RESEARCH   **Open Access**

New algorithm for mode shape estimation based on ambient signals considering model order selection

Chao Wu[1*], Chao Lu[2] and Yingduo Han[2]

Abstract

Using time-synchronized phasor measurements, a new signal processing approach for estimating the electromechanical mode shape properties from ambient signals is proposed. In this method, Bayesian information criterion and the ARMA($2n$,$2n-1$) modeling procedure are first used to automatically select the optimal model order, and the auto regressive moving averaging models are built based on ambient data, then the low-frequency oscillation modal frequency and damping ratio are identified. Next, Prony models of ambient signals are presented, and the mode shape information of multiple dominant interarea oscillation modes are simultaneously estimated. The advantages of the new ARMA-P method are demonstrated by its applications in both a simulation system and measured data from China Southern Power Grid.

Keywords: Electromechanical dynamics, Mode shape, Ambient signal, Model order selection, ARMA-P method

Introduction

Modal frequency, modal damping ratio, mode shape magnitude, and mode shape angle are the key parameters describing the electromechanical modal properties of a power system [1]. Similar to former two, the mode shape properties that describe the participation of state variables in a particular mode are of vital importance for the safety and reliable operation of the system. Near real-time knowledge of mode shape characteristics provide critical information for the optimization of generators and/or load shedding, in order to improve the damping of the dangerously low-damped modes of power systems.

In general, the analysis of mode shape properties can be accomplished using two basic approaches: the eigen-analysis of a small signal model [1], or as shown in this article, signal processing of the time-synchronized measurements. One important advantage of the signal-based method is that the identification is dependent on the large, complex system model. In this category, the Prony method [2,3] and the Eigensystem Realization Algorithm

[4,5] are widely used. However, their applications are generally limited only for ringdown signals, which are relatively few in actual power grids. Ambient signals, caused by low-level stochastic disturbances, are more frequently and easily collected in real systems. Some publications have offered algorithms for identifying the modal frequency and damping ratio from ambient data [6-9], which fully demonstrate that this kind of signal includes abundant information about the system. Only recently the mode shape has been considered [10-14]. Based on the relationship between the cross- and power spectral densities and mode shapes, an approach for estimating the mode shapes using spectral method is presented in [10]. Liu and Venkarasubramanian [11] applied the frequency domain decomposition method to the mode shape identification. Then, the channel matching method was introduced in [12] and refined in [13], in which a narrowband bandpass filter must firstly be used to extract one single mode from multiple modes. In this case, the changes of operation modes in actual power system would inevitably cause the ineffectiveness of the prefixed filter, and influence the estimation accuracy. In [14], the transfer function method was proposed which showed that mode shape could be calculated by evaluating a transfer function, constructed

* Correspondence: wuchao@szu.edu.cn
[1]College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518060, China
Full list of author information is available at the end of the article

between a pair of system outputs, at the mode of interest. However, as the key factors of transfer function, the selection of model order are not taken into account in these articles, this will directly affect the accuracy of the mode shape analysis results.

In this article, in order to identify the low-frequency oscillation mode shape properties based on ambient signals, a new multiple modes estimation method called the auto regressive moving averaging-Prony (ARMA-P) is proposed. In addition, the problem of optimal model order selection in ARMA modeling is considered, which would improve the efficiency and accuracy of the new ARMA-P method.

The remainder of this article is organized as follows. Section 2 describes the feasibility and the theoretical basis of using the ARMA and Prony models together to extract the mode shapes from ambient data. In Section 3, different model order selection criteria are comparatively discussed, and ARMA $(2n,2n − 1)$ modeling procedures are adopted to optimize the order selection path. To estimate the mode shape characteristics based on ambient signals, the equations of ARMA-P method are derived in Section 4. Sections 5 and 6 provide the simulation and actual system examples, respectively. The results demonstrate that the new ARMA-P method can effectively estimate the mode shape properties from ambient data. Conclusions are provided in Section 7.

Theoretical basis of ARMA-P method

It is well known that under small-signal disturbance conditions, the power system may be linearized and represented in state space form [1],

$$\dot{\underline{x}}(t) = A\underline{x}(t) + B\underline{q}(t) \tag{1}$$

where $\underline{x}$ is the $n \times 1$ system state vector, including machine rotor angles and velocities. Input vector $\underline{q}$ (order $m \times 1$) is a hypothetical random-noise source vector perturbing the system. Under ambient conditions, input $\underline{q}$ is typically conceptualized as noises produced by random loads switching in power systems. A (order $n \times n$) is the state matrix and B (order $n \times m$) is the input matrix. Measurement-based electromechanical mode estimation assumes that the power system is in the steady-state condition described by (1).

It has been well established that the eigensolution of the state matrix A in (1) provides all the required information to completely describe the modal properties of power system. The reader can refer to [1] for more detail. A brief review of these properties is described here.

The eigenvalues and eigenvectors associated with A are

$$|\lambda_k I − A| = 0 \tag{2a}$$

$$A\underline{u}_k = \lambda_k \underline{u}_k \tag{2b}$$

$$\underline{v}_k A = \lambda_k \underline{v}_k \tag{2c}$$

where λ_k is the kth eigenvalue ($k = 1 \ldots n$), $\underline{u}_k = [u_{1,k}, u_{2,k}, \ldots u_{n,k}]^T$ is the kth right eigenvector, $\underline{v}_k = [v_{k,1}, v_{k,2}, \ldots v_{k,n}]$ is the kth left eigenvector.

Then the linear transformation defined in (3) is applied to the system (1), and the system state $\underline{x}_i$ is calculated, shown in (4).

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} u_{1,1} & \cdots & u_{1,n} \\ \vdots & \ddots & \vdots \\ u_{n,1} & \cdots & u_{n,n} \end{bmatrix} \begin{bmatrix} z_1 \\ M \\ z_n \end{bmatrix} \tag{3}$$

$$x_i = \sum_{k=1}^{n} \left\{ u_{i,k} e^{\lambda_k t} \left[z_k(0) + \sum_{l=1}^{n} \sum_{j=1}^{m} v_{k,l} b_{l,j} \int_0^t e^{-\lambda_k \tau} q(\tau) d\tau \right] \right\} \tag{4}$$

where $b_{l,j}$ is the lth row jth column element in the input matrix B, $z_l(0)$ is the initial value of z_k.

Equation (4) provides information on how the modes are combined to create the system states. The element $u_{i,k}$ (the ith element of $\underline{u}_k$) provides critical information on how the ith state (generators and other dynamic devices) participates in the kth oscillation mode. The magnitude of $u_{i,k}$ determines the intensity level for the state variable x_i to participate the kth oscillation mode, and the angle of $u_{i,k}$ determines the oscillation phase of the state variable x_i in the kth oscillation mode.

Thus, comparing the elements $u_{i,k}$ and $u_{j,k}$ describes the mode shape information between states i and j. The mode shape magnitude is then defined as the ratio of the magnitudes of $u_{i,k}$ and $u_{j,k}$

$$\text{Magnitude} = \frac{|u_{i,k}|}{|u_{j,k}|} \tag{5}$$

and the mode shape angle is the difference between the angles of $u_{i,k}$ and $u_{j,k}$

$$\text{Angle} = \angle u_{i,k} − \angle u_{j,k} \tag{6}$$

Then rewrite (4) as

$$x_i(t) = \sum_{k=1}^{n} \Phi_{i,k}(t) e^{\lambda_k t} \tag{7}$$

where

$$\Phi_{i,k}(t) = u_{i,k} \left[z_k(0) + \sum_{l=1}^{n} \sum_{j=1}^{m} v_{k,l} b_{l,j} \int_0^t e^{-\lambda_k \tau} q(\tau) d\tau \right] \tag{8}$$

On the other hand, the signal x_i shown in (7) is sampled with the time interval T. Using the sum of

weighted exponential components to fit the sampled signal, Prony model is adopted to describe the approximate signal as follows

$$\hat{x}_i(\kappa) = \sum_{k=1}^{n} \Psi_{i,k}(\kappa)e^{\lambda_k T\kappa} = \sum_{l=1}^{n} A_{i,k}e^{j\theta_{i,k}}e^{(\alpha_k+j2\pi f_k)T\kappa}$$

(9)

where $\kappa = 0,1\ldots N-1$, N is the data length, n is the number of oscillation modes.

