Improved Subsynchronous Oscillation Parameter Identification with Synchrophasor Based on Matrix Pencil Method in Power Systems

Xiaoxue Zhang, Fang Zhang, *Senior Member*, *IEEE*, Wenzhong Gao, *Fellow*, *IEEE*, and Jinghan He, *Fellow*, *IEEE*

Abstract—The subsynchronous oscillations (SSOs) related to renewable generation seriously affect the stability and safety of the power systems. To realize the dynamic monitoring of SSOs by utilizing the high computational efficiency and noise-resilient features of the matrix pencil method (MPM), this paper proposes an improved MPM-based parameter identification with synchrophasors. The MPM is enhanced by the angular frequency fitting equations based on the characteristic polynomial coefficients of the matrix pencil to ensure the accuracy of the identified parameters, since the existing eigenvalue solution of the MPM ignores the angular frequency conjugation constraints of the two fundamental modes and two oscillation modes. Then, the identification and recovery of bad data are proposed by utilizing the difference in temporal continuity of the synchrophasors before and after noise reduction. The proposed parameter identification is verified with synthetic, simulated, and actual measured phase measurement unit (PMU) data. Compared with the existing MPM, the improved MPM achieves better accuracy for parameter identification of each component in SSOs, better real-time performance, and significantly reduces the effect of bad data.

Index Terms—Subsynchronous oscillations (SSOs), synchrophasor, parameter identification, matrix pencil method, bad data.

I. INTRODUCTION

W ITH the development of renewable generations in modern power systems, subsynchronous oscillations (SSOs) caused by the resonances between power electronic devices and series compensators or weak grids are of frequent occurrence [1] - [3]. Compared with the traditional SSOs caused by the resonances between electrical systems and turbine generators in which only one oscillation component exits [4], the SSOs caused by renewable generations

may contain one subsynchronous component together with another frequency-coupled supersynchronous component due to the power electronic converters [3], [5]. The oscillation features of these SSOs are more changeable and widely spread, which endanger the operation of the power system. Therefore, the parameters of each component in instantaneous data during oscillation should be accurately identified in order to monitor and mitigate SSOs effectively [6]-[8].

There has been much research on identifying SSOs using instantaneous data, which can accurately obtain SSO parameters [9]-[12]. However, since instantaneous data are only obtainable from fault recorders (FRs), which are primarily stored at different buses, it is challenging to collect FR data on time. Also, there are no unified timing standards for FR, which makes it difficult to observe SSOs dynamically and globally.

SSO parameter identification based on synchrophasors provided by phasor measurement units (PMUs) and wide area measurement systems (WAMSs) can realize dynamic and synchronized monitoring of SSOs by using the synchronous measurement mechanism and high reporting rate of synchrophasors [13], [14]. Among the existing SSO parameter identification methods with synchrophasors, the discrete fourier transform (DFT) based algorithms are widely studied. The feasibility of identifying the subsynchronous parameters by synchrophasors is explored in [15] for the first time, and the spectrum analysis is performed on the amplitude of synchrophasors by taking advantage of the linear transformation features of synchrophasor calculation. In [16], the interpolated DFT enhanced by the Hann window is adopted to identify SSOs with a 2-s data window. The limitation of the spectrum analysis is that the spectrum resolution ratio is greatly affected by the length of the data window. A long data window is required to ensure a sufficiently high spectrum resolution ratio. However, for the fast-changing SSOs, the longer the data window is, the worse the real-time performance and reliability of parameter identification are.

In addition, non-DFT methods based on synchrophasors have also been studied. Two classic modal parameter extraction (MPE) algorithms, i. e., Prony analysis [17], [18] and the matrix pencil method (MPM) [19], are adopted to identify the parameters of SSO with a 1-s data window in [20].

Based on historical data and the incremental learning of the multi-support vector basis model, the features of PMU data are classified and identified in [21], [22]. Compared with the spectrum analysis method, the non-DFT methods are less constrained by the spectrum resolution ratio, so the identification accuracy is less sensitive to the data window length, which can theoretically shorten the data window for SSO parameter identification. However, the shortening of the data window leads to fewer extractable features simultaneously, thus increasing the difficulty of parameter identification, and needs to be compensated by the computational volume.

As mentioned before, the supersynchronous component may exist during the SSO. Due to the coupling frequency of the subsynchronous and supersynchronous components, the positive and negative spectra of which are aliasing, the parameters obtained by the above methods are incorrect because the supersynchronous component is not considered. In the primary research of our team [23], the subsynchronous and supersynchronous parameters are extracted based on spectrum analysis, but the supersynchronous component has not yet been involved in SSO parameter identification based on MPE algorithms. In addition, for the SSOs caused by renewable generations, a 1-s data window still cannot reflect the changeable characteristics of SSO.

MPM is a classic MPE algorithm with advantages in computational efficiency and noise-resilient features [19]. The SSO parameter identification based on MPM is implemented in [20], in which the matrix pencil is constructed with synchrophasors, and solving the eigenvalues of the matrix pencil can extract the modal parameter. In solving the eigenvalue problem, the approximate calculation does not perform like the spectrum analysis method, so the calculation errors of MPM can theoretically be decreased to almost zero. However, the angular frequency conjugation constraints of the two fundamental modes and two oscillation modes are ignored when solving the eigenvalues directly, and the accuracy of the frequency obtained by the MPM is difficult to guarantee when the model error exists, which leads to the identified parameters suffering from low accuracy. Moreover, in practical applications of PMUs, in the event of measurement error of the PMU or data variation during communication, bad data may exist in synchrophasors uploaded by the PMU, which results in temporal discontinuity of synchrophasors and seriously affects the performance of MPM. Few existing studies on MPM have dealt with bad data.

To take advantage of the computational efficiency and noise-resilient features of MPM, the MPM is improved by solving issues in the matrix pencil solution and identification and recovery of bad data in this paper. The features of the proposed improved SSO parameter identification based on MPM are twofold.

1) By analyzing the principle of the MPM and the disadvantage of the matrix pencil solution, the angular frequency fitting equations corresponding to fundamental and oscillation frequencies are established based on the characteristic polynomial coefficients of the matrix pencil, in which the angular frequency conjugation constraints of the two fundamental modes and two oscillation modes are considered. Multi-order matrix pencil is proposed to further smooth model errors.

2) Identification and recovery methods of bad data are proposed. The significance of bad data is evaluated by utilizing the difference in temporal continuity of the synchrophasors before and after noise reduction. Based on the significance level of the bad data, whether to perform noise reduction or recovery of bad data can be determined, thus reducing the impact of bad data on the performance of the MPM.

