.0729 ± j2.5198 -0.1259 ± j3.0877 -0.3796 -2.6148 ± j137.0977 -12.7390 ± j29.1182	4.2461 3.3843 2.5391 0.5558 0.5400	2.6562	7.6327	9.3632
STLS-Prony	-0.0958 ± j2.4549 -0.0642 ± j3.1367 -0.6704 -7.5575	3.9339 2.6997 0.6480 0.5981	0.0345	41.1524	15.4492

VII. CONCLUSION

The contribution of this paper can be summarized as follows:

- PE-Prony analysis and TLS-Prony analysis of power system signals are enunciated very lucidly and their nuances are brought out.
- A formulation of STLS-Prony algorithm for power system signals is developed successfully.
- It is demonstrated that STLS-Prony algorithm performs better than PE-Prony and TLS-Prony algorithms when the test signal has a highly damped local mode.
- On the other hand, when the test signal has only lightly damped inter-area modes, both TLS-Prony and STLS-Prony algorithms are effective right down to an SNR of 1 dB whereas PE-Prony algorithm throws up some spurious electromechanical modes.
- Filtering a noisy oscillatory signal is a nontrivial problem. In this context also, STLS is shown to be relevant by virtue of its characteristic ability to offer unique filtered values to individual samples of the test signal.
- The solution of linear equations in Prony analysis of highly noisy signals often gets bogged down by rank deficiency issues. It is demonstrated that this problem can be circumvented by taking recourse to QR decomposition via Givens rotation. As compared to the commonplace technique of Householder transformation, Givens rotation is no doubt long-winded. However, this can be expedited by adopting Fast Givens rotation, which will be taken up as future research.
- One basic limitation of STLS-Prony algorithm is its protracted execution. Adoption of Fast Givens rotation is hoped to improve this, and hence enhance the appeal of STLS-Prony algorithm remarkably.

TLS-Prony and STLS-Prony algorithms, which work right down to an SNR of 1 dB when the signal has only inter-area modes, have a scope for application to mode identification from ambient signals as the latter have an SNR level above this [48], [49]. The primary advantage of these algorithms is that these require a signal window length of only 20 s whereas the mode-meter algorithms normally used for mode identification from ambient signals require a signal window length of 10 minutes [50]. This drastic reduction in the required signal window length would improve power system reliability significantly by virtue of a pronounced curtailment in latency, which enables rapid initiation of mitigating control action in the event of onset of system instability.

APPENDIX
DETERMINATION OF MATRIX L

The basic advantage of the Toeplitz structure of the perturbation matrix $[\boldsymbol{g}|\boldsymbol{E}]$ in STLS-Prony algorithm presented in Section V is that its Frobenius norm can be obtained as the scaled Eucledian norm of a small vector made up of distinct elements of $[\boldsymbol{g}|\boldsymbol{E}]$. That is,

$$\|\boldsymbol{g}|\boldsymbol{E}\|_F = \|\boldsymbol{L}\boldsymbol{\eta}\|_2$$

TABLE 5. Repetitions of different η_i elements of $[g|E]$ in the example.

$\eta_1 = 1$ time	$\eta_2 = 2$ times	$\eta_3 = 3$ times	$\eta_4 = 4$ times
$\eta_5 = 5$ times	$\eta_6 = 5$ times	$\eta_7 = 5$ times	$\eta_8 = 5$ times
$\eta_9 = 4$ times	$\eta_{10} = 3$ times	$\eta_{11} = 2$ times	$\eta_{12} = 1$ time

Expanding this, one obtains the matrix $\boldsymbol{L}$ as given below.

$$\left\| \left[\begin{array}{c|cccc} \eta_{p+1} & \eta_p & \eta_{p-1} & \cdots & \eta_2 & \eta_1 \\ \eta_{p+2} & \eta_{p+1} & \eta_p & \cdots & \eta_3 & \eta_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ \underbrace{\eta_{p+m}}_{\boldsymbol{g}} & \underbrace{\eta_{p+m-1} & \eta_{p+m-2} & \cdots & \eta_{m+1} & \eta_m}_{\boldsymbol{E}} \end{array} \right] \right\|_F = \left\| [\boldsymbol{L}] \underbrace{\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_l \\ \eta_{p+1} \\ \vdots \\ \eta_{p+m} \end{bmatrix}}_{\eta} \right\|_2 \tag{82}$$

This is better illustrated through an example.