Each term in (9) has four elements: the damping factor α_k, the frequency f_k, the magnitude $A_{i,k}$, and the angle $\theta_{i,k}$. Each exponential component with a different frequency is viewed as one unique mode of the original signal.

From (7) and (9), it can be deduced that on the sample point

$$\sum_{k=1}^{n} \Phi_{i,k}(\kappa T)e^{\lambda_k T\kappa} = \sum_{k=1}^{n} \Psi_{i,k}(\kappa)e^{\lambda_k T\kappa}$$

(10)

Obviously

$$\Psi_{i,k}(\kappa) = \Phi_{i,k}(\kappa T)$$

(11)

Taking (8) into consideration, $\Psi_{i,k}$ is time-varying, defined as

$$\Psi_{i,k}(\kappa) = \Psi_{i,k0} + \Delta\Psi_{i,k}(\kappa)$$

(12)

where

$$\Psi_{i,k0} = u_{i,k}z_k(0)$$

(13a)

$$\Delta\Psi_{i,k}(\kappa) = u_{i,k}\sum_{l=1}^{n}\sum_{j=1}^{m}v_{k,l}b_{l,j}\int_{0}^{t}e^{-\lambda_k\tau}q(\tau)d\tau$$

(13b)

It can easily be found that $\Psi_{i,k0}$ is time-invarying, and it is proportional to the right eigenvectors $u_{i,k}$ with a constant $z_k(0)$, whereas $\Delta\Psi_{i,k}(\kappa)$ is time-varying due to the change of the input vector q.

Thus, considering (5), (6), and (13a), the mode shape information of kth oscillation mode between the signals i and j can be estimated by

$$\text{Magnitude} = \left|\frac{u_{i,k}}{u_{j,k}}\right| = \left|\frac{\Psi_{i,k0}}{\Psi_{j,k0}}\right|$$

(14a)

$$\text{Angle} = \angle u_{i,k} - \angle u_{j,k} = \angle\Psi_{i,k0} - \angle\Psi_{j,k0}$$

(14b)

Assuming for the moment that all signals are time-synchronized samples, and a reference state or signal is chosen as the state having the high observability in the kth oscillation mode, the theoretical feasibility of ARMA-P method estimating mode shape properties of multiple modes based on ambient signals is certified.

Model order selection

In the ARMA modeling of ambient signal, the model order selection is an important step. The applicability of model order will influence the accuracy and efficiency of oscillation modal frequency and damping ratio analysis, and further affect the mode shapes identification. In addition, the mode shape characteristics are relevant to multiple nodes in power grids, that is to say, we have to build the ARMA models of multiple signals. Obviously, it is time-consuming. And for the online application of ARMA-P method, it is best to automatically select the model order. Therefore, the model order selection is studied in this section. Different model order selection criteria are comparatively discussed, and the modeling procedure is considered to improve the calculation efficiency.

Model order selection criteria

Model order is a key factor in the ARMA model of ambient signal. A model with too high an order will include too much irrelevant oscillation information, and a model with too low an order may not include enough essential information about the system. Only, the ARMA model, with an optimal model order, can precisely describe the dynamic characteristics of power grids.

Information criteria (IC), which are generally used to determine the model order, are referred to as penalized log-likelihood criteria where the penalized term depends on the number of free parameters in the model and/or the number of observations [15]. They can be written in a generalized form

$$\text{IC}(p) = -2\sum_{i=1}^{N}\log f\left(x_i\middle|\hat{\theta}_p\right) + pC(N) \quad (15)$$

where $f(x_i|\theta_k)$, $i = 1,\ldots,N$ describes the conditional probability density of the observations $x_1,\ldots,x_N$, $C(N)$ is an increasing function of the observations number N, $\hat{\theta}_p$ is the estimator for the unknown model parameter based on the observations, and the optimal order choice is such that $\hat{p} = \text{argmin IC}(p)$.

Obviously that the second term grows as the parameters becomes complex, while the first term has the opposite variation, so the minimization of IC realizes a compromise between the data fitting and the complexity of the chosen parameter.

For the model order selection, the most known criterion is surely the Akaike's information criterion (AIC) [16], written as follows

$$\text{AIC}(p) = -2\sum_{i=1}^{N}\log f\left(x_i\middle|\hat{\theta}_p\right) + 2p \quad (16)$$

Theoretical research in the AIC has helped to specify the asymptotic behavior of AIC. This criterion is unsatisfactory

since it asymptotically leads to a strictly positive over parameterization probability of the model order [17].

In order to overcome the inconsistency of AIC, Schwarz [18] suggests the widely known Bayesian information criterion (BIC) based on the Bayesian justification

$$\text{BIC}(p) = -2 \sum_{i=1}^{N} \log f\left(x_i \middle| \hat{\theta}_p\right) + p \log N \quad (17)$$

And a different approach was introduced by Rissanen [19]. This approach suggested using the minimization of the length of a code. This code is required to encode observations. This criterion is referred to as minimum description length (MDL), defines as

$$\begin{aligned}\text{MDL}(p) = &-2 \sum_{i=1}^{N} \log f\left(x_i \middle| \hat{\theta}_p\right) + p \log N \\ &+ (p+2) \log(p+2)\end{aligned} \quad (18)$$

Theoretical research has found that BIC and MDL criteria are almost surely convergent in that they help in finding the appropriate model order when the observations number $N \rightarrow \infty$ (strong consistency) and penalized the term of likelihood more than AIC [15]. Because of the similarity of the two criteria, we choose to discuss the performance of BIC in this article.

Another strongly consistent criterion, referred to as φ_β, was introduced by El and Hallin [20]. It is a generalization of Rissanen's work on stochastic complexity [21], written as follows

$$\phi_\beta(p) = -2 \sum_{i=1}^{N} \log f\left(x_i \middle| \hat{\theta}_p\right) + p N^\beta \log \log N \quad (19)$$

with the refined conditions shown in (20) [18], which allows for β to be adjusted according to the number of observations N.

$$0 < \frac{\log \log N}{\log N} \leq \beta \leq 1 - \frac{\log \log N}{\log N} < 1 \quad (20)$$

In this article, we proposed to apply AIC, BIC, and ϕ_β, to the estimation of the order of ARMA models shown in (21).

ARMA model is a representative with the assumption that the input is approximately white over the frequency band of interest [22], and it is applicable to describe the characteristics of ambient signals in power grids.

$$\begin{aligned}x_i(\kappa) = &\varsigma_1 x_i(\kappa-1) + \cdots + \varsigma_n x_i(\kappa-n) \\ &- \phi_1 a(\kappa-1) - \cdots - \phi_m a(\kappa-m) \\ &+ a(\kappa)\end{aligned} \quad (21)$$

where $a(\kappa)$ is the stochastic disturbance input, ς_h $(h = 1 \dots n)$ and ϕ_g $(g = 1 \dots m)$ are the coefficients of AR and MA parts, N is the observations number, $\kappa = 1 \dots N$.

The order p in the criteria above is defined as $p = n + m$. Omitting terms that do not depend on the model order p, it is well known that the first terms in these formulae become $N \log \hat{\sigma}_a^2$, where $\hat{\sigma}_a$ is the variance estimate of disturbance input a shown in (21). Thus, we obtain

$$\text{IC}(p) = N \log \hat{\sigma}_a^2 + p C(N) \quad (22a)$$

$$\text{AIC}(p) = N \log \hat{\sigma}_a^2 + 2p \quad (22b)$$

$$\text{BIC}(p) = N \log \hat{\sigma}_a^2 + 2 \log N \quad (22c)$$

$$\phi_\beta(p) = N \log \hat{\sigma}_a^2 + k N^\beta \log \log N \quad (22d)$$

The selected model order verifies $\hat{p} = \arg\min \text{IC}(p)$.