Finally, the proposed improved MPM is verified with synthetic, simulated, and actual measured PMU data. Also, it is compared with the existing MPM in [20]. The comparison results demonstrate that the proposed improved MPM can realize better accuracy of the frequency, amplitude, and phase of the fundamental, subsynchronous, and supersynchronous components in SSOs with a 200-ms data window. Meanwhile, the consideration for the identification and recovery of bad data can effectively guarantee the performance of the MPM.

This paper is organized as follows. Section II studies the synchrophasors model and principle of MPM. Section III introduces SSO parameter identification based on improved MPM. Section IV proposes the identification and recovery of improved MPM of the bad data. The performance of the proposed method is evaluated using synthetic, simulated, and actual measured PMU data in Section V. The conclusions are illustrated in Section VI.

II. SYNCHROPHASORS MODEL AND PRINCIPLE OF MPM

A. SSO Model with Supersynchronous Component

The instantaneous data $x(t)$ during SSO consists of fundamental, subsynchronous, and supersynchronous components, i.e.,

$$x(t) = x_0\cos(2\pi f_0 t + \phi_0) + x_{sub}\cos(2\pi f_{sub} t + \phi_{sub}) + x_{sup}\cos(2\pi f_{sup} t + \phi_{sup}) \tag{1}$$

where (f_0, x_0, ϕ_0), $(f_{sub}, x_{sub}, \phi_{sub})$, and $(f_{sup}, x_{sup}, \phi_{sup})$ are the frequencies, amplitudes, and phases of the fundamental, subsynchronous, and supersynchronous components of SSO, respectively. The frequencies of subsynchronous and supersynchronous components satisfy $f_{sub} + f_{sup} = 2f_N$, and f_N is the nominal frequency.

Synchrophasors are obtained by applying the DFT synchrophasor algorithm [24] on $x(t)$, similar to the derivations in [20], [23]. Let $\dot{X}_0(k)$, $\dot{X}_{sub}(k)$, and $\dot{X}_{sup}(k)$ denote the spectrum of the fundamental, subsynchronous, and supersynchronous components of the kth synchrophasor, respectively. Each component is composed of a positive spectrum and a negative spectrum, as shown in (2)-(4).

$$\begin{cases} \dot{X}_0^+(k) = Q^*(f_0, 1)x_0 e^{-j\phi_0}e^{j\alpha k} \\ \dot{X}_0^-(k) = Q(f_0, -1)x_0 e^{j\phi_0}e^{-j\alpha k} \end{cases} \tag{2}$$

$$\begin{cases} \dot{X}_{sub}^+(k) = Q^*(f_{sub}, 1)x_{sub} e^{-j\phi_{sub}}e^{j\beta k} \\ \dot{X}_{sub}^-(k) = Q(f_{sub}, -1)x_{sub} e^{j\phi_{sub}}e^{-j\beta k} \end{cases} \tag{3}$$

$$\begin{cases} \dot{X}_{sup}^+(k) = Q(f_{sup}, -1)x_{sup} e^{j\phi_{sup}}e^{j\beta k} \\ \dot{X}_{sup}^-(k) = Q^*(f_{sup}, 1)x_{sup} e^{-j\phi_{sup}}e^{-j\beta k} \end{cases} \tag{4}$$

where the subscripts "+" and "−" represent the positive and negative spectra, respectively; the superscript "*" represents the conjugation; α and β are the angular frequencies of f_0 and f_{sub} as shown in (5), respectively; and $Q(f, l)$ is illustrated in (6).

$$\begin{cases} \alpha = 2\pi \dfrac{f_N - f_0}{f_S} \\ \beta = 2\pi \dfrac{f_N - f_{sub}}{f_S} \end{cases} \tag{5}$$

$$Q(f, l) = \frac{1}{N} \sum_{n=0}^{N-1} e^{\left(j \frac{2\pi f}{f_N N} + j \frac{2\pi}{N} l \right) n} \tag{6}$$

where N is the number of instantaneous data in the DFT algorithm; and f_S is the reporting rate of synchrophasor, and $f_S = 2 f_N$.

Due to the coupling frequency of the subsynchronous components and the supersynchronous components, the frequencies of positive and negative spectra of $\dot{X}_{sub}(k)$ and $\dot{X}_{sup}(k)$ are the same. Thus, (3) and (4) can be sumed to positive and negative spectra of oscillation component, respectively, as shown in (7).

$$\begin{cases} \dot{X}_S^+(k) = \dot{X}_{sub}^+(k) + \dot{X}_{sup}^+(k) \\ \dot{X}_S^-(k) = \dot{X}_{sub}^-(k) + \dot{X}_{sup}^-(k) \end{cases} \tag{7}$$

where $\dot{X}_S^+(k)$ and $\dot{X}_S^-(k)$ are the positive and negative spectra of oscillation components, respectively.

Consequently, the synchrophasor $\dot{X}(k)$ consists of four parts, as expressed in (8): the positive and negative spectra of fundamental components vary at angular frequencies α and $-\alpha$, and the positive and negative spectra of oscillation components vary at angular frequencies β and $-\beta$.

$$\dot{X}(k) = \dot{X}_0^+(k) + \dot{X}_0^-(k) + \dot{X}_S^+(k) + \dot{X}_S^-(k) \tag{8}$$

Based on the MPE algorithm, $\dot{X}(k)$ can be transformed into a modal model consisting of four modes as shown in (9)-(11).

$$\dot{X}(k) = \sum_{m=1}^{4} R_m e^{w_m k} \tag{9}$$

$$\begin{cases} R_1 = Q^*(f_0, 1) x_0 e^{-j\phi_0} \\ R_2 = Q(f_0, -1) x_0 e^{j\phi_0} \\ R_3 = Q^*(f_{sub}, 1) x_{sub} e^{-j\phi_{sub}} + Q(f_{sup}, -1) x_{sup} e^{j\phi_{sup}} \\ R_4 = Q(f_{sub}, -1) x_{sub} e^{j\phi_{sub}} + Q^*(f_{sup}, 1) x_{sup} e^{-j\phi_{sup}} \end{cases} \tag{10}$$

$$\begin{cases} w_1 = j\alpha \\ w_2 = -j\alpha \\ w_3 = j\beta \\ w_4 = -j\beta \end{cases} \tag{11}$$

where R_m and w_m ($m = 1, 2, 3, 4$) are the amplitude and angular frequency of mode m, respectively; and $R_1 e^{w_1}$ and $R_2 e^{w_2}$ correspond to fundamental modes, and $R_3 e^{w_3}$ and $R_4 e^{w_4}$ correspond to oscillation modes.

