Journal of Control, Automation and Electrical Systems (2018) 29:614–624
https://doi.org/10.1007/s40313-018-0402-y

Evaluation of a Maximum Likelihood Estimator for the Identification of Power Systems Oscillation Modes

João C. Y. Menezes[1] **· Aguinaldo S. e Silva**[1]

Received: 24 October 2017 / Revised: 6 February 2018 / Accepted: 23 July 2018 / Published online: 13 August 2018

Abstract
In this paper, a maximum likelihood (ML) estimator is applied for the estimation of power system oscillation modes. A regularization term is used in order to improve the estimation. An index is proposed to rank the modes and separate spurious from real modes. The ML estimator is compared with Prony method, and its advantages and limitations are discussed. Both methods are applied to synthetic systems and to real Phasor Measurement Unit (PMU) data acquired from the Brazilian Interconnected Power System (BIPS). The results show that the proposed maximum likelihood estimator is useful to complement and validate the results obtained by Prony analysis.

Keywords Power systems · Oscillation modes · Maximum likelihood estimation · Prony method

1 Introduction

Power system controllers, such as Power System Stabilizers (PSSs), keep the small-signal stability, ensuring that oscillation modes are well damped. Controllers tuning is usually performed using a model obtained by the linearization of the system equations around one or a set of operating points.

Model-based controller design and evaluation are inherently limited, since only a subset of operating conditions is employed, which does not cover the wide range of topologies, load and generation dispatch in which power systems operate.

The measurement-based estimation of the oscillation modes allows the evaluation of the performance of power system controllers in a wide range of operating conditions. Wide Area Monitoring System (WAMS), based on the use of synchronized measurements, provides system-wide measurements, allowing the identification of the system oscillation modes. These measurements can be obtained after large disturbances, originating transient (or ringdown) data, or in the normal system operating conditions, generating ambient data (Zhou et al. 2012; Leandro et al. 2015).

Large disturbances excite several oscillation modes, making their identification easier. Therefore, the damping of the oscillations modes and the effect of PSS can be clearly evaluated. The identified oscillation modes can also be used for system-wide model validation (Decker et al. 2010).

Prony method has been widely used for the identification of oscillation modes using transient data, although several estimation methods can be applied (Lu et al. 2012; Kamwa et al. 2011; Bronzini et al. 2007; Messina and Vittal 2006). However, it presents several shortcomings specially for noisy signals. Variants were proposed to improve its performance but usually the noise is dealt with by increasing the model order, which introduces spurious modes that must be separated from the system modes.

Identification methods based on ML take noise into account and theoretically can asymptotically achieve the Cramer-Rao lower bound (CRLB) (Kay 1993), which is not the case of linear estimators like Prony method. ML has found few applications for the identification of oscillation modes in power systems.

In this paper, a ML estimator is proposed for the identification of oscillation modes using data from the power system transient data. An index is proposed to rank the oscillation modes and to separate spurious modes from the real modes. A regularization term is used in order to improve the performance. The ML estimator is compared with Prony method and its advantages and limitations are discussed. Both methods are applied to synthetic systems and to real Phasor

✉ João C. Y. Menezes
joaocym@gmail.com

Aguinaldo S. e Silva
aguinaldo.s.silva@ufsc.br

[1] Department of Electrical Engineering, Federal University of Santa Catarina, Santa Catarina, Brazil

Measurement Unit (PMU) data acquired from the Brazilian Interconnected Power System (BIPS). The results show that the proposed maximum likelihood estimator is useful to complement and validate the results obtained by Prony analysis. This paper is organized as follows. In Sect. 2, the formulation of the ML estimator is introduced and the solution method is described. In Sect. 3, the Prony method is reviewed. In Sect. 4, results for data generated by simulation of two synthetic systems are presented. In Sect. 5, the identification is performed using real data, acquired by a low voltage WAMS installed in the BIPS. The conclusions are presented in Sect. 6.

2 ML Estimator

If the estimator does not attain the CRLB in all situations or it is impossible to achieve the CRLB, then, there is not a minimum variance unbiased (MVU) estimator. This situation is common and, in many cases, linear estimators are not a choice either, because the estimator can be highly nonlinear on data. Therefore, ML estimators become a viable alternative in many occasions (Kay 1993). The ML estimator is asymptotically Gaussian, unbiased and efficient.

In order to apply the ML to power systems, the transient response is modeled as a sum of exponentials:

$$\begin{aligned} s_d[n] &= \sum_{i=1}^{N_\lambda} (a_i + jb_i)(u_i + jv_i)^n \\ &= \sum_{i=1}^{N_\lambda} R_i z_i^n, \quad n = 0, 1, \ldots, N-1 \end{aligned} \tag{1}$$

where $R_i = a_i + jb_i$ is the i-th complex residue associated with the i-th discrete complex eigenvalue $z_i = u_i + jv_i$ and N_λ is the model order.

The i-th continuous eigenvalue is:

$$\lambda_i = \frac{\ln z_i}{T} = \alpha_i + j\omega_i \tag{2}$$

where T is the sampling period.

The observation model is:

$$x[n] = s_d[n; \boldsymbol{\theta}] + w[n], \quad n = 0, 1, \ldots, N-1 \tag{3}$$

where x is the observed data, $\boldsymbol{\theta}$ is a vector of parameters and w, in this case, is modeled as a White Gaussian Noise (WGN), with distribution $\mathcal{N}(0, \sigma^2)$.

The ML method is based on the maximization of the likelihood function (Kay 1993):

$$p(\mathbf{x}; \theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} \left(x[n] - s_d[n; \theta]\right)^2\right] \tag{4}$$

In order to maximize (4), the following cost function can be used:

$$\min J_{ML}(\mathbf{x}) = \sum_{n=0}^{N-1} \left(x[n] - s_d[n; \theta]\right)^2 \tag{5}$$

$s_d[n; \theta]$ is rewritten in the following form:

$$\begin{aligned} s_d[n; \boldsymbol{\theta}] = \sum_{i=1}^{N_m} e^{\alpha_i nT} \big(K_{1,i} \cos(\omega_i nT) \\ + K_{2,i} \sin(\omega_i nT)\big) \end{aligned} \tag{6}$$

where $\boldsymbol{\theta} = \left[\mathbf{K_1}' \ \mathbf{K_2}' \ \boldsymbol{\alpha}' \ \boldsymbol{\omega}'\right]'$ and $\mathbf{K_1}$, $\mathbf{K_2}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\omega}$ are the vectors formed by $K_{1,i}$, $K_{2,i}$, α_i and ω_i. This ensures that the time response will be a real vector, grouping together complex conjugate eigenvalues. In the case of real eigenvalues, $K_{2,i}$ and ω_i are set to 0.