ARMA(2*n*,2*n* – 1) modeling procedure

One immediate disadvantage of using the model order selection criterion is that since it is aimed at finding the optimal order among optional items. That is to say, a great many of ARMA models with different orders have to be built first. Obviously, it is time consuming and not good for the online application of the new ARMA-P method. In order to improve the calculation efficiency, the modeling procedure that specifies the search path of the optimal model order is considered in this article.

The ARMA(2*n*,2*n* – 1) modeling procedure proposed by Wu and Pandit [23] is employed. In this approach, first the ARMA(2*n*,2*n* – 1) model with the initial value $n = 1$ is modeled, then let $n = n + 1$. Only when the order is adequate as judged by the model order selection criterion, this step stops. Following that, the order of the AR and MA parts are reduced, respectively, and the model order selection criterion is applied until the optimal order is found. As shown in Figure 1, comparing with the traditional box modeling procedure, the computation efficiency is highly improved using the ARMA(2*n*,2*n* – 1) modeling procedure. It is better for the online application of the ARMA-P method in identifying the mode shape properties in interconnected power grids.

ARMA-P algorithms

Based on the theoretical deduction of the ARMA-P method in Section 2, the equations of this new approach for extracting the mode shape information from ambient signals are derived in this section.

Estimating the oscillation modal frequency and damping by ARMA model

With the assumption that the input is approximately white over the frequency band of interest, the ARMA model is proposed to estimate the oscillation mode characteristics from ambient data. As we know, small fluctuations associated with power system operation are the results of low-level stochastic disturbances inherent in

Figure 1 Modeling procedure. (a) ARMA($2n,2n - 1$) modeling procedure. **(b)** Box modeling procedure.
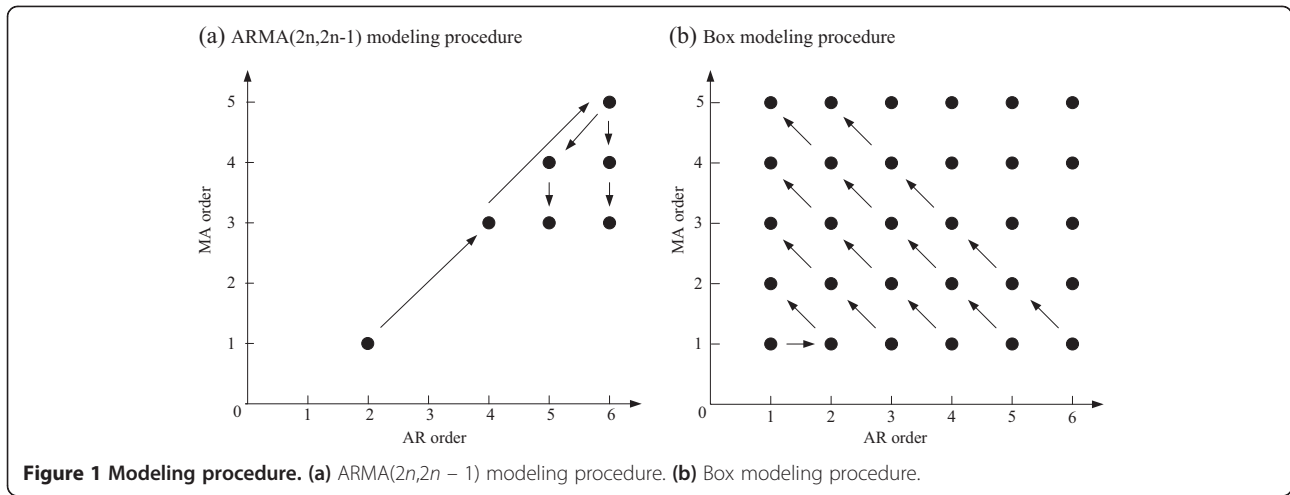
power grids. This kind of disturbance is assumed to relatively be statistically stationary for a block of data over the frequencies of interest [22,24].

The ARMA model of ambient signal x_i is shown in (21).

First, the optimal model order is selected using IC and the ARMA($2n,2n - 1$) modeling procedure. Then when $k = m + 1, m + 2 . . . m + M$ ($M > n$), a matrix equation is formed as follows

$$\begin{bmatrix} R_{m+1} \\ R_{m+2} \\ \vdots \\ R_{m+M} \end{bmatrix} = \begin{bmatrix} R_m & R_{m-1} & \cdots & R_{m-n+1} \\ R_{m+1} & R_m & \cdots & R_{m-n+2} \\ \vdots & \vdots & \vdots & \vdots \\ R_{m+M-1} & R_{m+M-2} & \cdots & R_{m+M-n} \end{bmatrix} \begin{bmatrix} \varsigma_1 \\ \varsigma_2 \\ \vdots \\ \varsigma_n \end{bmatrix} \quad (23)$$

where R_k is the autocorrelation function of signal x_i.

$$R_k = \frac{1}{N} \sum_{l=k+1}^{N} x_i(l) x_i(l - k)$$

Equation (23) is called the Modified Yule-Walker equation. The solution of (23) is the estimated coefficient vector of AR part.

A new time series $y_i(\kappa)$ is defined based on the information of the observations $x_1, . . . x_N$ and the AR part coefficient estimation $\hat{\varsigma}_1, . . . \hat{\varsigma}_n$.

$$y_i(\kappa) = x_i(\kappa) - \hat{\varsigma}_1 x_i(\kappa - 1) - \cdots - \hat{\varsigma}_n x_i(\kappa - n) \quad (24)$$

And

$$y_i(\kappa) = a(\kappa) - \phi_1 a(\kappa - 1) - \cdots - \phi_m a(\kappa - m) \quad (25)$$

And the spectral density function of y_i is calculated

$$S_{y_i y_i}(\omega) = \sigma_a^2 |\phi(B)|^2_{B=e^{-i\omega T}} = \sigma_a^2 |\prod_{j=1}^{m} \left(1 - \eta_{i,j} B\right)|^2_{B=e^{-i\omega T}} \quad (26)$$

where $\phi(B)$ is the polynomial of MA part, B is the backward operator, and T is the sample time.

Obviously when $B = 1/\eta_{i,j}$, (26) equals to zero.

On the other hand, based on the definition of spectral density, the spectral density function of the signal y_i is obtained

$$S_{y_i y_i}(\omega) = \sum_{k=0}^{m} R_{yi,k} B^k|_{B=e^{-i\omega T}} \quad (27)$$

where $R_{yi,k}$ is the autocorrelation function of y_i.

Deduced from (26) and (27), we obtained

$$\sum_{k=0}^{m} R_{yi,k} \left(\frac{1}{\eta_{i,j}}\right)^k = 0 \quad (28)$$

Then $\eta_{i,j}$ are calculated from (28), and substituted into the MA polynomial.

$$\prod_{j=1}^{m} \left(1 - \eta_{i,j} B\right) = 1 - \sum_{j=1}^{m} \phi_j B^j \quad (29)$$

By comparing the homogenous exponential coefficients of operator in (29), the coefficients of MA part are obtained. Thus, the ARMA model of ambient signal is built up.

The conjugate eigenvalues λ_k, λ_k^* can be calculated by solving the AR polynomial. And the low-frequency oscillation modal frequency f_k and damping ratio ξ_k are calculated

$$\begin{cases} f_k = \dfrac{\sqrt{\ln\lambda_k \ln\lambda_k^*}}{2\pi T} \cdot \sqrt{1 - \xi_k^2} \\ \xi_k = -\dfrac{\ln|\lambda_k|}{\sqrt{\ln\lambda_k \ln\lambda_k^*}} \end{cases} \quad (30)$$

where $k = 1 . . . n_d$, n_d is the number of dominant oscillation modes.

Estimating the mode shape magnitude and angle by Prony model

The approach of Prony model to estimate the electromechanical properties of power grids can be broken down into two parts [2,3]: first, calculating the eigenvalues of discrete model for estimating the modal frequency and damping ratio; second, computing the weights or coefficients to further extract the mode shapes properties.