Consider a case with the number of rows in $[\boldsymbol{E}]$ = the number of rows in $[\boldsymbol{g}] = m = 8$.

Let the number of columns in $[\boldsymbol{E}] = p = 4$.

With these, (82) becomes:

$$\left\| \left[\begin{array}{c|cccc} \eta_5 & \eta_4 & \eta_3 & \eta_2 & \eta_1 \\ \eta_6 & \eta_5 & \eta_4 & \eta_3 & \eta_2 \\ \eta_7 & \eta_6 & \eta_5 & \eta_4 & \eta_3 \\ \eta_8 & \eta_7 & \eta_6 & \eta_5 & \eta_4 \\ \eta_9 & \eta_8 & \eta_7 & \eta_6 & \eta_5 \\ \eta_{10} & \eta_9 & \eta_8 & \eta_7 & \eta_6 \\ \eta_{11} & \eta_{10} & \eta_9 & \eta_8 & \eta_7 \\ \underbrace{\eta_{12}}_{\boldsymbol{g}} & \underbrace{\eta_{11} & \eta_{10} & \eta_9 & \eta_8}_{\boldsymbol{E}} \end{array} \right] \right\|_F = \left\| [\boldsymbol{L}] \underbrace{\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \\ \eta_9 \\ \eta_{10} \\ \eta_{11} \\ \eta_{12} \end{bmatrix}}_{\eta} \right\|_2 \tag{83}$$

From the Table 5, one can infer that if the Frobenius norm of $[\boldsymbol{g}|\boldsymbol{E}]$ is to become equal to Eucledian norm of $[\boldsymbol{L}\eta]$, then $\boldsymbol{L}$ in this example must be formed as a square diagonal matrix as follows:

$$\boldsymbol{L} = diag(\sqrt{1}, \sqrt{2}, \sqrt{3}, \sqrt{4}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{5}, \sqrt{4}, \sqrt{3}, \sqrt{2}, \sqrt{1}). \tag{84}$$

ACKNOWLEDGMENT

The authors wish to thank the reviewers wholeheartedly for their comprehensive review and valuable comments.

REFERENCES

[1] K. R. Padiyar and A. M. Kulkarni, *Dynamics and Control of Electric Transmission and Microgrids*. Hoboken, NJ, USA: Wiley, 2019.

[2] *Analysis and Control of Power System Oscillations*, Final Report of the Task force 07 of Advisory Group of Study Committee 38 of CIGRE, CIGRE Task force 07, Paris, France, Dec. 1996.

[3] D. N. Kosterev, C. W. Taylor, and W. A. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 967–979, Aug. 1999.

[4] *IEC/IEEE Standard for synchrophasor for Power Systems—Measurements*, Standard IEC/IEEE Std 60255-118-1 TM 2018, Dec. 2018.

[5] G. R. B. Prony, "Essai experimental at analytique, etc," *Paris, J. de L'Ecole Polytechnique*, vol. 1, no. 2, pp. 24–76, 1795.

[6] F. B. Hildebrand, *Introduction to Numerical Analysis*. New York, NY, USA: McGraw-Hill, 1956.

[7] M. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. Antennas Propag.*, vol. AP-23, no. 6, pp. 777–781, Nov. 1975.

[8] J. F. Hauer, C. J. Demeure, and L. L. Scharf, "Initial results in Prony analysis of power system response signals," *IEEE Trans. Power Syst.*, vol. 5, no. 1, pp. 80–89, Feb. 1990.

[9] K. Rao and K. N. Shubhanga, "MAPE—An alternative fitness metric for Prony analysis of power system signals," *Int. J. Emerg. Electr. Power Syst.*, vol. 19, no. 6, Dec. 2018, Art. no. 20180091, doi: 10.1515/ijeeps-2018-0091.