After adding together the complex conjugate eigenvalues, (6) relates to (1) by:

$$\begin{aligned} K_{1,i} &= 2a_i \\ K_{2,i} &= -2b_i \end{aligned} \tag{7}$$

The main advantage of the ML estimator is its statistical properties, specially the asymptotic optimality. It is also easy to formulate, resulting in a least square problem when the noise is modeled as WGN, but it is a difficult problem to solve, due to the attributes of the cost function described in Sect. 2.2. This can potentially result in high computing time. However, ML works well even in low signal-to-noise ratio (SNR) conditions. In addition, it results in low order models with few to none spurious modes.

There are some implementations of the ML criterion, in general, resorting to ARMA and ARMAX models. Bresler and Macovski (1986) addresses the estimation problem with an ARMA model and a recursive algorithm, separating the computation of eigenvalues, which are nonlinear, and the computation of residues, which are linear. Tufts and Kumaresan (1980) and other works have a similar approach. These methods have been applied to acoustics, speech and signal processing but have not been explored deeply for power system applications, except for Dosiek et al. (2013), which uses a ML algorithm for the estimation of oscillation modes using ambient data. An ARMAX model and a Recursive Maximum Likelihood algorithm are employed while in this paper, transient data are used with the model given by a sum of exponentials.

2.1 Regularization Term

The chance of the model over-fitting the data increases with the number of degrees of freedom of the model. This generates spurious modes as some modes are used to accommodate noise, specially when low energy modes must be estimated and in case of low SNR signals, when the cost function decreases more due to fitting to the noise than to those low energy states (Bishop 1995).

If the signal generated by the identified model fits to the noise, it oscillates more rapidly resulting in regions with high curvature (second derivative). This can be prevented by penalizing the integral of the second derivative of the estimated response (Bishop 1995), using a regularization term. The regularization term, if well tuned, can reduce the computing time or help to find modes that are difficult to estimate.

The regularization function Ω is added to the cost function as a penalty function in order to control the effective complexity of the chosen model, enabling $s_d[n; \boldsymbol{\theta}]$ to fit more easily to the data $\mathbf{x}$ and reducing the effects of noise:

$$\tilde{J}_{ML} = J_{ML} + \nu \Omega \tag{8}$$

and is given by:

$$\Omega = \int \left(\frac{d^2 s_d}{dt^2} \right)^2 dt \tag{9}$$

where $t = nT$ and ν is the weight of the penalty function. The choice of the penalty weight is not straightforward, but several experiments for the test systems have shown that values around $\nu = 0.1/N$ usually lead to good results.

2.2 Optimization Method

The cost function in (8) is nonconvex with various local minima (Bresler and Macovski 1986). In Tufts and Kumaresan (1980), some example plots of the objective function of the ML problem are shown and analyzed. This requires changes in the formulation that makes it easier to solve or more sophisticated optimization tools must be employed. In this paper, the optimization is performed by SolvOpt, an implementation by Kuntsevich of Shor's r-algorithm (Shor et al. 1985; Kappel and Kuntsevich 2000).

2.3 Spurious Modes Detection

In order to rank and detect spurious modes, a modified information criterion based on Trudnowski (1994), called Geometric Mean Information Criterion (GIC), is used with the Prony method and with the ML estimator. The method is described in Algorithm 1.

If the weights α and β in (15) are tuned equally (which is usually the case), it results in the geometric mean of $\gamma_{T,i}$ and $\gamma_{F,i}$ raised to an arbitrary power, hence the name of the criterion.

Algorithm 1 Iterative GIC

Require: Estimated residues and eigenvalues.
Ensure: Reduced order model.

1: Calculate the model response to an impulse in time domain from the residues and eigenvalues:

$$g_{model}(kT) = \sum_{i=1}^{N_\lambda} R_i e^{kT\lambda_i} \quad k = 0, 1, \ldots, N-1 \tag{10}$$

where N is the number of samples. The frequency domain response is determined through Fourier transform:

$$G_{model}(j\omega_k) = \mathcal{F}(g_{model}(t_k)) \tag{11}$$

The frequency interval $\Delta\omega$ and the number of intervals N_ω are chosen accordingly to the bandwidth characteristics of the system.

2: Compute the impulse responses in time and frequency domains for each of the N_λ pairs of residue/eigenvalue (R_i, λ_i):

$$\begin{aligned} g_i(kT) &= R_i e^{kT\lambda_i} \quad k = 0, 1, \ldots, N-1 \\ G_i(j\omega_k) &= \mathcal{F}(g_i(kT)) \end{aligned} \tag{12}$$

for $i = 1, 2, \ldots, N_\lambda$.

3: Compute a time norm and a frequency norm for each pair (R_i, λ_i) defined as:

$$\gamma_{T,i} = \left(\frac{\|g_{model} - g_i\|^2}{N \|g_{model}\|^2} \right) \tag{13}$$

$$\gamma_{F,i} = \left(\frac{\|G_{model} - G_i\|^2}{N_\omega \|G_{model}\|^2} \right) \tag{14}$$

4: Compute GIC:

$$\Gamma_i = \gamma_{T,i}^{\alpha} \gamma_{F,i}^{\beta} \tag{15}$$

where α and β are nonnegative constants tuned to weight time and frequency domain responses.

5: The residue/eigenvalue pair with lowest Γ_i is denoted $(R_{m(1)}, \lambda_{m(1)})$ and its GIC is $\Gamma_{m(1)}$.

6: The values $g_{model}(kT)$ and $G_{model}(j\omega_k)$ are updated removing $(R_{m(1)}, \lambda_{m(1)})$:

$$\begin{aligned} g_{model}(kT) &= g_{model}(kT) - R_{m(1)} e^{t_k \lambda_{m(1)}} \\ &\quad k = 0, 1, \ldots, N-1 \\ G_{model}(j\omega_k) &= \mathcal{F}(g_{model}(kT)) \end{aligned} \tag{16}$$

7: The steps 3, 4, 5 and 6 are repeated for each $N_\lambda - 1$ remaining residue/eigenvalue pair.