In this part, the dominant oscillation modes, which play important roles in describing the dynamic characteristics of power system, are primarily studied. In general, the energies of the dominant modes take up large proportions in ambient signals. So in this case, the rest modes can be omitted in this article. Moreover, since the eigenvalues corresponding to one oscillation mode are conjugate, in the following part only the complex pair eigenvalues are considered. Consequently, based on the dominant modal information shown in (30), the eigenvalues of discrete model in Prony algorithm are calculated as

$$\gamma_k, \gamma_k^* = \exp\left[\left(-2\pi f_k \frac{\xi_k}{\sqrt{1-\xi_k^2}} \pm j2\pi f_k\right)T\right] \quad (31)$$

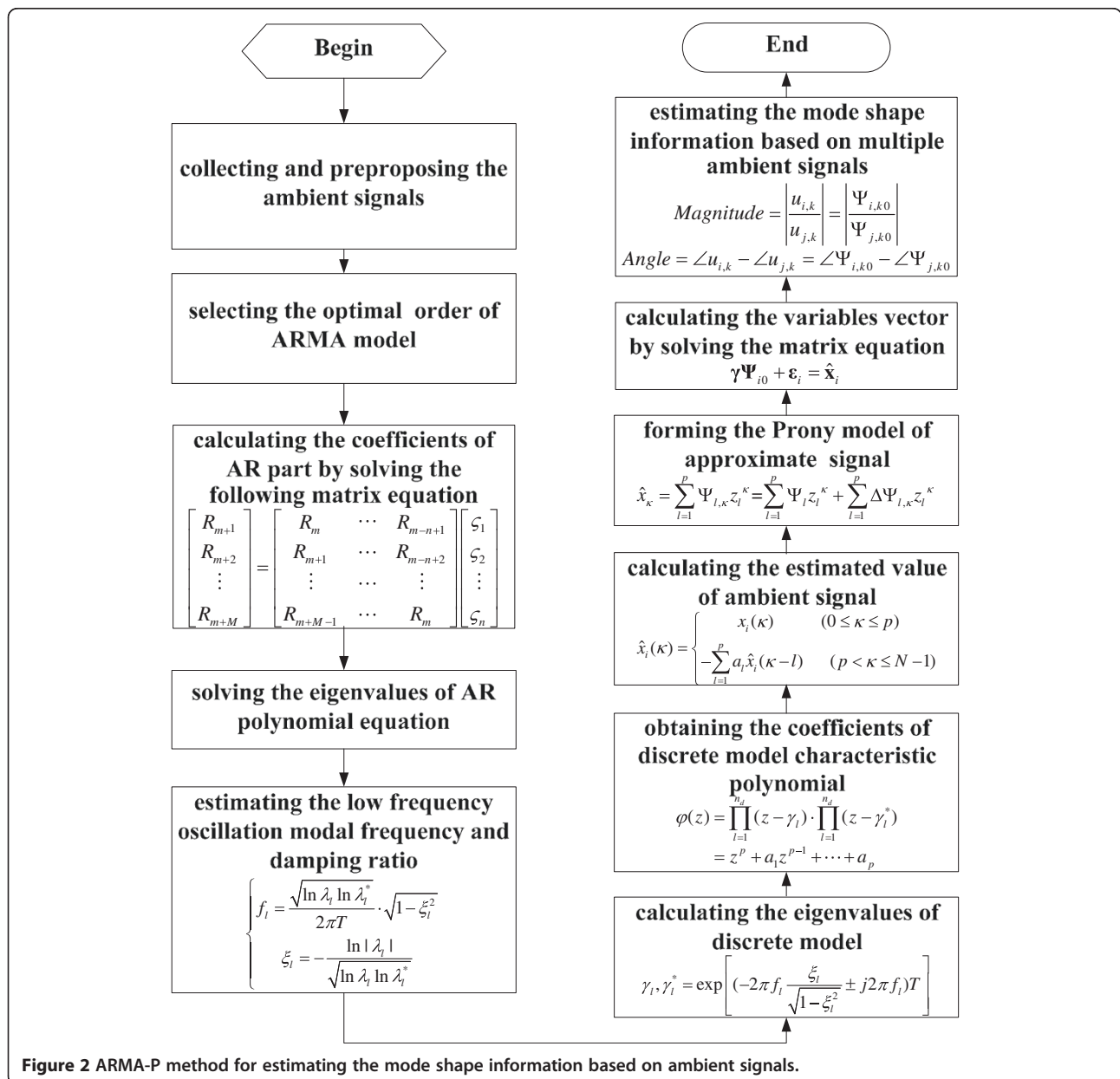where $k = 1\ldots n_d$, n_d is the number of dominant oscillation modes.

Figure 2 ARMA-P method for estimating the mode shape information based on ambient signals.

And the characteristic polynomial of discrete model is built

$$\varphi(z) = \prod_{l=1}^{n_d} (z - \gamma_l) \cdot \prod_{l=1}^{n_d} (z - \gamma_l^*) = z^p + a_1 z^{p-1} + \ldots + a_p$$

(32)

Then the coefficients a_l ($l = 1 \ldots p$, $p = 2 \times n_d$) are obtained from (32).

Following that, the approximate signal $\hat{x}$ is calculated

$$\hat{x}_i(\kappa) = \begin{cases} x_i(\kappa) & (0 \leq \kappa \leq p) \\ -\sum_{l=1}^{p} a_l \hat{x}_i(\kappa - l) & (p < \kappa \leq N - 1) \end{cases}$$

(33)

In Prony algorithm, the signal $\hat{x}_i$ is used to fit the true signal x_i. Following its principle shown in (9) and (12), the approximate signal $\hat{x}_i$ can be described as

$$\hat{x}_i(\kappa) = \sum_{l=1}^{p} \Psi_{i,l} \gamma_l^{\kappa} = \sum_{l=1}^{p} \Psi_{i,l0} \gamma_l^{\kappa} + \sum_{l=1}^{p} \Delta \Psi_{i,l}(\kappa) \gamma_l^{\kappa}$$

(34)

Defining an item ε_i as follows

$$\varepsilon_i(\kappa) = \sum_{l=1}^{p} \Delta \Psi_{i,l}(\kappa) \gamma_l^{\kappa}$$

(35)

Considering (13b), the item ε_i is a weighted sum of disturbance inputs. In this article, the inputs are assumed to relatively be statistically stationary over the frequencies of interest, so the item ε_i can also be termed as relatively statistically stationary.

Then a matrix equation is created.

$$\begin{bmatrix} 1 & \cdots & 1 \\ \gamma_1^1 & \cdots & \gamma_p^1 \\ \vdots & \cdots & \vdots \\ \gamma_1^{N-1} & \cdots & \gamma_p^{N-1} \end{bmatrix} \begin{bmatrix} \Psi_{i,10} \\ \Psi_{i,20} \\ \vdots \\ \Psi_{i,p0} \end{bmatrix} + \begin{bmatrix} \varepsilon_i(0) \\ \varepsilon_i(1) \\ \vdots \\ \varepsilon_i(N-1) \end{bmatrix} = \begin{bmatrix} \hat{x}_i(0) \\ \hat{x}_i(1) \\ \vdots \\ \hat{x}_i(N-1) \end{bmatrix}$$

(36)

It is rewritten in short as

$$\boldsymbol{\gamma} \boldsymbol{\Psi}_{i0} + \boldsymbol{\varepsilon}_i = \hat{\mathbf{x}}_i$$

A least square solution of (36) yields the variables vector as shown in (37).