[10] K. Rao, "Performance analysis and improvement of power systems ring-down electromechanical mode identification algorithms," Ph.D. dissertation, Dept. Elect. Electron. Eng., Nat. Inst. Technol. Karnataka (NITK), Surathkal, India, 2021.

[11] *Identification of Electromechanical Modes in Power Systems*, Standard PES-TR15, IEEE Electromechanical Mode Identification Task Force, IEEE Power & Energy Society, Jun. 2012.

[12] R. Kumaresan and D. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, no. 6, pp. 833–840, Dec. 1982.

[13] M. D. Rahman and K.-B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 10, pp. 1440–1454, Oct. 1987.

[14] N. Zhou, J. Pierre, and D. Trudnowski, "Some considerations in using Prony analysis to estimate electromechanical modes," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Vancouver, BC, Canada, Jul. 2013, pp. 1–5, doi: 10.1109/PESMG.2013.6672888.

[15] K.-H. Lee, G.-H. Choi, and J.-H. Lee, "Perturbation theory-based performance analysis of TLS-prony method for natural frequency extraction," *Digit. Signal Process.*, vol. 85, pp. 17–28, Feb. 2019.

[16] Y. Li, K. J. R. Liu, and J. Razavilar, "A parameter estimation scheme for damped sinusoidal signals based on low-rank Hankel approximation," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 481–486, Feb. 1997.

[17] H. Park, L. Zhang, and J. B. Rosen, "Low rank approximation of a Hankel matrix by structured total least norm," *BIT Numer. Math.*, vol. 39, no. 4, pp. 757–779, 1999.

[18] P. Lemmerling, "Structured total least squares: Analysis, algorithms and applications," Ph.D. dissertation, Dept. Elect. Eng., KU Leuven, Leuven, Belgium, 1999.

[19] W. Zheng, C. Xu, J. Yang, J. Gao, and F. Zhu, "Low-rank structure preserving for unsupervised feature selection," *Neurocomputing*, vol. 314, pp. 360–370, Nov. 2018.

[20] H. Jin, Z.-H. Xiao, Q.-Y. Song, and Z.-Z. Qi, "Structure-preserving model order reduction for K-power bilinear systems via Laguerre functions," *Int. J. Syst. Sci.*, vol. 54, no. 8, pp. 1648–1660, Jun. 2023.

[21] J. F. Hauer, "Application of Prony analysis to the determination of modal content and equivalent models for measured power system response," *IEEE Trans. Power Syst.*, vol. 6, no. 3, pp. 1062–1068, Aug. 1991.

[22] C. E. Grund, J. J. Paserba, J. F. Hauer, and S. L. Nilsson, "Comparison of Prony and eigenanalysis for power system control design," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 964–971, Aug. 1993.