8: The procedure is repeated for a total of N_v times, until the required order of the system model is achieved or all the modes are removed. The conjugate complex eigenvalues are evaluated and removed together.

9: The reduced order model is determined as:

$$G_r(s) = \sum_{i=1}^{N_v} \frac{R_{m(i)}}{s - \lambda_{m(i)}} \tag{17}$$

In Sects. 4 and 5, Algorithm 1 is used to sort out the eigenvalues of the models obtained with Prony analysis and the ML method, indicating which ones are more important for the responses of each model, according to (15).

The same metric can be used to measure the relative quality of a model as:

$$\begin{aligned}\gamma_{T,\mathrm{sub}} &= \left(\frac{\|g_{\mathrm{measure}} - g_{\mathrm{sub}}\|^2}{N\,\|g_{\mathrm{measure}}\|^2}\right)\\ \gamma_{F,\mathrm{sub}} &= \left(\frac{\|G_{\mathrm{measure}} - G_{\mathrm{sub}}\|^2}{N_\omega\,\|G_{\mathrm{measure}}\|^2}\right)\\ \Gamma_{\mathrm{sub}} &= \gamma_{T,\mathrm{sub}}^{\alpha}\gamma_{F,\mathrm{sub}}^{\beta}\end{aligned} \tag{18}$$

where g_{measure} corresponds to the noisy observations, G_{measure} is $\mathcal{F}(g_{\mathrm{measure}})$, g_{sub} is the time response of a model to be assessed, composed by a subset of the estimated model residues/eigenvalues pairs, and $G_{\mathrm{sub}} = \mathcal{F}(g_{\mathrm{sub}})$.

3 Conventional Prony Method

This method is found in many references such as Hauer et al. (1990) and Bos (2007). The advantages of the conventional Prony method include low computing times, even for high order models, and easy implementation. It is a simplification of the ML criteria (Bresler and Macovski 1986; Kumaresan et al. 1984) and assumes that the observations are an exact description of an exponential sum model, and the noise is not modeled (Bos 2007). Ideally, it should be used with noiseless signals, but in most cases it works well with noise, providing a good fit to the data, specially in high SNR scenarios. The drawbacks are the spurious modes that can be numerous when a high order model is used to accommodate noise. For low SNR, the performance of the method can be impaired. Several variants tackle these issues. In Trudnowski et al. (1999), the accuracy of the method is improved, by using multiple input signals. Trudnowski (1994) and Zhou et al. (2012) propose methods for spurious modes detection and order reduction. In Zhou et al. (2010), the issue of real-time mode estimation with PMU data is tackled with a recursive algorithm.

4 Results: Synthetic Data

In this section, synthetic data from two test systems are used to evaluate the ML method and compare it with Prony method. For the synthetic data, the eigenvalues and residues are known for the output signals.

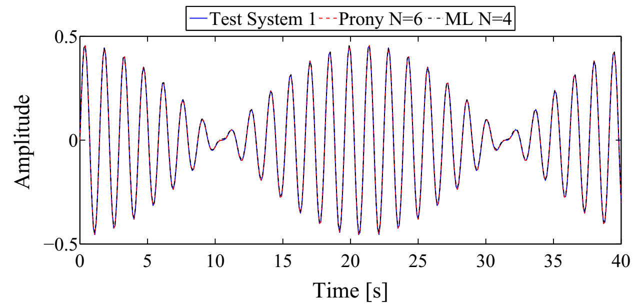

Fig. 1 Test System 1—Case 1 comparison

4.1 Test System 1

The first test system is a fourth-order system (Hauer et al. 1990). Although this is not a power system model, it is used for an initial assessment of the ML method, as compared with Prony method, and the performance of the GIC.

$$G_{T1}(s) = \frac{1}{s^2 + (2\pi/1.5)^2} + \frac{1}{s^2 + (2\pi/1.4)^2} \tag{19}$$

The system has two pairs of modes with zero damping, with frequencies of 0.67 Hz and 0.71 Hz. Three cases are considered:

- *Case* 1 noiseless signal with sampling time $T = 0.01$ s and a data time window of 40 s
- *Case* 2 SNR of 25 dB with sampling time $T = 0.01$ s and a data time window of 40 s
- *Case* 3 SNR of 10 dB with reduced sampling time $T = 0.001$ s and a data time window of 25 s

The SNR is measured by function awgn in MATLAB.

4.1.1 Test System 1: Case 1

In Fig. 1, the time responses from Test System 1 and the models obtained with the ML method, with order four, and with Prony method, with order six, are presented. These orders ensure that Prony and ML models have a time response that approximates the system time response.

In Table 1, the most significant eigenvalues, as given by the GIC, for each model, are presented. Prony method selected order led to two spurious eigenvalues. In the ML model, there are no spurious modes.

The GIC, $\Gamma_{\mathrm{sub},i}$, was calculated for each individual mode using (18), with $g_{\mathrm{sub},i}$ equal to the response of the individual mode i. The objective is to compare the discrepancy indexes obtained with the Prony method and the ML method with the indexes of the individual modes of the original Test System 1, using the known eigenvalues and residues. The results are shown in Table 2. The GICs of each estimated mode from

Table 1 Test System 1: noiseless–eigenvalues

Mode	Test System 1	Prony	ML
1	$\pm 4.488j$	$\pm 4.488j$	$\pm 4.488j$
2	$\pm 4.189j$	$\pm 4.189j$	$\pm 4.189j$

Table 2 Test System 1: noiseless–GIC

GIC	Test System 1	Prony	ML
$\Gamma_{sub,1}$	3.9954×10^{-6}	3.9954×10^{-6}	4.0214×10^{-6}
$\Gamma_{sub,2}$	4.5869×10^{-6}	4.5869×10^{-6}	4.5650×10^{-6}