$$\hat{\boldsymbol{\Psi}}_{i0} = \left(\boldsymbol{\gamma}^T \boldsymbol{\gamma} \right)^{-1} \boldsymbol{\gamma}^T \hat{\mathbf{x}}_i$$
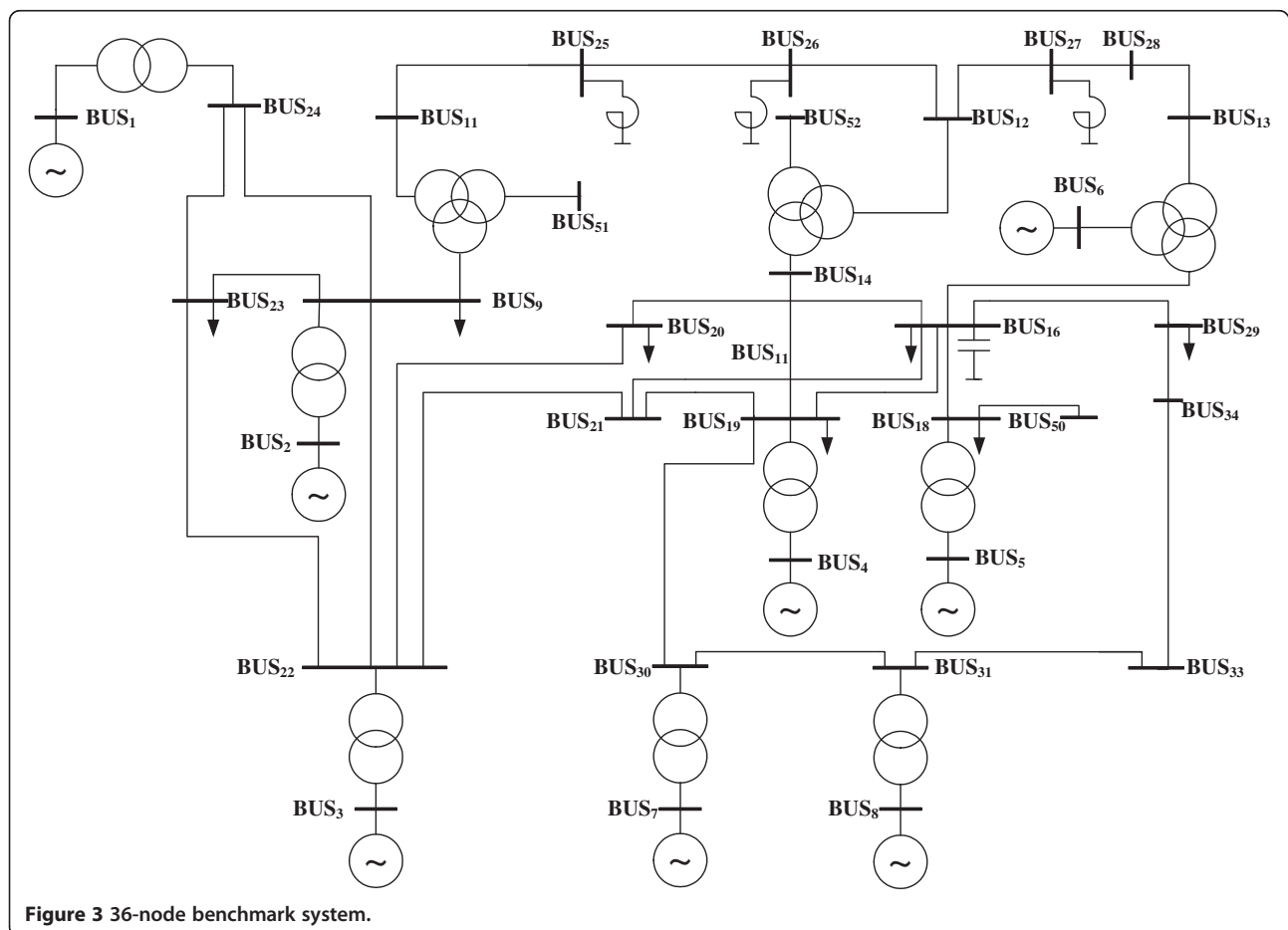
(37)

Figure 3 36-node benchmark system.

Thus, the mode shape magnitude and angle information of multiple dominant interarea modes in power grids can be estimated from the variables $\hat{\Psi}_{i0}$ according to (14).

Therefore, based on ambient signals synchronously collected from power grids, following these steps shown in Figure 2, the system's mode shape properties are identified.

Simulation examples

A 36-node benchmark system shown in Figure 3 is used to demonstrate the performance of the ARMA-P method, using the Power System Analysis Software Package as the simulation tool. The system contains eight generators at major generation buses 1 through 8. The generators are represented using detailed two-axis transient models. Each load in this system is split into a portion consisting of constant power and random power. The random portion of both the real and reactive loads is obtained by passing independent Gaussian white noise through low pass filters.

The dominant interarea low-frequency oscillation modes shown in Table 1 are calculated by conducting an eigenalaysis of the entire system's small-signal model under nominal steady-state operating conditions.

In this article, Mode I (0.778 Hz, 1.123%) is mainly discussed. The mode shape magnitude and angle information of this mode is given in Table 2. Taking Gen8 as the reference, the eight generators in 36-node benchmark systems can be classified into two groups: Group A (including Gen1, Gen2), Group B (including Gen3, Gen4, Gen5, Gen6, Gen7, and Gen8). Among them, Gen1, Gen7, and Gen8 have high degrees of participation in Mode I.

For the examples that follow, a typical time-domain simulation is comprised of driving the system with random load variations. The system's responses consist of small random variations in the system states. As an example, Figure 4a describes the resulting random variations of Gen8 frequency signal for a 10-min simulation. The spectrum of this signal shown in Figure 4b illustrates that the ambient signal includes rich information about the dominant oscillation modes in 36-node benchmark system, especially Mode I.
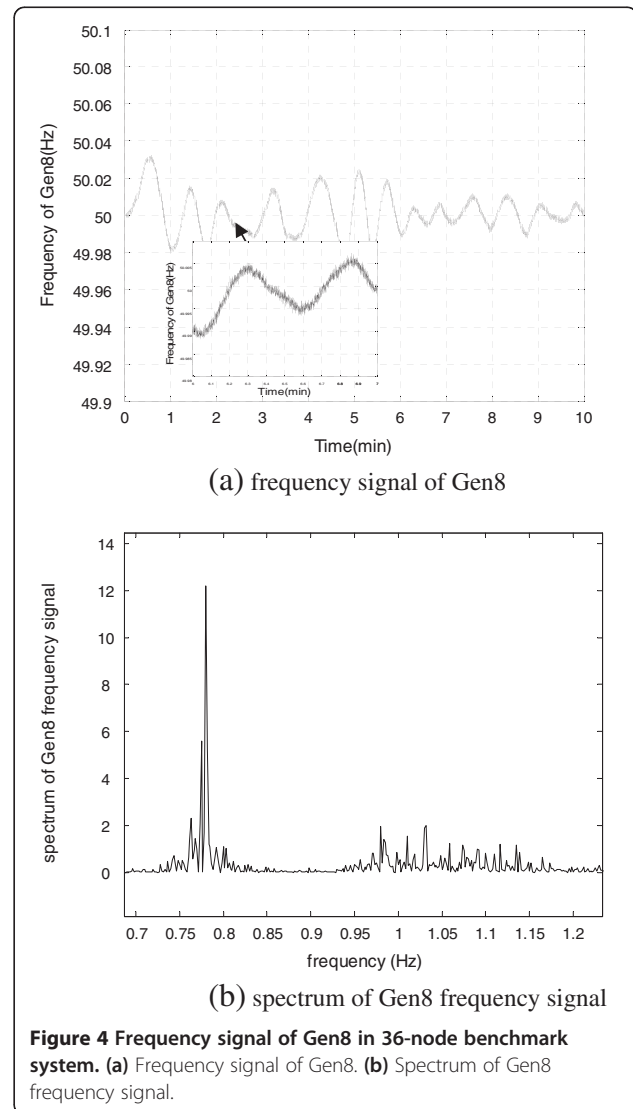
Table 1 Dominant low-frequency oscillation modes of 36-node benchmark system

Mode	Eigenvalue	Frequency (Hz)	Damping ratio (%)
I	$-0.0549 + j4.885$	0.778	1.123
II	$-0.270 + j6.152$	0.980	4.348

Table 2 Mode shape information of mode I in 36-node benchmark system

Gen	Magnitude (p.u.)	Angle (rad)
1	0.683	3.156
2	0.120	3.076
3	0.474	6.096
4	0.425	−0.050
5	0.597	6.140
6	0.426	−0.067
7	0.983	0.016

ARMA-P method is applied to estimate the mode shape properties from the frequency signals of the eight generators in the simulation system. First, these ambient signals are preprocessed. Considering the frequency range of electromechanical mode (generally

(a) frequency signal of Gen8

(b) spectrum of Gen8 frequency signal

Figure 4 Frequency signal of Gen8 in 36-node benchmark system. (a) Frequency signal of Gen8. **(b)** Spectrum of Gen8 frequency signal.