Note that only k is a variable in (9) representing the features of angular frequency changes with time. As shown in Fig. 1, the difference between adjacent synchrophasors is sig-

nificant, which is formed by the superposition of the angular frequency of four modes. It is difficult to extract the features of angular frequency directly.
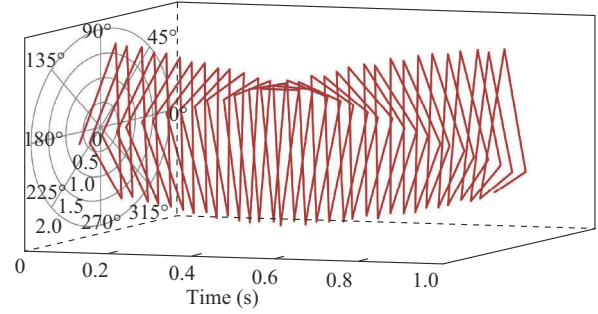

Fig. 1.   Single-phase current synchrophasors.

B. Process of MPM

The basic idea of MPM is shown in Fig. 2.
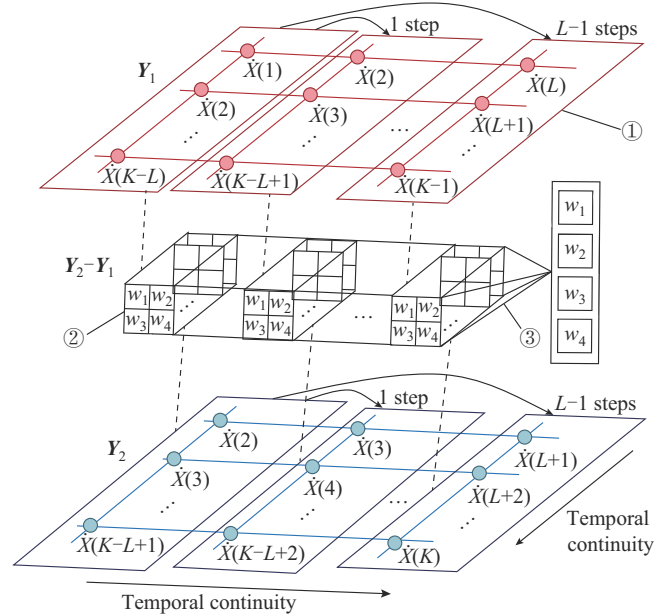
Fig. 2.   Basic idea of MPM.

The first step is to construct a Hankel matrix $\boldsymbol{Y}$ based on synchrophasors as:

$$\boldsymbol{Y} = \begin{bmatrix} \dot{X}(1) & \dot{X}(2) & \cdots & \dot{X}(L+1) \\ \dot{X}(2) & \dot{X}(3) & \cdots & \dot{X}(L+2) \\ \vdots & \vdots & & \vdots \\ \dot{X}(K-L) & \dot{X}(K-L+1) & \cdots & \dot{X}(K) \end{bmatrix}_{(L+1)\times(K-L)} \tag{12}$$

where K is the number of synchrophasors; $L+1$ is the number of columns of the Hankel matrix; and $\boldsymbol{Y}$ can be rewritten as:

$$\boldsymbol{Y} = \boldsymbol{Z}_1 \boldsymbol{R} \boldsymbol{Z}_2 \tag{13}$$

$$\boldsymbol{Z}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ e^{w_1} & e^{w_2} & e^{w_3} & e^{w_4} \\ \vdots & \vdots & \vdots & \vdots \\ e^{w_1(K-L-1)} & e^{w_2(K-L-1)} & e^{w_3(K-L-1)} & e^{w_4(K-L-1)} \end{bmatrix} \tag{14}$$

$$Z_2 = \begin{bmatrix} 1 & e^{w_1} & \cdots & e^{w_1 L} \\ 1 & e^{w_2} & \cdots & e^{w_2 L} \\ 1 & e^{w_3} & \cdots & e^{w_3 L} \\ 1 & e^{w_4} & \cdots & e^{w_4 L} \end{bmatrix} \tag{15}$$

$$R = \text{diag}(R_1, R_2, R_3, R_4) \tag{16}$$

where matrix Z_1 represents the time continuity of each column in Y; matrix Z_2 represents the time continuity of each row in Y; and matrix R represents the amplitude of the four modes. Then, Y can be decomposed into two submatrices Y_1 and Y_2 as:

$$\begin{cases} Y_1 = Y(:, 1:L) = Z_1 R Z_2(:, 1:L) \\ Y_2 = Y(:, 2:L+1) = Z_1 R Z_2(:, 2:L+1) \end{cases} \tag{17}$$

As illustrated in Fig. 2, Y_1 denotes $K - L$ phasors shifted forward by 1 to $L-1$ steps, and each row and column in Y_1 has the features of temporal continuity, so does Y_2 (see part ①). As the difference between the $K-L$ phasors corresponding to Y_1 and Y_2 is one step, each point in $Y_2 - Y_1$ denotes temporal continuity between two adjacent synchrophasors (see part ②), which is equal to the superposition of changes after the angular frequencies of each phasor in Y_1 increase by w_1, w_2, w_3, and w_4, reflecting the features of the angular frequencies of the four modes (see part ③). Thus, by performing feature extraction to average the temporal continuity of each point in $Y_2 - Y_1$, the angular frequencies of the four modes can be obtained. Therefore, the matrix pencil $Y_2 - \lambda Y_1$ can be constructed based on Y_1 and Y_2 as:

$$Y_2 - \lambda Y_1 = Z_1 R (Z_0 - \lambda I) Z_2(:, 1:L) \tag{18}$$

where $Z_0 = \text{diag}(e^{w_1}, e^{w_2}, e^{w_3}, e^{w_4})$; and I is a 4×4 unit matrix.

When λ is equal to any one of the exponential of the angular frequency among the four modes, the order of $Y_2 - \lambda Y_1$ can be reduced, and the characteristic polynomial of $Y_2 - \lambda Y_1$ is equal to 0 as:

$$|Y_2 - \lambda Y_1| = 0 \tag{19}$$

Finally, by solving (19), the generalized eigenvalues of $Y_2 - \lambda Y_1$, i.e., w_m can be obtained, and R_m can be further calculated based on w_m and (9).