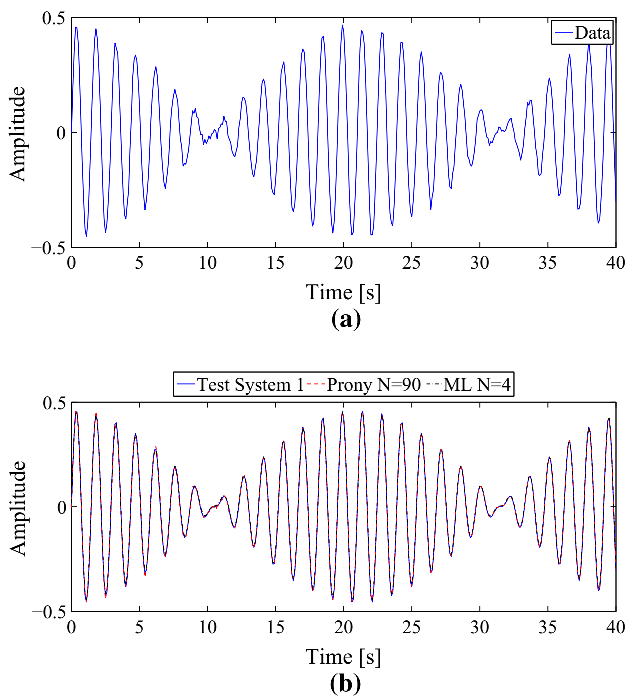

Fig. 2 Test System 1—Case 2. **a** Test System 1 impulse response with 25 dB SNR. **b** Test System 1—Case 2 comparison

Prony and ML models are close to the original, from Test System 1. Thus, the methods estimate the modes accurately.

4.1.2 Test System 1: Case 2

In Fig. 2a, the noisy system output is shown. In Fig. 2b, time responses of the models obtained by the ML method, with order four, and Prony method, with order 90, are presented. These orders give a close approximation of the real-time response. In this case, the same sample time, the same time window and approximately the same noise are used as in Hauer et al. (1990).

In Table 3, the main eigenvalues of each model are presented. In this case, there are 86 spurious eigenvalues in the 90th-order Prony model. In the ML model, there are no spurious modes.

Table 3 Test System 1: 25 dB SNR–eigenvalues

Mode	Test System 1	Prony	ML
1	$\pm 4.488j$	$\pm 4.488j$	$\pm 4.488j$
2	$\pm 4.189j$	$\pm 4.189j$	$\pm 4.189j$

Table 4 Test System 1: 25 dB SNR–GIC

GIC	Test System 1	Prony	ML
$\Gamma_{sub,1}$	4.0238×10^{-6}	4.0495×10^{-6}	4.0781×10^{-6}
$\Gamma_{sub,2}$	4.5687×10^{-6}	4.5756×10^{-6}	4.5591×10^{-6}

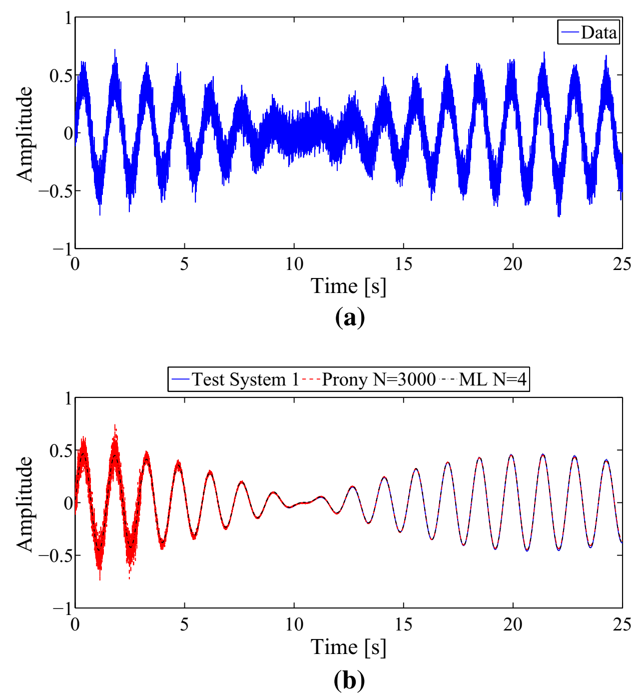

Fig. 3 Test System 1—Case 3. **a** Test System 1—impulse response with 10 dB SNR. **b** Test System 1—Case 3 comparison

The GICs calculated for each individual mode are shown in Table 4. The GICs of the estimated modes are close to the GICs of the modes of Test System 1.

4.1.3 Test System 1: Case 3

In Fig. 3a, the system output, contaminated with a 10 dB SNR noise, is presented. In Fig. 3b, time responses of the models obtained by the ML method, with order four, and Prony method, with order 3000, are presented. These orders give a close approximation of the real-time response.

Table 5 Test System 1: 10 dB SNR–eigenvalues

Mode	Test System 1	Prony	ML
1	$\pm 4.488j$	$-0.006 \pm 4.488j$	$\pm 4.488j$
2	$\pm 4.189j$	$-0.003 \pm 4.189j$	$\pm 4.189j$

Table 6 Test System 1: 10 dB SNR–GIC

GIC	Test System 1	Prony	ML
$\Gamma_{\text{sub},1}$	9.3795×10^{-8}	9.2712×10^{-8}	9.3855×10^{-8}
$\Gamma_{\text{sub},2}$	1.1384×10^{-7}	1.1471×10^{-7}	1.1349×10^{-7}

Table 7 Test System 1: model orders and computing times

Case	Order		Computing time (s)	
	Prony	ML	Prony	ML
Noiseless	6	4	0.0036	0.1571
25 dB	90	4	0.0234	0.1347
10 dB	3000	4	303.19	2.1947

In the Prony method, the selected order of the model results in 2996 spurious eigenvalues. In the ML model, there are no spurious modes.

A comparison of the four most significant eigenvalues, indicated by the GIC, for both methods, is shown in Table 5.

In Table 6, the GICs for each individual mode are shown. Again, the estimated modes are relatively close to modes of Test System 1, using the GIC metric.

4.1.4 Test System 1: Conclusions

The model orders and computing times for each case with Prony and ML methods are shown in Table 7.

Both the methods performed well, except for Prony method in the third case. For $T = 0.001$ s, even for higher SNR such as 25 dB, Prony method still performs poorly.

The algorithms were tested in a Windows 10 machine with Intel Core i5-3210M, 2.50 GHz and MATLAB. In the first and second cases, Prony method outperformed the ML method. In the last case, Prony method takes a longer time to find the solution.

The GICs of the individual modes estimated for Prony and ML methods, presented in Tables 2, 4 and 6, are close to those from the original system. Since the criterion also takes into account the residues, it is an indication that the methods could find the modes of Test System 1 accurately.