[0.1 Hz, 2.5 Hz]) in power systems, the signals are low pass filtered with a cutoff frequency of 3 Hz, then decimated from 50 samples per second to 5 samples per second, and finally high pass filtered to remove any low-frequency trends.

Then according to the procedure of ARMA-P method shown in Figure 2, the model order must be determined first. In order to fully compare the performances of the three typical criteria in power grids, AIC, BIC, and ϕ_β are, respectively, applied to select the optimal order of ARMA model based on the simulative ambient data. And the ARMA $(2n, 2n - 1)$ modeling procedure is employed to specify the search path.

In this part, the frequency signal of Gen8, which includes rich information about the dynamic characteristic of system, is chosen to be the analysis object. The number of observations is $N = 3000$. According to (20), the bounds on the value β in ϕ_β criterion are calculated,

Table 3 Oscillation modes results of three typical model order selection criteria

Criterion	Model order	Mode I		Mode II	
		Frequency (Hz)	Damping ratio (%)	Frequency (Hz)	Damping ratio (%)
AIC	(16,15)	0.774	1.702	0.959	4.905
BIC	(12,11)	0.776	1.270	0.963	4.549
$\varphi_{0.3}$	(11,9)	0.774	1.553	1.016	5.699
$\varphi_{0.4}$	(8,7)	0.773	1.541	1.089	8.858
$\varphi_{0.5}$	(7,5)	0.776	1.819	1.127	10.749

$\beta_{min} = 0.2598$, $\beta_{max} = 0.7402$. So in this case, the ϕ_β criterion has been calculated for values of β equal to 0.3, 0.4, and 0.5.

The research on the model order selection is based on 50 experiments. The eigenvalues results of different model order selection criteria are plotted in Figure 5, and the average value of the order for the 50 experiments is given in Table 3. The oscillation mode properties in each case are also calculated.

We can notice that AIC leads to over-parameterization while ϕ_β ($\beta = 0.4, 0.5$) under-parameterize because of the high penalty. The order results of BIC and ϕ_β ($\beta = 0.3$) are similar. Considering the relatively limited system dynamic information that include in ambient data, a small change in model order would lead to a large variation in estimated modal parameters and influence the identification accuracy. So, the oscillation mode characteristic results corresponding to BIC and ϕ_β ($\beta = 0.3$) are further compared, in order to discuss the applicability of these criteria in power systems. Comparing with the eigenanalysis results shown in Table 1, the accuracy of the modal information corresponding to BIC are much better. Considering Occam's Razor [25], BIC is appropriate to select the optimal order of ARMA model, and its feasibility is testified.
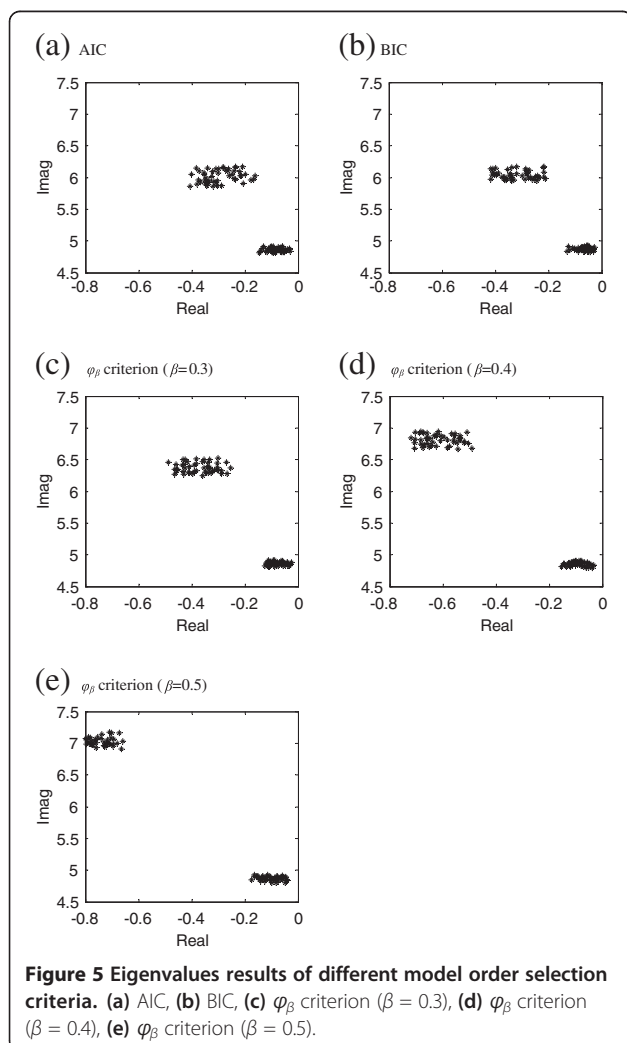
Figure 5 Eigenvalues results of different model order selection criteria. (a) AIC, **(b)** BIC, **(c)** φ_β criterion ($\beta = 0.3$), **(d)** φ_β criterion ($\beta = 0.4$), **(e)** φ_β criterion ($\beta = 0.5$).

Table 4 Analysis results of Mode I based on frequency signals of eight generators

Gen	Frequency (Hz)	Damping ratio (%)
1	0.778	0.995
2	0.774	2.014
3	0.772	1.331
4	0.771	1.566
5	0.775	1.136
6	0.779	1.400
7	0.775	1.341
8	0.776	1.270

Table 5 Results of variables vector $\hat{\Psi}_{i0}$ of Mode I

Gen	$\hat{\psi}_{i0}$
1	−0.611 + j0.206
2	−0.099 + j0.041
3	0.518 − j0.011
4	0.439 − j0.021
5	0.620 − j0.108
6	0.461 − j0.031
7	0.993 + j0.015
8	1.000 + j0.000

Then BIC and the ARMA($2n,2n − 1$) modeling procedure is applied to process the frequency signals of the eight generators in 36-node benchmark system, and the estimated modal information of Mode I are listed in Table 4. Obviously, the analysis results are basically close to the eigenanalysis results.

Based on the dominant oscillation mode information above, the eigenvalues of discrete model are calculated. And following the procedure of ARMA-P method shown in Figure 2, the variables vector $\hat{\Psi}_{i0}$ are solved. Using Gen8 as the reference, the analysis results of Mode I are shown in Table 5.

And the mode shape magnitude and angle are extracted by (14), listed in Table 6.

The ARMA-P method results closely approximate the eigenanalysis results shown in Table 2. The relative errors are mostly no more than 10%, the accuracy of mode shape angle is somewhat better than that of magnitude, and the effectiveness of model order selection is verified again. The mode shape properties can be estimated accurately using the ARMA-P method based on ambient data.