III. Improved SSO Parameter Identification Based on MPM

A. Deficiencies in Solving Matrix Pencil

In [20], the angular frequencies of the four modes were calculated by solving the generalized eigenvalues of the matrix pencil. Since solving the eigenvalues transforms the matrix pencil into a set of nonlinear correlated basis vectors, each eigenvalue denotes the projected lengths of the matrix pencil on the basis vector, and the angular frequency conjugation constraints of the two fundamental modes and two oscillation modes as (11) are ignored. The previous derivation is based on an ideal model, so there is a one-to-one correspondence between the eigenvalues of the matrix pencil and the angular frequencies of the four modes. However, in practical applications, in the presence of model errors, w_1 and w_2 obtained by solving the eigenvalues of the matrix pencil do

not meet the $\pm j\alpha$ conjugate constraint, and w_3 and w_4 do not meet the $\pm j\beta$ conjugate constraint either. Therefore, the accuracy of α and β is difficult to guarantee and will cause additional errors for parameter identification of each component in SSO.

Furthermore, as f_0 is close to nominal frequency, the value of $|Q^*(f_0, 1)|$ is quietly small, and even tends to zero when f_0 tends to 50 Hz, as shown in Fig. 3. The fundamental mode $R_1 e^{w_1}$ corresponding to $Q^*(f_0, 1)$ is highly sensitive to model errors. Not only w_1 and w_2 do not meet the conjugate constraint, but the calculation error of w_1 is significant. It is difficult to distinguish which eigenvalue corresponds to α, and the parameters of the fundamental component identified by α suffer from low accuracy.
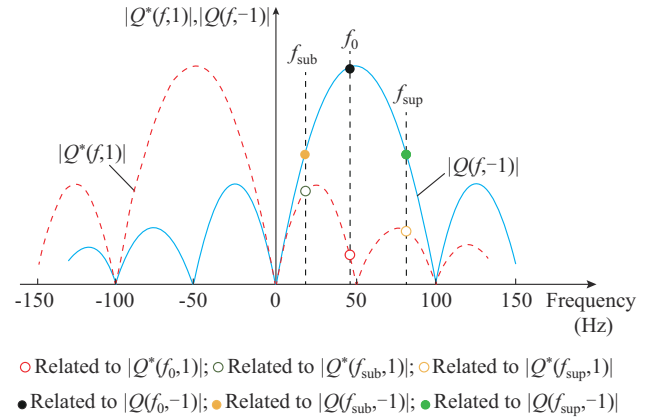

Fig. 3. $|Q^*(f, 1)|$ and $|Q(f, -1)|$ varing with frequency.

The fitness of MPM is checked by the coefficient of determination in [20], which reflects the difference between original synchrophasors and reconstructed synchrophasors based on the identified parameters. However, even if the fitness is successful, it just means the original and reconstructed synchrophasors are close to each other and does not mean α is solved accurately, because the amplitude of the fundamental mode $R_1 e^{w_1}$ tends to zero and has little effect on the overall synchrophasors.

B. Basic Idea of Improved SSO Parameter Identification Based on MPM

The basic idea of the proposed improved MPM is to establish the angular frequency fitting equations corresponding to fundamental and oscillation frequencies based on the characteristic polynomial coefficients of the matrix pencil, in which the angular frequency conjugation constraints of the two fundamental modes and two oscillation modes are considered, thus guaranteeing the accuracy of calculated α and β.

Based on the properties of the matrix characteristic polynomial, (19) can be derived as:

$$\sum_{i=0}^{4} a_i \lambda^i = 0 \tag{20}$$

where a_i is the characteristic polynomial coefficient of the matrix pencil.

Since a_i can be obtained based on the eigenvalues λ of the matrix pencil, the characteristic polynomial coefficients are

equivalent to eigenvalues. In the presence of model errors, a_i calculated by λ suffers from calculation errors; the value on the left side of (20) is no longer zero when the correct α or β is substituted, and residuals appear instead. By searching α or β that minimizes the residuals, α or β can be solved exactly.

Therefore, after constructing the matrix pencil, the characteristic polynomial coefficients of the matrix pencil are calculated first, and then the fitting equations of α and β are established based on the characteristic polynomial coefficients of the matrix pencil as:

$$\begin{bmatrix} \sum_{i=0}^{4} a_i \mathrm{e}^{\mathrm{j}\alpha i} \\ \sum_{i=0}^{4} a_i \mathrm{e}^{-\mathrm{j}\alpha i} \end{bmatrix} + \begin{bmatrix} \varepsilon(\alpha) \\ -\varepsilon(\alpha) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{21}$$

$$\begin{bmatrix} \sum_{i=0}^{4} a_i \mathrm{e}^{\mathrm{j}\beta i} \\ \sum_{i=0}^{4} a_i \mathrm{e}^{-\mathrm{j}\beta i} \end{bmatrix} + \begin{bmatrix} \varepsilon(\beta) \\ -\varepsilon(\beta) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{22}$$

where $\varepsilon(\alpha)$, $\varepsilon(-\alpha)$, $\varepsilon(\beta)$, and $\varepsilon(-\beta)$ are the residuals of the fitting equations.

On the one hand, the fitting equations of α and β replace the eigenvalues of the matrix pencil with the characteristic polynomial coefficients, which can approximately characterize the features of the angular frequencies of the four modes. On the other hand, the conjugate constraints of $\pm \mathrm{j}\alpha$ and $\pm \mathrm{j}\beta$ are considered. Solving the minimum residuals corresponding to α and β will eliminate α and β with more significant errors. The obtained α and β can satisfy the conjugate constraints of $\pm \mathrm{j}\alpha$ and $\pm \mathrm{j}\beta$, and they are closest to the eigenvalues of the matrix pencil.

The numerical solution of (21) and (22) can be obtained as α and β by minimizing the residuals $\left\| \varepsilon(\alpha) \right\|_2 + \left\| \varepsilon(-\alpha) \right\|_2$ and $\left\| \varepsilon(\beta) \right\|_2 + \left\| \varepsilon(-\beta) \right\|_2$, respectively. Then, the amplitude of each mode R_m can be obtained based on (9), and the parameters of the fundamental, subsynchronous, and supersynchronous components can be obtained based on (10).

C. Angle Frequency Fitting Equations Based on Multi-order Matrix Pencil

In the existing studies, Hankel matrix Y is decomposed into two submatrices to construct the matrix pencil, which contains model errors corresponding to K phasors in solving the eigenvalues of the matrix pencil. The multi-order matrix pencil is proposed to reduce the impact of the randomness of model errors in this subsection. Y is decomposed into multiple submatrices so that the matrix pencil can be extended to multiple orders. Based on the mean value of the characteristic polynomial coefficients of the multi-order matrix pencil, the angle frequency fitting equations of α and β can be constructed. Hence, the model errors are smoothed to further improve the accuracy of α and β.