The GIC, relative to the observation data, given by Equation (18), was calculated for the Test System 1, the 4th-order Prony and the 4th-order ML estimations. In this case, g_{sub} in (18) is determined using four eigenvalues, that is, all significant modes are considered. In Table 8, the GICs calculated

Table 8 Test System 1: GIC

Case	Test System 1	Prony	ML
Noiseless	0	1.2711×10^{-12}	2.5493×10^{-12}
25 dB	7.4363×10^{-11}	6.9298×10^{-11}	7.5923×10^{-11}
10 dB	2.4148×10^{-11}	3.7152×10^{-11}	2.3612×10^{-11}

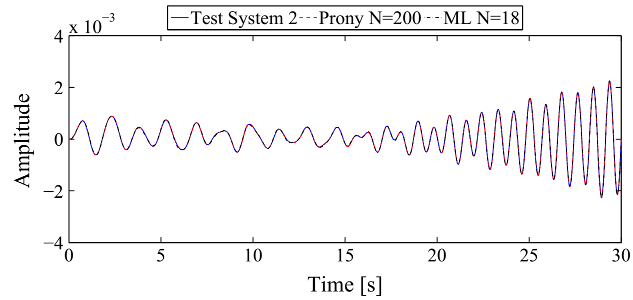

Fig. 4 Test System 2—Case 1 comparison

for each case is shown. In Cases 2 and 3, the indexes for Test System 1 are different from 0 because the noiseless response is being compared with the noisy data. If the estimated model has an index close to index of Test System 1, it is assumed that it is a good estimate.

While in the noiseless case (Case 1), the GIC is closer to the GIC of Test System 1 for the Prony model, for the cases with noise (Cases 2 and 3), the ML model gives a closer GIC.

4.2 Test System 2

Test System 2 is the New England-New York system (Pal and Chaudhuri 2005). The system has 95 states and two unstable modes.

Three data sets are used:

- *Case* 1 noiseless signal with $T = 1/60$ s and data time window of 30 s
- *Case* 2 SNR of 25 dB with $T = 1/60$ s and a data time window of 30 s
- *Case* 3 SNR of 10 dB with $T = 0.001$ s and a data time window of 25 s

The disturbance is a unit impulse in the electric power at Bus 13, and the output is the speed of Generator 9.

4.2.1 Test System 2: Case 1

In Fig. 4, time responses of Test System 2 and the models obtained with the ML method, with order 18, and with the Prony method, with order 200, are presented. These orders give a close approximation of the time response of Test System 2.

Table 9 Test System 2: noiseless–eigenvalues

Mode	Test System 2	Prony	ML
1	$0.161 \pm 7.29j$	$0.162 \pm 7.291j$	$0.16 \pm 7.29j$
2	$-0.049 \pm 4.131j$	$-0.048 \pm 4.129j$	$-0.049 \pm 4.139j$
3	$0.079 \pm 6.983j$	$0.079 \pm 6.979j$	$0.085 \pm 6.986j$
4	$-0.08 \pm 2.517j$	$-0.08 \pm 2.517j$	$-0.082 \pm 2.511j$

Table 10 Test System 2: noiseless–GIC

GIC	Test System 2	Prony	ML
$\Gamma_{sub,1}$	1.7049×10^{-7}	1.7177×10^{-7}	1.7460×10^{-7}
$\Gamma_{sub,2}$	4.8454×10^{-6}	4.8412×10^{-6}	4.9790×10^{-6}
$\Gamma_{sub,3}$	4.7575×10^{-6}	4.7660×10^{-6}	4.6182×10^{-6}
$\Gamma_{sub,4}$	6.7241×10^{-6}	6.7232×10^{-6}	6.7226×10^{-6}

In Table 9, the eight most significant eigenvalues, as given by the GIC, of each model, are shown. These modes are close to the dominant modes of the system. The remaining 192 eigenvalues, in the Prony model, and ten eigenvalues, in the ML model, are spurious.

The individual GIC for each mode is presented in Table 10. Modes 3 and 4 in the ML model are slightly different. Those modes are more difficult to estimate as discussed in Sect. 4.2.4.

4.2.2 Test System 2: Case 2

In Fig. 5a, the Test System 2 impulse response, contaminated with a 25 dB SNR noise, is presented. In Fig. 5b, time responses of the models obtained by the ML method, with order 16, and Prony method, with order 500, are presented. These orders give a close approximation of the real time response.

In Table 11, the eight most significant eigenvalues, as given by the GIC, for both methods, are shown. Again, these modes are close to the dominant modes of the system. The remaining 492 eigenvalues, in the Prony model, and eight eigenvalues, in the ML model, are spurious. In this case, a good fit with the ML method was achieved for a lower order than in Case 1.

The GIC calculated for each individual mode can be seen in Table 12. The indexes of all modes are close to what is expected, except $\Gamma_{sub,3}$ obtained with the ML method.

4.2.3 Test System 2: Case 3

In Fig. 6a, the 10 dB SNR time response of Test System 2 is shown. In Fig. 6b, time responses of the models obtained by the ML method, with order 18, and Prony method, with order 3500, are presented. These orders give a close approximation of the real-time response.

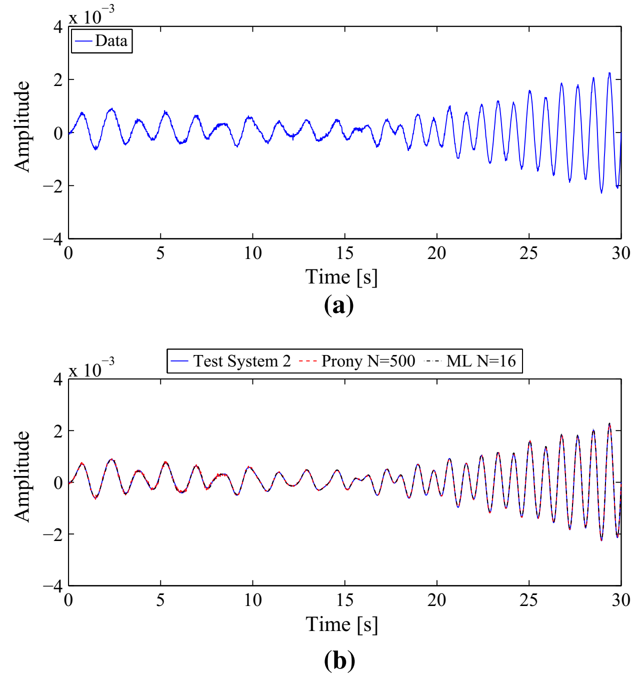

Fig. 5 Test System 2—Case 2. **a** Test System 2 impulse response with 25 dB noise. **b** Test System 2—Case 2 comparison