In actual operating condition, noise disturbances such as measurement errors exist continuously in power systems. Since the fluctuation amplitude of ambient signal is relatively small, this type of disturbance cannot be ignored. In order to testify the feasibility of ARMA-P

Table 7 Mode shape results of Mode I from ambient signals with different SNR values

SNR (dB)	Gen	Magnitude (p.u.)	Error (%)	Angle (rad)	Error (%)
20	1	0.647	5.324	3.229	2.318
	2	0.135	12.510	3.369	9.523
	3	0.530	11.648	5.428	10.970
	4	0.458	7.878	−0.044	10.890
	5	0.588	1.422	6.282	2.319
	6	0.375	11.942	−0.073	8.444
	7	0.974	0.851	0.014	10.798
12	1	0.601	12.019	3.169	0.431
	2	0.131	8.971	2.783	9.538
	3	0.407	14.211	5.468	10.312
	4	0.372	12.426	−0.054	9.959
	5	0.517	13.430	6.352	3.458
	6	0.387	9.045	−0.060	10.424
	7	0.861	12.383	0.014	10.364

method in this case, colored noises with certain amplitude are added to the original ambient signals. And the identified mode shape results of Mode I from ambient data with 20-dB signal-to-noise ratio (SNR) and 12-dB SNR are indicated in Table 7.

It can be seen from the table that the relative errors of angle results are all less than 11%, which means that the mode shape angle can be identified accurately from ambient data; the relative errors of magnitude results are a little bigger, yet still less than 15%, which also meets the accuracy requirements of engineering.

Moreover, to test the statistical accuracy of the ARMA-P method, 100 Monte Carlo simulations are run, the mean and root mean square error (RMSE) are calculated. Monte Carlo trials work as to take several independent measurements on the power system in order to get a sense of the method's statistical performance.

Table 6 Mode shape results of Mode I in 36-node benchmark system

Gen	Magnitude (p.u.)	Error (%)	Angle (rad)	Error (%)
1	0.645	5.564	2.817	10.741
2	0.108	10.000	2.751	10.566
3	0.518	9.283	6.263	2.740
4	0.439	3.294	−0.048	4.000
5	0.629	5.360	6.111	0.472
6	0.462	8.451	−0.066	1.493
7	0.993	1.017	0.015	6.250

Table 8 Mode shape with mean and RMSE value of Mode I based on ambient signals

Gen	Magnitude (p.u.)		Angle (rad)	
	Mean	RMSE	Mean	RMSE
1	0.717	0.056	3.390	0.051
2	0.109	0.009	3.355	0.055
3	0.517	0.032	5.930	0.070
4	0.440	0.024	−0.048	0.032
5	0.566	0.040	6.177	0.062
6	0.457	0.033	−0.069	0.029
7	0.971	0.010	0.0170	0.031

Figure 6 Structure diagram of China Southern Power Grid.

The results of the Monte Carlo trials are listed in Table 8. Obviously, there is a good agreement between the new ARMA-P method and the traditional eigenanalysis. The feasibility of AMRA-P method is further verified.

Actual system example

Now consider the ambient data from China Southern Power Grid as shown in Figure 6. The system contains several interarea modes including two dominant modes, one is the Yunnan-Guizhou mode, about 0.6–0.7 Hz, and another is the Yunnan&Guizhou-Guangdong mode, about 0.4–0.5 Hz. The angle relationship of generators in the Yunnan-Guizhou mode is shown in Figure 7, where the ones on the left are the generators in Yunnan Province, and the ones on the right are the generators in Guizhou Province.

For demonstration purposes, five locations spread across the power grid are selected, including Anshun substation, Gaopo converting plant, and Xingren

converting plant in Guizhou province, and Luoping substation in Yunnan Province, Luodong substation in Guangdong Province.

Ten-minute frequency data were collected from 02:10 to 02:20, on June 14th, 2009. The frequency signal of Luoping substation is shown in Figure 8. Owing to the limitation of phasor measurement, the accuracy of frequency signal shown is limited, and since the power flow was regulated during the data-collecting process, this signal has apparent fluctuations. However, they would not influence the identification of oscillation mode information.

The ARMA-P method is used to identify the mode shape properties from ambient signals. Luoping substation was selected as the reference. The analysis results in Table 9 and Figure 9 show the mode shape angle information in China Southern Power Grid.

From Table 9, it can be seen that the system includes two dominant interarea oscillation modes, with the
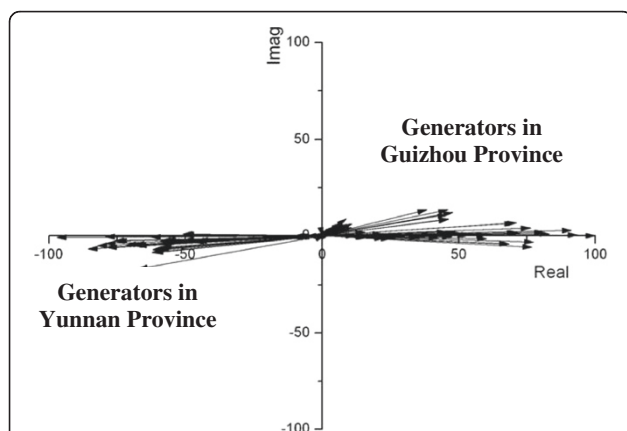
Figure 7 Angle relationship of generators in Yunnan-Guizhou mode.
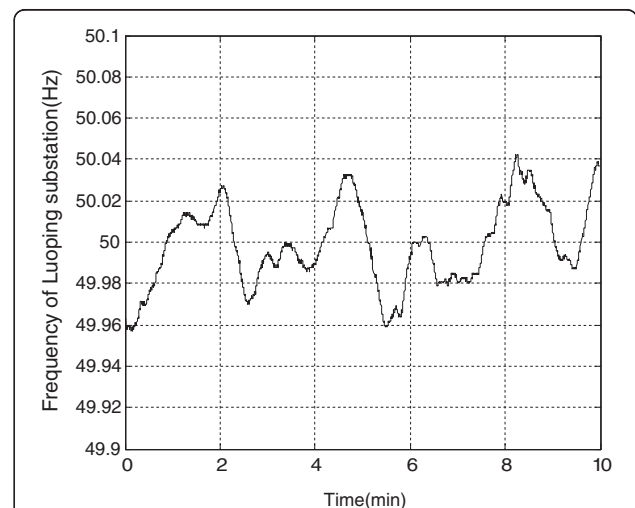
Figure 8 Measured frequency signal of Luoping substation.

Table 9 Results of mode shape angle in China Southern Power Grid

Location	Mode	Frequency (Hz)	Angle (rad)
Anshun	Mode I	0.677	3.621
	Mode II	0.484	0.122
Gaopo	Mode I	0.685	3.370
	Mode II	0.479	−0.310
Xingren	Mode I	0.693	3.102
	Mode II	0.478	−0.176
Luodong	Mode I	\	\
	Mode II	0.490	3.555
Luoiping	Mode I	0.659	0.000
	Mode II	0.471	0.000

frequency of mode I at about 0.47 Hz, and the frequency of mode II at about 0.68 Hz. Anshun, Gaopo, Xingren, Luoping all participate in the two modes, whereas Luodong only participates in mode I. The plots in Figure 9 indicate that in mode I Anshun, Gaopo, and Xingren swing together against Luoping, which conforms to the angle relationship shown in Figure 7, and in mode II Anshun, Gaopo, Xingren, and Luoping swing together against Luodong. It can be deduced that mode I is the Yunnan-Guizhou mode, and mode II is the Yunnan&Guizhou-Guangdong mode. The analysis results conform to the mode shape characteristics information that analyzed in advance. The ARMA-P method performs well in estimating the mode shape of multiple modes simultaneously based on actual ambient signals in China Southern Power Grid.

Conclusion

A methodology considering the model order selection called ARMA-P, used for estimating the mode shape properties from time-synchronized phasor measurements, is presented. Based on its theoretical analysis basis, the approach is applied to a simulation system and measured data from China Southern Power Grid. The results demonstrate that the optimal model order can be selected automatically and efficiently using BIC and the ARMA($2n$,$2n - 1$) modeling procedure. This method works well in estimating the mode shape information of multiple oscillation modes simultaneously based on ambient signals with different SNR value. And further based on Monte Carlo studies, it is shown that the ARMA-P method can estimate mode shapes with reasonably good accuracy.