First, Y is decomposed into multiple submatrices as:

$$Y_p = \begin{bmatrix} \dot{X}(p) & \dot{X}(p+1) & \dots & \dot{X}(l+p) \\ \dot{X}(p+1) & \dot{X}(p+2) & \dots & \dot{X}(l+p+1) \\ \vdots & \vdots & & \vdots \\ \dot{X}(K-P+p-l) & \dot{X}(K-P+p-l+1) & \dots & \dot{X}(K-P+p) \end{bmatrix} \tag{23}$$

where $l+1$ is the number of columns of the submatrix Y_p and $l < L$; and P is the number of submatrices, $p = 1, 2, ..., P$.

The multi-order matrix pencil is constructed as:

$$Y_{p+1} - \lambda^{(p)} Y_p \tag{24}$$

where the superscript (p) indicates the pth matrix pencil, $p = 1, 2, ..., P-1$.

Multiple sets of characteristic polynomial equations can be obtained by substituting $\lambda^{(1)}$-$\lambda^{(p-1)}$ into (20). Since α and β are solutions of all characteristic polynomial equations, by superimposing these characteristic equations directly, the angle frequency fitting equations of α and β based on the multi-order matrix pencil can be constructed as:

$$\begin{bmatrix} \sum_{p=1}^{P-1} \sum_{i=0}^{4} a_i^{(p)} \mathrm{e}^{\mathrm{j}\alpha i} \\ \sum_{p=1}^{P-1} \sum_{i=0}^{4} a_i^{(p)} \mathrm{e}^{-\mathrm{j}\alpha i} \end{bmatrix} + \begin{bmatrix} \varepsilon(\alpha) \\ -\varepsilon(\alpha) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{25}$$

$$\begin{bmatrix} \sum_{p=1}^{P-1} \sum_{i=0}^{4} a_i^{(p)} \mathrm{e}^{\mathrm{j}\beta i} \\ \sum_{p=1}^{P-1} \sum_{i=0}^{4} a_i^{(p)} \mathrm{e}^{-\mathrm{j}\beta i} \end{bmatrix} + \begin{bmatrix} \varepsilon(\beta) \\ -\varepsilon(\beta) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{26}$$

The solutions of (25) and (26) are the same as that of (21) and (22), respectively.

D. Data Window

For identification of SSO based on the synchrophasors, the shorter the data window is, the less low-frequency information in the synchrophasors is, and it is more difficult to identify SSO when the subsynchronous frequency f_{sub} is close to the fundamental frequency f_0.

Since the DFT-based algorithm performs an approximation calculation in the process of identification, the errors of identification results are still significant when f_{sub} is close to f_0, even with a 2-s data window [16], [23]. On the contrary, the MPM does not perform any approximation calculation in the modeling process and can accurately identify SSO when f_{sub} is below 45 Hz with a shorter data window.

The existing studies of MPM ignore the angular frequency conjugation constraints of the two fundamental modes and two oscillation modes when solving the eigenvalues of the matrix pencil directly, so it has to use a 1-s data window to ensure the accuracy of parameter identification. It is precise because the proposed improved MPM compensates for this deficiency that the advantage of not performing the approximation calculation of the MPM can be fully utilized, and the data window for identifying SSO can be shortened to 200 ms.

IV. IDENTIFICATION AND RECOVERY OF BAD DATA

A. Noise Reduction

Noise is one of the factors leading to model errors and will affect the accuracy of parameter identification based on MPM. Before implementing the MPM, singular value decomposition (SVD) based rank reduction is usually applied to the Hankel matrix Y [20] as:

$$Y = UDV^{\mathrm{T}} \tag{27}$$

where D is a diagonal matrix with singular values of Y; and U and V are two unitary matrices. As the first four singular values in D correspond to each of the four modes, by intercepting the first four singular values as $D' = D(1:4, 1:4)$ and the corresponding singular phase quantities as $U' = U(:, 1:4)$ and $V' = V(:, 1:4)$, the noise that is less significant than the four modes can be removed. In turn, the noise reduction is realized, and the Hankel matrix after noise reduction Y' is given as:

$$Y' = U'D'V'^{\mathrm{T}} \tag{28}$$

As only the first four singular values are retained after the noise reduction, the larger the dimension of the matrix D is, the better the noise reduction performance is. For a fixed number of synchrophasors, the dimension of D is the largest when the Hankel matrix is a square matrix, so the Hankel matrix will be constructed as a square matrix in this paper to maximize the noise elimination.

The reason why SVD can still retain the features of four modes despite the error of w_1 is significant is that the amplitude corresponding to w_1 is small enough, thus weakening the effect of w_1 on the whole fundamental mode $R_1 e^{w_1}$. As long as the noise is less significant to the four modes, noise can be reduced to some extent.

B. Identification of Bad Data

Although noise will cause deviations in all of the synchrophasors, and bad data will cause deviation in a single data point, the effects of noise and bad data on the performance of the MPM are not substantially different, as these deviations are averaged to each row and column of the Hankel matrix during the feature extraction process. When the bad data deviates seriously, the temporal continuity of synchrophasors in the Hankel matrix may be broken. Hence, the parameters obtained by the MPM will be totally wrong, and it is difficult to eliminate the impact of bad data through noise reduction.

The process of noise reduction with single bad data is illustrated in Fig. 4. As the SVD will average the deviations of bad data to each row and column of the Hankel matrix, the singular values are all affected by bad data. Since only the temporal continuity at the position of bad data is broken, while the other phasors are not, the first four singular values can still retain the features of the four modes. After the rank reduction, the characteristics of bad data are weakened, and then the features of the four modes together with the weakened bad data are reconstructed into Y', so the breaking extent of temporal continuity in each row and column of Y' is substantially reduced. After differentiating Y and Y', the dif-

ference of the phasors where the temporal continuity is most severely broken is most significant, i.e., the bad data. Therefore, to evaluate the significance of bad data, the ratio of the maximum difference to the average difference of the synchrophasors before and after noise reduction is calculated in this paper.
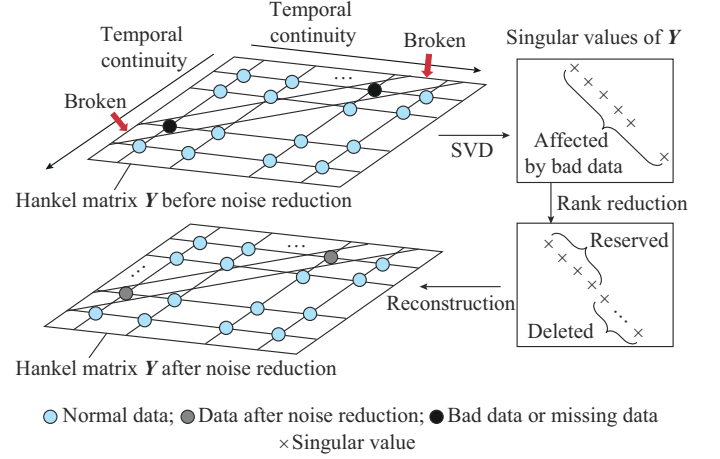

Fig. 4. Process of noise reduction with single bad data.