Table 11 Test System 2: 25 dB SNR–eigenvalues

Mode	Test System 2	Prony	ML
1	$0.161 \pm 7.29j$	$0.159 \pm 7.291j$	$0.163 \pm 7.29j$
2	$-0.049 \pm 4.131j$	$-0.049 \pm 4.129j$	$-0.051 \pm 4.129j$
3	$0.079 \pm 6.983j$	$0.076 \pm 6.976j$	$0.089 \pm 6.977j$
4	$-0.08 \pm 2.517j$	$-0.081 \pm 2.518j$	$-0.074 \pm 2.51j$

Table 12 Test System 2: 25 dB SNR–GIC

GIC	Test System 2	Prony	ML
$\Gamma_{sub,1}$	1.7235×10^{-7}	1.7403×10^{-7}	1.7427×10^{-7}
$\Gamma_{sub,2}$	4.8448×10^{-6}	4.8368×10^{-6}	4.8303×10^{-6}
$\Gamma_{sub,3}$	4.7536×10^{-6}	4.7668×10^{-6}	4.3612×10^{-6}
$\Gamma_{sub,4}$	6.7330×10^{-6}	6.7316×10^{-6}	6.7626×10^{-6}

In Table 13, the eight most significant eigenvalues, as given by the GIC, for both methods, are presented. These modes are close to the system modes. The remaining 3492 eigenvalues, in the Prony model, and ten eigenvalues, in the ML model, are spurious.

In Table 14, the GICs for each individual mode are shown. The ML model gives a better estimation as compared with Prony method. The 4th mode obtained with Prony method presents the GIC with the biggest discrepancy.

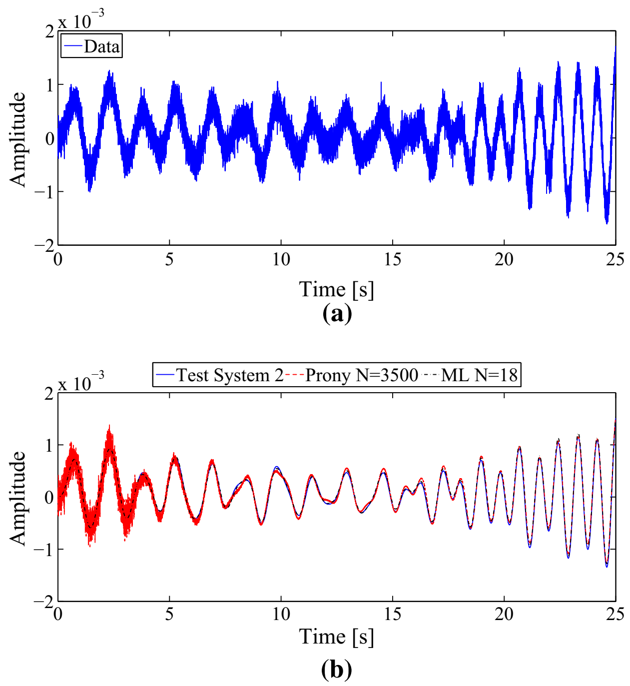

Fig. 6 Test System 2—Case 3. **a** Test System 2 impulse response with 10 dB SNR. **b** Test System 2—Case 3 comparison

Table 13 Test System 2: 10 dB SNR–eigenvalues

Mode	Test System 2	Prony	ML
1	$0.161 \pm 7.29j$	$0.168 \pm 7.273j$	$0.174 \pm 7.3j$
2	$-0.049 \pm 4.131j$	$-0.05 \pm 4.124j$	$-0.047 \pm 4.13j$
3	$-0.08 \pm 2.517j$	$-0.089 \pm 2.511j$	$-0.072 \pm 2.517j$
4	$0.079 \pm 6.983j$	$-0.001 \pm 6.839j$	$0.116 \pm 6.923j$

Table 14 Test System 2: 10 dB SNR–GIC

GIC	Test System 2	Prony	ML
$\Gamma_{\mathrm{sub},1}$	1.2501×10^{-7}	1.1351×10^{-7}	1.2592×10^{-7}
$\Gamma_{\mathrm{sub},2}$	2.2289×10^{-7}	2.2078×10^{-7}	2.2290×10^{-7}
$\Gamma_{\mathrm{sub},3}$	5.2038×10^{-7}	5.1961×10^{-7}	5.2329×10^{-7}
$\Gamma_{\mathrm{sub},4}$	3.8993×10^{-7}	5.2516×10^{-7}	3.7453×10^{-7}

In order to illustrate how the GIC can be used to detect spurious modes, the criterion was applied to the model identified with the ML method considering the two, four, six, eight and ten most significant eigenvalues. The results are shown in Table 15. The GIC gets smaller when 8 eigenvalues, known to be correct from the eigenvalues of Test System 2, are added to the response. After the inclusion of the first spurious mode, the GIC value rises. However, the criterion value may continue roughly the same or even reduce slightly after the inclusion of a spurious mode in the reduced model. Therefore, after calculating the GIC for different model orders, the user has to define a threshold ε given by

$$\frac{\left\|\Gamma_{\mathrm{sub},i} - \Gamma_{\mathrm{sub},i-2}\right\|}{\left\|\Gamma_{\mathrm{sub},i-2}\right\|} < \varepsilon \tag{20}$$

to determine if a mode should be considered spurious. A value $\epsilon = 0.01$ proved adequate for the results presented in this paper.

An alternative is to define a desired model order and reduce the order of the estimated system, to represent only the system dominant dynamics (Trudnowski 1994). Another procedure that can help to detect spurious modes is changing the order of the identified ML or Prony model. Spurious modes tend to vary between different order models, while the system modes remain the same.

Table 15 Test System 2: different order GIC

Order	GIC
2	1.2592×10^{-7}
4	9.4895×10^{-9}
6	3.8145×10^{-9}
8	8.3048×10^{-10}
10	8.2363×10^{-9}

Table 16 Test System 2: model orders and computing times

Case	Order		Computing time	
	Prony	ML	Prony	ML
Noiseless	200	18	0.3144	5.9812
25 dB	500	16	1.2111	5.6473
10 dB	3500	18	397.152	78.113

4.2.4 Test System 2: Conclusions

This test system was specially difficult to optimize in the ML method. The unstable mode $0.079 \pm 6.983j$ is hard to estimate and its identification depends on the optimization parameters such as the weight of regularization term and the order of the model, which had to be fine tuned to find this mode. The identification of this mode may also depend on the optimization starting point.