The algorithm proposed in this article shows great promise for estimating the mode shape properties of power systems. Future work will more rigorously investigate its performance including the data length required and the calculation speed.
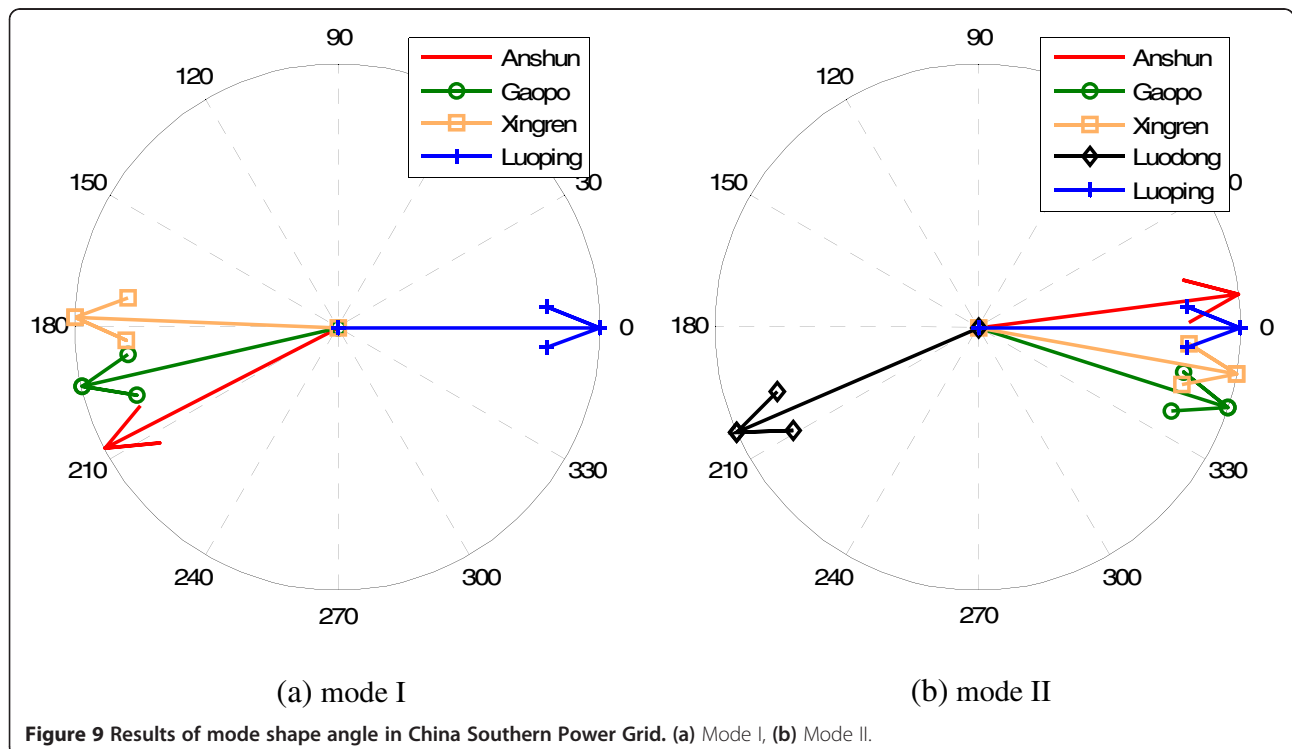
Figure 9 Results of mode shape angle in China Southern Power Grid. (a) Mode I, **(b)** Mode II.

Author details
[1]College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518060, China. [2]Department of Electrical Engineering, Tsinghua University, Beijing 100084, China.

Reference
1. P. Kunder, *Power System Stability and Control* (McGrawHill, Inc., New York, 1994)
2. J.F. Hauer, C.J. Demeure, L.L. Scharf, Initial results in Prony analysis of power response signals. IEEE Trans. Power Syst. **5**(1), 80–89 (1990)
3. D.J. Trudnowski, J.W. Pierre, *Signal processing methods for estimating small-signal dynamic properties from measured responses, in Analysis of Nonlinear and Non-stationary Inter-area Oscillations: A Time Frequency Perspective* (Springer, New York, 2009)
4. I. Kamwa, G. Trudel, L. Gerin-Lajoie, *Low-order black-box models for control system design in large power systems, in IEEE Proceeding of Power Industry Computer Application Conference* (Salt Lake City, 1995), pp. 190–198
5. J.J. Sanchez-Gasca, J.H. Chow, Performance comparison of three identification methods for the analysis of electromechanical oscillations. IEEE Trans. Power Syst. **14**(3), 995–1002 (1999)
6. J.W. Pierre, D.J. Trudnowski, M.K. Donnelly, Initial results in electromechanical mode identification from ambient data. IEEE Trans. Power Syst. **12**(3), 1245–1251 (1997)
7. IEEE Task Force Report, *Identification of electromechanical modes in power systems* (The Institute of Electrical and Electronics Engineers, Inc, 2012)
8. L. Vanfretti, L. Dosiek, J.W. Pierre, D.J. Trudnowski, J.H. Chow, R. García-Valle, U. Aliyu, Application of ambient analysis techniques for the estimation of electromechanical oscillations from measured PMU data in four different power systems. Eur. Trans. Electr. Power **21**(4), 1640–1656 (2011)
9. C. Wu, C. Lu, Y.D. Han, Closed-loop identification of power system based on ambient data. Math. Probl. Eng. (2012). Article ID 632897, 16 (2012)
10. D.J. Trudnowski, Estimating electromechanical mode shape from synchrophasor measurements. IEEE Trans. Power Syst. **23**(3), 1188–1195 (2008)
11. G. Liu, V. Venkatasubramanian, *Oscillation monitoring from ambient PMU measurements by frequency domain decomposition, in IEEE Proceeding of Symposium Circuits and Systems* (Seattle, 2008), pp. 2821–2824
12. L. Dosiek, D.J. Trudnowski, J.W. Pierre, *New algorithm for mode shape estimation using measured data, in IEEE Proceeding of PES General Meeting* (Pittsburgh, 2008), pp. 1–8
13. L. Dosiek, J.W. Pierre, D.J. Trudnowski, N. Zhou, *A channel matching approach for estimating mode shape, in IEEE Proceedings of PES General Meeting* (Calgary, 2009), pp. 1–8
14. N. Zhou, Z.Y. Huang, L. Dosiek, D.J. Trudnowski, J.W. Pierre, *Electromechanical mode shape estimation based on transfer function identification using PMU measurements, in IEEE Proceedings of PES General Meeting* (Calgary, 2009), pp. 1–7
15. S. Patrick, *Optimization in Signal and Image Processing* (Wiley, New Jersey, 2009)
16. H. Akaike, *Information theory and an extension of the maximum likelihood principle, in Proceeding of 2nd International Symposium on Information Theory* (Budapest, 1973), pp. 267–281
17. R. Shibata, Selection of the order of an AR model by AIC. Biometrika **63**, 117–126 (1976)
18. G. Schwarz, Estimating the dimension of a model. Ann. Stat. **6**, 461–464 (1978)
19. J. Rissanen, Modeling by shortest data description. Automatica **14**, 465–471 (1978)
20. M. El, M. Hallin, *Order selection, stochastic complexity and Kullback–Leibler information, in Athens Conference on Applied Probability and Time Series Analysis, vol. II (1995), volume 115 of Lecture Notes in Statist* (Springer, New York, 1996), pp. 291–299
21. J. Rissanen, *Stochastic Complexity in Statistical Inquiry* (World Scientific, New Jersey, 1989)
22. B.E.P. George, G.M. Jenkins, G.C. Reinsel, *Time Series-Forecasting and Controls* (Wiley, New Jersey, 2008)
23. S.M. Pandit, S.M. Wu, *Time-Series and System Analysis with Applications* (Krieger Publishing Company, FL, 2006)
24. H. James Douglas, *Time Series Analysis* (Princeton University Press, Princeton, 1994)
25. A. Blumer, A. Ehrenfeucht, D. Haussler, M. Warmuth, Occam's Razor. Inf. Process. Lett. **24**, 377–380 (1987)