By defining $|E| = |Y - Y'|$, the points in E are as:

$$\varepsilon_{ij} = \left| \dot{X}_{ij} - \dot{X}_{ij}' \right| \quad \dot{X}_{ij} \in Y, \ \dot{X}_{ij}' \in Y' \tag{29}$$

where E denotes the difference of Y and Y'; ε_{ij} is the point of the ith row and j^{th} column in E; and $\dot{X}_{ij}$ and $\dot{X}_{ij}'$ are the points in Y and Y', respectively.

The difference of the kth synchrophasors before and after taking the noise reduction $\varepsilon(k)$ is calculated as:

$$\varepsilon(k) = \bar{\varepsilon}_{ij} \quad i + j = k + 1 \tag{30}$$

Then, the ratio of the maximum difference to the average difference of the synchrophasors before and after taking noise reduction r is:

$$r = \frac{\max(\varepsilon(k))}{\dfrac{1}{K} \sum_{k=1}^{K} \varepsilon(k)} \tag{31}$$

The value of r reflects the prominence of the difference in bad data compared with the average difference. Hence, the significance of bad data can be evaluated by comparing the calculated r with a threshold value. When the calculated r is less than the threshold value, the significance of bad data is relatively low, and the impact of bad data is minimal, which can be treated as noise. The calculated error of parameter identification can be decreased to an acceptable range by noise reduction. In contrast, when the calculated r is greater than the threshold value, the difference in bad data is significant enough to be identified by locating $\max(\varepsilon(k))$. In addition, the presence of bad data may yield obvious errors for parameter identification, the effect of which is difficult to eliminate by noise reduction, so the recovery of the bad data is necessary.

Note that the threshold value of r proposed in this paper is not an indicator to distinguish bad data and non-bad data, but a robustness indicator to reduce the impact of bad data

on the performance of the MPM. From this perspective, the threshold value does not need to be solved precisely, while it is necessary to ensure that the impact of bad data is minimal enough when the calculated r is close to the threshold value. Although the misclassification or omission of bad data may occur under these circumstances, performing either noise reduction or the recovery of bad data can eliminate the impact of bad data.

C. Recovery of Bad Data

The basic idea of the recovery of bad data is to correct bad data by performing iterative SVD-based rank reduction of Y, and replacing the points related to bad data in Y with the points related to bad data in Y' until the difference of bad data satisfies the recovery accuracy criterion. The process of the recovery of bad data is as follows.

Step 1: begin with the nth SVD-based rank reduction of existing Hankel matrix $Y_{(n)}$, and obtain reconstruction matrix $Y'_{(n)}$, where the subscript (n) indicates the number of iterations, and set $n = 1$.

Step 2: calculate the difference between $Y_{(n)}$ and $Y'_{(n)}$, obtain $\varepsilon_{ij(n)}$, and then calculate the difference of bad data $\varepsilon(k_{error})_{(n)}$ based on $\varepsilon_{ij(n)}$ as:

$$\varepsilon(k_{error})_{(n)} = \bar{\varepsilon}_{ij(n)} \quad i+j = k_{error}+1 \tag{32}$$

where k_{error} is the position of bad data, and it is obtained previously as introduced in Section IV-B.

Step 3: determine if $\varepsilon(k_{error})_{(n)}$ meets the recovery accuracy criterion. If $\varepsilon(k_{error})_{(n)}$ is less than the recovery accuracy criterion, the current $Y'_{(n)}$ is sufficiently accurate, and the recovery of bad data is unnecessary. If not, obtain the position of the point with the minimum difference of bad data as $\min(\varepsilon_{ij(n)}, i+j = k_{error}+1)$, and then search the point $\dot{X}'_{ij_{\min}(n)}$ corresponding to the bad data with the minimum difference in the same position from $Y'_{(n)}$. Set the existing Hankel matrix in the next iteration as $Y_{(n+1)} = Y_{(n)}$, and replace the point corresponding to bad data in $Y_{(n+1)}$ with $\dot{X}'_{ij_{\min}(n)}$ as:

$$\dot{X}_{ij(n+1)} \quad i+j = k_{error}+1 = \dot{X}'_{ij_{\min}(n)} \tag{33}$$

Finally, continue to the next iteration with $n = n+1$, until $\varepsilon(k_{error})_{(n)}$ satisfies the recovery accuracy criterion.

D. Missing Data

Actual measurements may have missing data in addition to bad data. Since the synchrophasors provided by PMU have unified time stamps, if the synchrophasor corresponds to any time is missing, the position of missing data can be determined based on the unified time stamps.

After determining the position of missing data, when assigning an extremely large or small value compared to other synchrophasors to the position of missing data, it is equivalent to the presence of bad data in the synchrophasors and the recovery of the missing data can be addressed based on the identification and recovery of the bad data, as shown in Fig. 4.

V. VERIFICATION

The overall characteristics of the proposed improved MPM is verified with synthetic and simulated PMU data in Section V-A and V-B, respectively. Meanwhile, the performance of the proposed identification and recovery of bad data are verified based on the synthetic PMU data in Section V-C.

A. Verification I: Synthetic PMU Data

The synthetic PMU data are modeled as (1), and the fundamental frequency f_0 is set as $[49, 49.5, 49.7, 50, 50.5, 51, 51.5]$Hz. Other parameters of fundamental, subsynchronous, and supersynchronous components are set as $(x_0, \phi_0) = (100, \pi/3)$, $(f_{sub}, x_{sub}, \phi_{sub}) = (20, 20, \pi/2)$, and $(f_{sup}, x_{sup}, \phi_{sup}) = (80, 10, \pi/4)$, respectively.

The sampling frequency of the instantaneous data is 1.6 kHz, and the reporting rate of PMU is 100 Hz. Performing the DFT algorithm on the instantaneous data can obtain the synthetic PMU data. Then, the improved SSO parameter identification based on MPM is implemented with a 200-ms data window, i.e., 21 synchrophasor data.

The primary research [23] indicated that identification accuracy is most affected by f_0 and f_{sub}. Therefore, two sets of tests are conducted in this subsection with f_{sub} varies in the range of $[5, 45]$Hz, and x_{sub} varies in the range of $[5, 100]$, with one parameter changing under various conditions while other parameters keep the above settings.