Although the original system order is 95, only eight of the system eigenvalues could be found in all cases. The modes that could not be estimated are nondominant or have low energy, making them more difficult to estimate. However, the most significant modes were estimated with good precision.

In Table 16, the orders used to obtain the models with Prony and ML methods along with the computing times in each case are shown. Prony method outperforms the ML method unless in case of heavy noise and lower sample time.

Table 17 Test System 2: GIC

Case	Test System 2	Prony	ML
Noiseless	1.2392×10^{-10}	1.2028×10^{-10}	3.7887×10^{-10}
25 dB	1.6897×10^{-10}	1.6631×10^{-10}	1.6752×10^{-10}
10 dB	9.2709×10^{-10}	1.4623×10^{-9}	8.3048×10^{-10}

The GIC of the individual modes estimated by Prony and ML methods in Tables 10, 12 and 14 were close to the ones of the original system, except in some cases such as the mode $0.079 \pm 6.983j$ estimated by the ML method in Cases 1 and 2. This is due to the fact that it is a difficult mode to estimate with the ML method. In Case 3, with a low SNR, Prony method could not estimate the modes accurately.

The GIC algorithm removes one mode at a time and then reevaluates the indexes. Therefore, there are discrepancies between the ranking order and the GIC of individual modes, as seen in Cases 1, 2 and 3.

In Table 17, the GICs of the 8th-order reduced Test System 2, 8th-order reduced Prony model and 8th-order reduced ML model, relative to the input data of each case, are shown.

Again, in Cases 2 and 3, the GIC obtained with the ML method is closer to the GIC of Test System 2 than that of Prony method. In Case 1, the noiseless case, the GIC of Prony model is closer than that of ML model.

In order to show the importance of the regularization term, in Case 1, (8) was used with ν set to 0. The ML method fails to estimate at least the eight main modes correctly. Orders up to 60 were tested. The 60th-order model took 58.78 seconds to optimize. In Cases 2 and 3 the objective function without the regularization term also fails.

5 Results: Real Data

In this section, real data from a disturbance in the BIPS, acquired by PMU of a low voltage WAMS (Prioste et al. 2011), are used. This was a large disturbance detected at 12:30 PM, January, 30th, 2017.

The data from the PMUs, the difference between the voltage angle of a bar in Northeastern Brazil and the voltage angle of a bar in Southern Brazil, are shown in Fig. 7a. The input data window is taken after the large disturbance near 2000 s. The results of the model identification using Prony method and the ML method are shown in Fig. 7b.

For the results, orders of 200 and 16 were used for Prony method and for the ML method, respectively. These orders ensure a close approximation between the system and the identified models time responses. The computing time was 0.16 s for the Prony method and 10.05 s for the ML method.

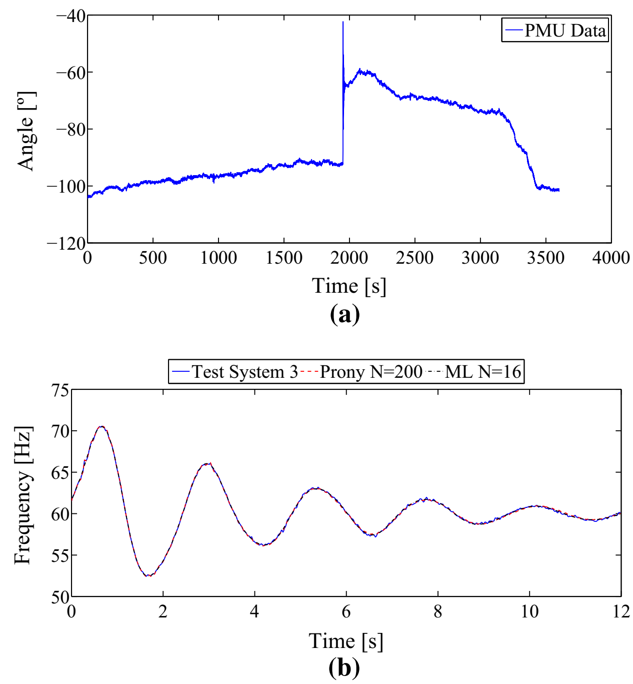

Fig. 7 Test System 3. **a** Test System 3 PMU data. **b** Test System 3—Response comparison

Table 18 Test System 3: eigenvalues

Mode	Prony	ML	Subspace
1	$-0.249 \pm 2.62j$	$-0.249 \pm 2.633j$	-0.233 ± 2.576
2	$-0.435 \pm 5.944j$	$-0.36 \pm 6.023j$	-0.454 ± 5.655
3	$-0.195 \pm 3.833j$	$-0.278 \pm 3.713j$	-0.1889 ± 3.77
4	$-1.646 \pm 11.407j$	$-1.841 \pm 3.167j$	

In Table 18, the most significant modes estimated by the Prony method and by the ML method, in the range of electromechanical oscillations, ranked according to the residues and GIC, are shown.

In Table 19, the GIC evaluated for each individual mode is shown. The GIC index of Mode 3 gives the largest discrepancy between the Prony and ML models.

The GIC for the 6th-order Prony model is 1.3100×10^{-15} and for the 6th-order ML model is 1.1114×10^{-13}. Those values indicate that the response of the 6th-order model obtained with Prony method is closer to the noisy observation data. However, it does not necessarily mean that the modes were estimated correctly.

Since a detailed analytical model is not available for this operating condition, in order to compare with the estimated modes, the canonical variate algorithm (CVA) subspace identification method (Overschee and Moor 1996) is used to track the oscillation modes using ambient data in a time window of ten minutes after the large disturbance. The results for the subspace method, with a model of order twelve, provided

Table 19 Test System 3: GIC

GIC	Prony	ML
$\Gamma_{\text{sub},1}$	2.1581×10^{-13}	2.0558×10^{-13}
$\Gamma_{\text{sub},2}$	8.3615×10^{-10}	8.5006×10^{-10}
$\Gamma_{\text{sub},3}$	8.6841×10^{-10}	7.9038×10^{-10}
$\Gamma_{\text{sub},4}$	8.5544×10^{-10}	8.5636×10^{-10}

three dominant modes, also given in Table 18. These results show that the Prony method and the ML method give consistent estimation of the most significant modes, according to GIC, and close to the estimated by the subspace method. The remaining pairs in Table 18 are discrepant and are deemed to be spurious. When the model order is increased, they do not show consistency between different model orders and estimation methods.