[23] K. Rao and K. N. Shubhanga, "A comparison of power system signal detrending algorithms," in *Proc. 7th Int. Conf. Power Syst. (ICPS)*, Dec. 2017, pp. 404–409.
[24] R. J. Hyndman, "Another look at forecast accuracy metrics for intermittent demand," *Foresight*, vol. 4, pp. 43–46, Jun. 2006.
[25] N. Zhou, J. W. Pierre, and D. Trudnowski, "A stepwise regression method for estimating dominant electromechanical modes," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 1051–1059, May 2012.
[26] G. Strang, *Linear Algebra and Its Applications*. Cengage India Pvt. Ltd., Noida, India, 2005.
[27] J. A. Cadzow, "Signal enhancement—A composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 1, pp. 49–62, Jan. 1988.
[28] D. W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood," *Proc. IEEE*, vol. 70, no. 9, pp. 975–989, Sep. 1982.
[29] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, USA, 1991.
[30] I. Markovsky and S. Van Huffel, "Overview of total least-squares methods," *Signal Process.*, vol. 87, no. 10, pp. 2283–2302, Oct. 2007.
[31] G. Holub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM Numer. Anal.*, vol. 17, no. 6, pp. 883–893, Dec. 1980.
[32] S. J. Leon, *Linear AlgebraWith Applications*. Boston, MA, USA: Pearson, 2014.
[33] B. De Moor, "Total least squares for affinely structured matrices and the noisy realization problem," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3104–3113, Nov. 1994.
[34] I. Dologlou, S. Van Huffel, and D. Van Ormondt, "Improved signal enhancement procedures applied to exponential data modeling," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 799–803, Mar. 1997.
[35] *MATLAB User's Manual Version R2014a*, MathWorks, Inc., Natick, MA, USA, 2014. [Online]. Available: https://www.mathworks.com
[36] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Philadelphia, PA, USA: SIAM, 1995.
[37] IEEE Task Force on Benchmark Systems for Stability Controls, "Benchmark models for the analysis and control of small-signal dynamics on power systems," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 715–722, Jan. 2017.
[38] D. J. Trudnowski, "Characteristics of identifying linear dynamic models from impulse response data using Prony analysis," Pacific Northwest Laboratory, Richland, WA, USA, 1992, doi: 10.2172/6843209.
[39] M. Klein, G. J. Rogers, and P. Kundur, "A fundamental study of inter-area oscillations in power systems," *IEEE Trans. Power Syst.*, vol. 6, no. 3, pp. 914–921, Aug. 1991.
[40] K. R. Padiyar, *Power System Dynamics - Stability and Control*. Hyderabad, India: BS Publications, 2008.
[41] K. R. Padiyar, *Analysis of Subsynchronous Resonance in Power Systems*. Norwell, MA, USA: Kluwer, 1999.
[42] P. Kundur, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill, 1994.
[43] H. Huang, Z. Xu, and W. Hua, "Estimation of interarea modes in large power systems," *Int. J. Electr. Power Energy Syst.*, vol. 53, pp. 196–208, Dec. 2013.
[44] K. N. Shubhanga and Y. Anantholla, "Manual for a multi-machine small signal stability programme," NITK, Surathkal, Mangalore, India, Tech. Rep., 2009. [Online]. Available: https://www.ee.iitb.ac.in/~peps/download/Small-signal_Stability_Programs/
[45] K. N. Shubhanga, *Power System Analysis: A Dynamic Perspective*, 1st ed. Noida, India, U.K.: Pearson India Education Services, 2018.
[46] G. Rogers, *Power Systems Oscillations*. Norwell, MA, USA: Kluwer, 2000.
[47] A. A. Anda and H. Park, "Self-scaling fast rotations for stiff and equality-constrained linear least squares problems," *Linear Algebra Appl.*, vol. 234, pp. 137–161, Feb. 1996.
[48] M. Larsson and D. S. Laila, "Monitoring of inter-area oscillations under ambient conditions using subspace identification," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Calgary, AB, Canada, Jul. 2009, pp. 1–6, doi: 10.1109/PES.2009.5275227.
[49] J. M. Seppänen, J. Turunen, M. Koivisto, N. Kishor, and L. C. Haarla, "Modal analysis of power systems through natural excitation technique," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1642–1652, Jul. 2014.
[50] D. J. Trudnowski and J. W. Pierre, "Overview of algorithms for estimating swing modes from measured responses," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Calgary, AB, Canada, Jul. 2009, pp. 1–8, doi: 10.1109/PES.2009.5275444.

KRISHNA RAO (Graduate Student Member, IEEE) received the B.E. degree in electrical engineering from the NMAM Institute of Technology (NMAMIT), Nitte, India, in 1995, and the M.Tech. and Ph.D. degrees in power systems from the National Institute of Technology Karnataka (NITK), Surathkal, India, in 2013 and 2021, respectively. He is currently an Assistant Professor with the Department of Electrical and Electronics Engineering, NMAMIT, Nitte (Deemed to be University). His research interests include power system dynamics and signal processing applications to power systems.

K. N. SHUBHANGA (Senior Member, IEEE) received the B.E. degree in electrical engineering and the M.Tech. degree in power systems from Mangalore University, India, in 1991 and 1994, respectively, and the Ph.D. degree from the Indian Institute of Technology, Bombay, in 2003.

He is currently a Professor with the National Institute of Technology Karnataka (NITK), Surathkal, India. He has authored a book titled *Power System Analysis: A Dynamic Perspective* published by Pearson India Education Services Pvt., Ltd., Noida, Uttara Pradesh, India, in 2018. His research interests include FACTS and power system dynamics.