As the measurement signal to noise ratio (SNR) of PMUs is around 45 dB [16], to study the characteristics of identification accuracy under noise interference, the noise conditions with 40 dB SNR and 30 dB SNR are considered. The function of SNR is $SNR = 10\lg((x_0^2 + x_{sub}^2 + x_{sup}^2)/(2\sigma^2))$, and σ is variance of noise. Furthermore, the results of the existing MPM [20] and the proposed improved MPM are compared. Since the data window length in [20] is 1 s, a 200-ms data window and a 1-s data window are used on the existing MPM for further comparison. The relative errors of both the improved and existing MPMs under ideal and noisy conditions are shown in Table I, in which "IMPM" represent the proposed improved MPM.

1) Relative Errors Under Ideal Conditions

The relative errors of parameters of each component in SSO obtained by the improved MPM are on the order of $10^{-16} \text{-} 10^{-13}$, which is closer to the accuracy of the existing MPM with the 1-s data window and superior to that of the existing MPM with the 200-ms data window.

2) Relative Errors Under Noise Conditions

Whether using a 200-ms data window or a 1-s data window, the relative errors of x_0 and ϕ_0 obtained by the existing MPM are approximately 10%, and the main reason is stated in Section III-A. As the conjugate constraints of $\pm j\alpha$ are considered in the fitting equations of α, the relative errors of the fundamental components obtained by the improved MPM are below the order of 10^{-2}.

According to the theory of spectral analysis, the shorter the data window is, the lower the spectrum resolution ratio is, and the interaction between the fundamental component and the oscillation component is severe. Therefore, under the noise condition, corrupted by the low identification accuracy of the fundamental components, the relative error of the oscillation components obtained by the existing MPM with the

200-ms data window is significant, among which the relative error of the amplitude and phase reaches 15% and 40%, respectively. In contrast, as the 1-s data window is long enough, the interaction between the fundamental component and the oscillation component is insignificant, and the relative error of the oscillation components obtained by the existing MPM with the 1-s data window is approximately 5%. Since the improved MPM can identify the fundamental component parameters with high accuracy, the relative error of the oscillation components obtained by the improved MPM with the 200-ms data window is still below 5%.

TABLE I
RELATIVE ERRORS OF IDENTIFICATION BASED ON IMPROVED MPM AND EXISTING MPM UNDER IDEAL AND NOISY CONDITIONS

Test set	Method	SNR (dB)	$\lg\frac{|\hat{f}_0-f_0|}{f_0}$	$\lg\frac{|\hat{x}_0-x_0|}{x_0}$	$\lg\frac{|\hat{f}_{sub}-f_{sub}|}{f_{sub}}$	$\lg\frac{|\hat{x}_{sub}-x_{sub}|}{x_{sub}}$	$\lg\frac{|\hat{x}_{sup}-x_{sup}|}{x_{sup}}$	$\lg\frac{|\hat{\phi}_{sup}-\phi_{sup}|}{\phi_{sup}}$	$\lg\frac{|\hat{x}_{sup}-x_{sup}|}{x_{sup}}$
$f_{sub}\in[5,55]$Hz	IMPM (200 ms)	∞	-16	-14	-15.0	-13.0	-13.0	-13.0	-14.0
		40	-3	-2	-3.0	-2.0	-2.0	-2.0	-1.3
		30	-3	-2	-2.0	-1.3	-1.3	-1.3	-1.3
	MPM (200 ms)	∞	-10	-10	-12.0	-10.0	-9.0	-10.0	-10.0
		40	-2	**-1**	-2.0	**-1.0**	**-1.0**	-1.3	**-1.0**
		30	-2	**-1**	-1.3	-1.0	**-1.0**	-1.0	**-0.5**
	MPM (1 s)	∞	-14	-12	-15.0	-13.0	-13.0	-13.0	-13.0
		40	-3	**-1**	-4.0	-2.0	-2.0	-2.0	-1.3
		30	-3	**-1**	-3.0	-1.3	-1.3	-1.3	-1.3
$x_{sub}\in[5,100]$	IMPM (200 ms)	∞	-16	-15	-16.0	-14.0	-14.0	-14.0	-14.0
		40	-3	-2	-3.0	-2.0	-2.0	-2.0	-2.0
		30	-3	-2	-3.0	-1.3	-1.3	-2.0	-2.0
	MPM (200 ms)	∞	-13	-12	-14.0	-12	-12.0	-13	-12.0
		40	-2	**-1**	-3.0	**-1.0**	**-1.0**	-1.3	-1.3
		30	-2	**-1**	-3.0	**-1.0**	**-1.0**	-1.3	-1.3
	MPM (1 s)	∞	-15	-13	-15.0	-14.0	-13.0	-14.0	-13.0
		40	-3	**-1**	-4.0	-2.0	-2.0	-2.0	-2.0
		30	-2	**-1**	-4.0	-1.3	-1.3	-2.0	-2.0

Note: ① the bolded number means the relative errors are beyond $10^{-1.3}\approx 5\%$. ② $SNR=\infty$ means the ideal condition. ③ The relative errors of $|\hat{\phi}_0-\phi_0|/\phi_0$ and $|\hat{f}_{sup}-f_{sup}|/f_{sup}$ are similar to $|\hat{x}_0-x_0|/x_0$ and $|\hat{f}_{sub}-f_{sub}|/f_{sub}$, respectively, so both of them are not listed here.

Consequently, compared with the existing MPM, the proposed improved MPM can accurately identify the parameters for the fundamental, subsynchronous, and supersynchronous components with a 200-ms data window under both frequency-varying conditions and noisy conditions, which significantly improves the real-time performance.

B. Verification II: Simulated PMU Data and Actual Measured PMU Data

To further verify the effectiveness of the proposed improved MPM, cases with both simulated PMU data and actual measured PMU data are conducted. The simulated PMU data are generated in the PSCAD platform based on the model in [1], as shown in Fig. 5, and the details of the model and the simulation parameters are the same with those in [1]. The simulation system is wind farm with doubly-fed induction generator (DFIG) connected to series-compensated system. The system operates in a stable state by setting suitable parameters, and the SSO occurs as the wind speed decreases. The actual measured PMU data come from [23], which are recorded in an actual SSO incident that occurred in the North China power grid. The current instantaneous data and synchrophasors are illustrated in Fig. 6 and Fig. 7 for simulated PMU data and actual measured PMU data, respectively.