6 Conclusions

Power systems oscillation modes are correctly estimated by the ML method from low order models, with reduced number of spurious modes. A regularization term may improve the estimation of the system modes. Some oscillation modes are more difficult to identify, that is, their identification may depend on the regularization function weight and also on the initial parameters in the optimization procedure. The optimization in the ML method takes most of the computing time, making the ML method computationally costlier than Prony method. Therefore, improvements in the optimization can reduce the computing time.

The conventional Prony method has in most cases a good performance. However, for signals with heavy noise, it requires high dimension models in order to fit the noise. This generates many spurious modes which must be sorted out from the real ones, requiring good spurious modes detection method with sophisticated statistics. Unless a high model order is required, Prony method is computationally faster than the ML method. For noisy measurements and lower model orders, Prony method may fail to estimate the main system modes correctly. For both methods, it is difficult to initially determine the model order that gives a good fit compared with the real system.

Since power system data may present low SNR, a ML estimator may be a valuable tool to complement and validate the results obtained from Prony analysis.

The performance of the ML estimator depends on the optimization method. It can be improved by suited optimization methods. The nonlinearity and nonstationarity of power systems still require more development of the identification methods and are topics for further research.

Acknowledgements The first author acknowledges CAPES for financial support.

References

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford: Oxford university press.

Bos, A. V. (2007). *Parameter estimation for scientists and engineers*. New York: Wiley.

Bresler, Y., & Macovski, A. (1986). Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *34*(5), 1081–1089. https://doi.org/10.1109/TASSP.1986.1164949.

Bronzini, M., Bruno, S., Benedictis, MD., Scala, ML. (2007). Power system modal identification via wavelet analysis. In: *2007 IEEE lausanne power tech* (pp. 2041–2046). https://doi.org/10.1109/PCT.2007.4538632.

Decker, IC., e Silva, AS., da Silva, RJG., Agostini, MN., Martins, N., Prioste, FB. (2010). System wide model validation of the Brazilian interconnected power system. In: *2010 IEEE power and energy society general meeting* (pp. 1–8). IEEE.

Dosiek, L., Pierre, J. W., & Follum, J. (2013). A recursive maximum likelihood estimator for the online estimation of electromechanical modes with error bounds. *IEEE Transactions on Power Systems*, *28*(1), 441–451. https://doi.org/10.1109/TPWRS.2012.2203323.

Hauer, J. F., Demeure, C. J., & Scharf, L. L. (1990). Initial results in Prony analysis of power system response signals. *IEEE Transactions on Power Systems*, *5*(1), 80–89. https://doi.org/10.1109/59.49090.

Kamwa, I., Pradhan, A. K., & Joos, G. (2011). Robust detection and analysis of power system oscillations using the Teager–Kaiser energy operator. *IEEE Transactions on Power Systems*, *26*(1), 323–333. https://doi.org/10.1109/TPWRS.2010.2046503.

Kappel, F., & Kuntsevich, A. V. (2000). *Computational Applications and Optimization*. Berlin: Springer.

Kay, S. M. (1993). *Fundamentals of statistical signal processing, volume I: Estimation theory*. Upper Saddle River: Prentice Hall.

Kumaresan, R., Tufts, D. W., & Scharf, L. L. (1984). A Prony method for noisy data: Choosing the signal components and selecting the order in exponential signal models. *Proceedings of the IEEE*, *72*(2), 230–233. https://doi.org/10.1109/PROC.1984.12849.

Leandro, R., e Silva, A. S., Decker, I., & Agostini, M. (2015). Identification of the oscillation modes of a large power system using ambient data. *Journal of Control, Automation and Electrical Systems*, *26*(4), 441–453.

Lu, W., Du, X., Ding, J., Wang, X. (2012). Modal parameter identification based on fast Fourier transform and Hilbert Huang transform. In: *2012 2nd international conference on consumer electronics, communications and networks (CECNet)* (pp. 2703–2706). https://doi.org/10.1109/CECNet.2012.6201909.

Messina, A. R., & Vittal, V. (2006). Nonlinear, non-stationary analysis of interarea oscillations via Hilbert spectral analysis. *IEEE Transactions on Power Systems*, *21*(3), 1234–1241. https://doi.org/10.1109/TPWRS.2006.876656.

Pal, B., & Chaudhuri, B. (2005). *Robust control in power systems*. Berlin: Springer.

Prioste, FB., e Silva, AS., Decker, IC. (2011). Monitoring oscillations modes of the Brazilian interconnected power system using ambient data. In: *2011 IEEE trondheim PowerTech* (pp. 1–7) .IEEE.

Shor, N. Z., Kiwiel, K. C., & Ruszczynski, A. (1985). *Minimization methods for non-differentiable functions* (Vol. 3). Berlin: Springer.

Trudnowski, D. I. (1994). Order reduction of large-scale linear oscillatory system models. *IEEE Transactions on Power Systems*, *9*(1), 451–458. https://doi.org/10.1109/59.317578.

Trudnowski, D. J., Johnson, J. M., & Hauer, J. F. (1999). Making Prony analysis more accurate using multiple signals. *IEEE Transactions on Power Systems*, *14*(1), 226–231. https://doi.org/10.1109/59.744537.

Tufts, D., Kumaresan, R. (1980). Improved spectral resolution ii. In: *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '80* (vol. 5, pp. 592–597). https://doi.org/10.1109/ICASSP.1980.1170988.

Van Overschee, P., & De Moor, B. L. (1996). *Subspace identification for linear systems: Theory, implementation, applications* (Vol. 3). Dordrecht: Kluwer academic publishers .

Zhou, N., Huang, Z., Tuffner, F., Pierre, J., Jin, S. (2010). Automatic implementation of Prony analysis for electromechanical mode identification from phasor measurements. In: *IEEE PES General Meeting* (pp. 1–8). https://doi.org/10.1109/PES.2010.5590169.

Zhou, N., Pierre, J. W., & Trudnowski, D. (2012). A stepwise regression method for estimating dominant electromechanical modes. *IEEE Transactions on Power Systems*, *27*(2), 1051–1059. https://doi.org/10.1109/TPWRS.2011.2172004.