Electric Power Systems Research 133 (2016) 113–120

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

# Distributed Prony analysis for real-world PMU data

Javad Khazaei[a], Lingling Fan[a,*], Weiqing Jiang[b], Durgesh Manjure[b]

[a] Department of Electrical Engineering, University of South Florida, Tampa, FL 33620, USA
[b] MISO Energy, 720 City Center Drive, Carmel, IN 46302, USA

ARTICLE INFO

*Article history:*
Received 1 July 2015
Received in revised form 6 November 2015
Accepted 9 December 2015
Available online 5 January 2016

*Keywords:*
Prony analysis
PMU data
Distributed optimization

ABSTRACT

Prony analysis has been applied in power system oscillation identification for decades. For a single PMU signal with 30 Hz sampling rate, merely applying Prony analysis cannot give accurate results of oscillating modes of power systems. This paper presents an analysis to show the effect of sampling rate on estimation accuracy and the mitigation methods to obtain accurate estimation. The methods include sampling rate reduction and multiple-signal Prony analysis. For multiple-signal Prony analysis, this paper proposes a distributed Prony analysis algorithm using consensus and subgradient update. This algorithm can be applied to multiple signals from multiple locations collected at the same period of time. This algorithm is scalable and can handle a large-dimension of PMU data by solving least square estimation problems with small sizes in parallel and iteratively. Real-world PMU data are used for analysis and validation. The proposed distributed Prony analysis shows being robust against sampling rate and generates reconstructed signals with better matching degree compared to the conventional Prony analysis for multiple signals.



## 1. Introduction

A power system is a massive system that can be perturbed by load changes, generator trips, faults or networks changes. Power system oscillations are common issues. To mitigate oscillations, oscillations should be identified and studied in a timely manner. There are two separate approaches to identify power system oscillations. The first approach is based on detailed dynamic model of the system such as: eigenvalue analysis or state space modeling [1]. Detailed modeling of a huge complicated power system is challenging and prone to errors. The second approach is based on measurements to identify oscillation modes. Measurement-based approach has been adopted by control engineers in practice. For example, equivalent system models will be constructed based on the measurement and further control strategies will be developed based on the identified system models.

With phasor measurement unit (PMU) data collected, electromechanical oscillation modes can be identified from these measurements. Several measurement-based system identification have been proposed for PMU data-based estimation, such as Kalman filters [2–4], least square estimation [5], and subspace algorithm [6]. Prony analysis is one of the most common measurement-based identification approaches to identify oscillatory modes. Prony analysis has been introduced by Hauer *et al* in power systems in 1990 [7,8]. The main idea is to directly estimate the frequency, damping and phase of modal components of a measured signal. An extension to Prony analysis is then introduced which allowed multiple signals to be analyzed at the same time resulting in one set of oscillatory modes [9].

Since then, Prony analysis has been applied in power system oscillation identification for decades. For PMU data with 30 Hz sampling rate, it is found that merely applying Prony analysis cannot give accurate results of oscillating modes of power systems. Zhou et al. have identified this issue in [10,11] and provided a solution. By re-sampling the PMU data to a lower sampling rate, the estimation will be more accurate. In this paper, an analysis is presented to show the effect of sampling rate on accuracy.

Mitigation methods are also presented in this paper to obtain accurate estimation. The two mitigation methods investigated include sampling rate reduction and multiple-signal Prony analysis. For multiple-signal Prony analysis, scalability is an issue given the large size of PMU data. A distributed algorithm is proposed in this paper to handle the issue of scalability. The objective of the algorithm is to have multiple Phasor Data Centers (PDCs) to conduct estimation at the same time. These PDCs will only utilize the local PMU data with limited information exchange from other PDCs. The computation effort is thus drastically reduced for each computing agent.

* Corresponding author. Tel.: +1 813 974 2031; fax: +1 813 974 5250.
*E-mail address:* linglingfan@usf.edu (L. Fan).

http://dx.doi.org/10.1016/j.epsr.2015.12.008
0378-7796/

Application of distributed optimization techniques has recently been introduced in system modes identification [12–14]. For example, in [12], distributed Prony analysis using alternating direction method of multipliers (ADMM) has been combined with centralized Prony method to estimate the slow frequency eigenvalues. Simulation data generated by PST [15] toolbox of IEEE 39-bus system is used to conduct Prony analysis.

While [13,12] have discussed the ADMM implementation, many details on Prony analysis have not been elaborated, e.g., sampling rate effect and validation through signal reconstruction. Further, the PMU data in [13] come from computer simulation. In this paper, real-world PMU data from Eastern Interconnection will be used for tests. The real-world PMU data has more complex characteristics.

This paper will develop a distributed Prony analysis algorithm using consensus and subgradient update. This algorithm can be applied to multiple signals from multiple locations collected the same period of time. This algorithm can handle a large-dimension of PMU data by solving least square estimation (LSE) problems with small sizes in parallel and iteratively. Moreover, convergence analysis is carried out to examine convergence. Robustness of the algorithm against sampling rate will also be examined. The rest of the paper is as follows: Section 2 describes the fundamentals of Prony analysis. An analysis of effect of sampling rate on estimation is presented in Section 3. Distributed Prony analysis including the general description and convergence analysisis described in Section 4. Section 5 presents case study results. Conclusion is presented in Section 6.

## 2. Fundamentals of Prony analysis

Consider a Linear-Time Invariant (LTI) system with the initial state of $x(t_0)$=$x_0$ at the time $t_0$, if the input is removed from the system, the dynamic system model can be represented as [16]:

$$\dot{x}(t) = Ax(t) \tag{1}$$

$$y(t) = Cx(t) \tag{2}$$

where $y \in \mathbb{R}$ is defined as the output of the system, $x \in \mathbb{R}^n$ is the state of the system, $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{1\times n}$ are system matrices. The order of the system is defined by $n$. If the $\lambda_i$, $p_i$, and $q_i$ are the $i$-th eigenvalues, right eigenvectors, and left eigenvectors of $n \times n$ matrix $A$, respectively, the (1) can be solved as:

$$\begin{aligned} x(t) &= \sum_{i=1}^{n} (q_i^T x_0) p_i e^{\lambda_i t} \\ &= \sum_{i=1}^{n} R_i x_0 e^{\lambda_i t} \end{aligned} \tag{3}$$

where $x_0$ is the initial state and $R_i = p_i q_i^T$ is a residue matrix. Based on (2), the $y(t)$ can be expressed as:

$$y(t) = \sum_{i=1}^{n} CR_i x_0 e^{\lambda_i t}. \tag{4}$$

Prony analysis directly estimates the parameters for the exponential terms in (4) by defining a fitting function in a basic form of:

$$\hat{y}(t) = \sum_{i=1}^{n} B_i e^{\sigma_i t} \cos(2\pi f_i t + \varphi_i) \tag{5}$$

The observed or measured $y(t)$ consists of $N$ samples which are equally spaced by $\Delta t$ as: $y(t_k)$=$y(k)$, $k$=1, ..., $N-1$. The basic assumption is to consider the signal record to be noise free and the order of the system can be set as: $n$ = $N$/2 [7]. Therefore, (5) can be recast in the exponential form as:

$$\begin{aligned} \hat{y}(t_k) &= \Re\left(\sum_{i=1}^{n} B_i e^{\lambda_i k \Delta t}\right) \\ &= \Re\left(\sum_{i=1}^{n} B_i z_i^k\right), \quad k = 1, \ldots, N \end{aligned} \tag{6}$$

where $N$ is the number of samples, $z_i$ are the eigenvalues of the system in discrete time domain, and $B_i$ is the residue of $z_i$. $z_i$ can be expressed as:

$$z_i = e^{\lambda_i \Delta t} \tag{7}$$

Due to the fact that $k$ = 1, ..., $N$, (6) can be expressed in matrix form as:

$$\begin{bmatrix} B_1 z_1^0 + \ldots + B_n z_n^0 \\ B_1 z_1^1 + \ldots + B_n z_n^1 \\ \vdots \\ B_1 z_1^{N-1} + \ldots + B_n z_n^{N-1} \end{bmatrix} = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix}. \tag{8}$$

Or in a simple form: $ZB$ = $Y$ as shown in (9).

$$\begin{bmatrix} z_1^0 & z_2^0 & \cdots & z_n^0 \\ z_1^1 & z_2^1 & \cdots & z_n^1 \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{N-1} & z_2^{N-1} & \cdots & z_n^{N-1} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix} \tag{9}$$

As the $z_i$ are the roots of the characteristic polynomial function of the system, in order to find the $z_i$, the coefficients of the polynomial need to be found first. The polynomial is formed as:

$$z^n - (a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_n z^0) = 0. \tag{10}$$

While the roots $z_i$ might be complex numbers, the system polynomial coefficients $a_i$ are real numbers. This feature helps develop algorithms since real numbers will be handled by computer algorithms while complex numbers cannot be directly handled.

From (10), we have

$$z^n = a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_n z^0. \tag{11}$$

Further, a linear prediction model (12) can be formulated since $y(k)$ is the linear combination of $z_i(k)$ based on (6). Therefore,

$$y(n) = a_1 y(n-1) + a_2 y(n-2) + \ldots + a_n y(0). \tag{12}$$

Enumerate the signal samples from $n$ step to $N$ step, we have (13): $Y$ = $Da$.

$$\underbrace{\begin{bmatrix} y(n) \\ \vdots \\ y(n+k) \\ \vdots \\ y(N) \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} y(n-1) & y(n-2) & \cdots & y(0) \\ \vdots & \vdots & \ddots & \vdots \\ y(n+k-1) & y(n+k-2) & \cdots & y(k) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & \cdots & y(N-n) \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} a(1) \\ \vdots \\ a(k) \\ \vdots \\ a(n) \end{bmatrix}}_{a} \tag{13}$$

*Remarks:* The dimension of $D$ matrix is $N-n+1$, $n$. If $n < N/2$, this is an over-determined linear equation and will be solved by the least square estimation (LSE). If $n > N/2$, the linear equations are under-determined and there are multiple solutions for $a$. When the $D$ matrix is square, there is a unique solution of $a$ and the match will be the best for. That is the reason that $n$ is selected to be close to $N/2$ [7].