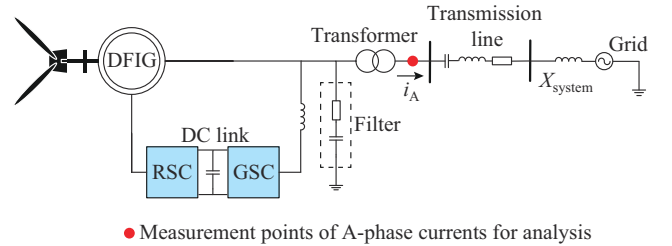● Measurement points of A-phase currents for analysis

Fig. 5. Diagram of simulation model.

The simulated PMU data with a 6-s data window are analyzed as in Fig. 6. In order to ensure the reliability of the interpolated DFT as a reference and make a reasonable comparison to verify the accuracy of the proposed improved MPM, the data window is selected at a stabilization stage of SSO. The actual measured PMU data with a 10-s data window are analyzed as in Fig. 7, which is at a fast-changing stage of SSO. The proposed improved MPM is performed on a 200-ms data window. Also, the existing MPM [20] with 1-s data window and the interpolated DFT [16] with 2-s data window are employed for comparison and reference, respectively. The identified results of simulated PMU data case and actual measured PMU data case are illustrated in Fig. 8 and Fig. 9, respectively.
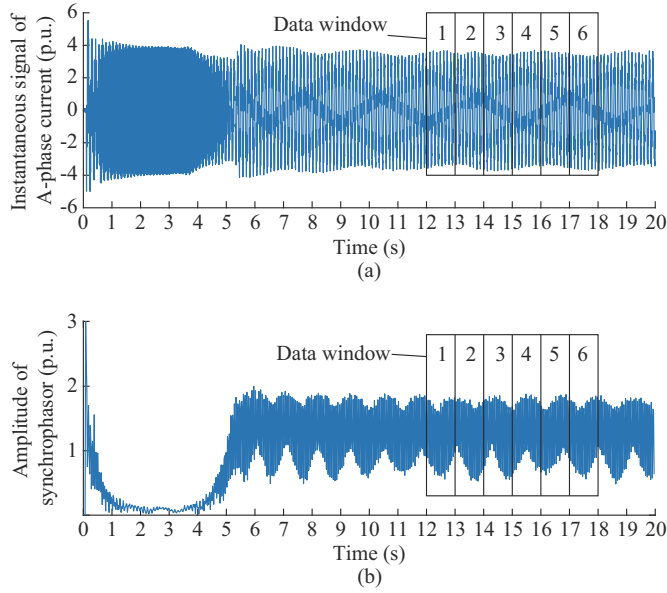
Fig. 6. Instantaneous data and synchrophasors for simulated PMU data case. (a) Instantaneous data. (b) Amplitude of synchrophasor.
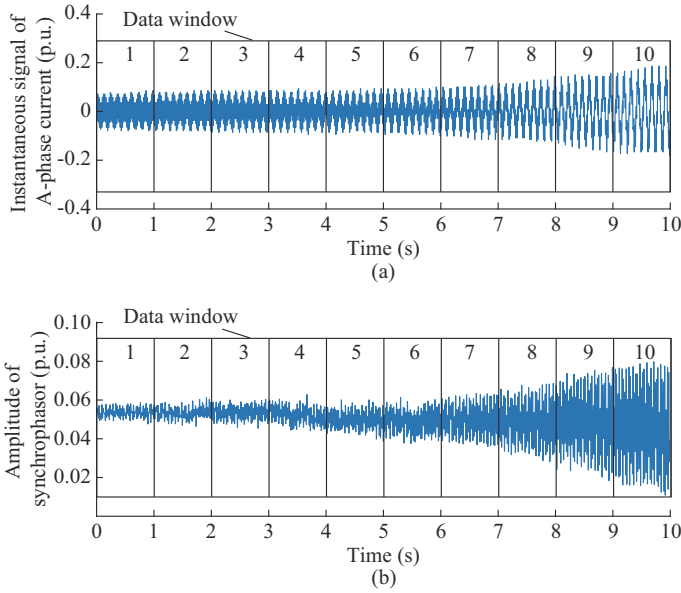

Fig. 7. Instantaneous data and synchrophasors for actual measured PMU data case. (a) Instantaneous data. (b) Amplitude of synchrophasor.

1) Fundamental Component

In the simulated PMU data case, the results of f_0 obtained by the improved MPM fluctuate at approximately 49.47 Hz to 49.57 Hz, and those of interpolated DFT are around 49.50 Hz. However, f_0 in the 5th calculation of the existing MPM is 49.90 Hz, which deviates from the improved MPM and the interpolated DFT. The results of x_0 obtained by the improved MPM are similar to those of the interpolated DFT. Affected by the deviation of f_0, the results of x_0 in the 5th calculation of the existing MPM are inaccurate. In the actual measured PMU data case, the variation ranges of f_0 and x_0 obtained by the improved MPM are similar to those of the existing MPM and interpolated DFT but with more details. Therefore, compared with the existing MPM, the proposed im-

proved MPM can identify the frequency and amplitude of fundamental component more accurately.
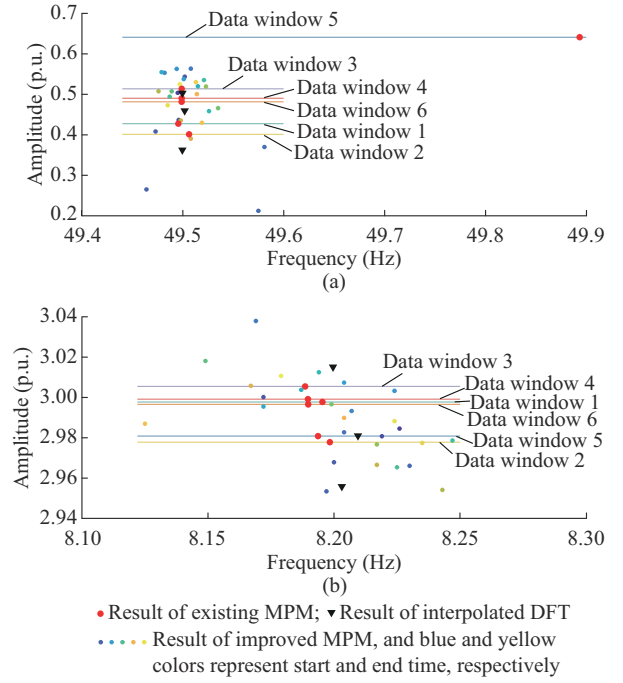

Fig. 8. Identified results of frequency and amplitude for simulated PMU data case. (a) Fundamental component. (b) Subsynchronous component.
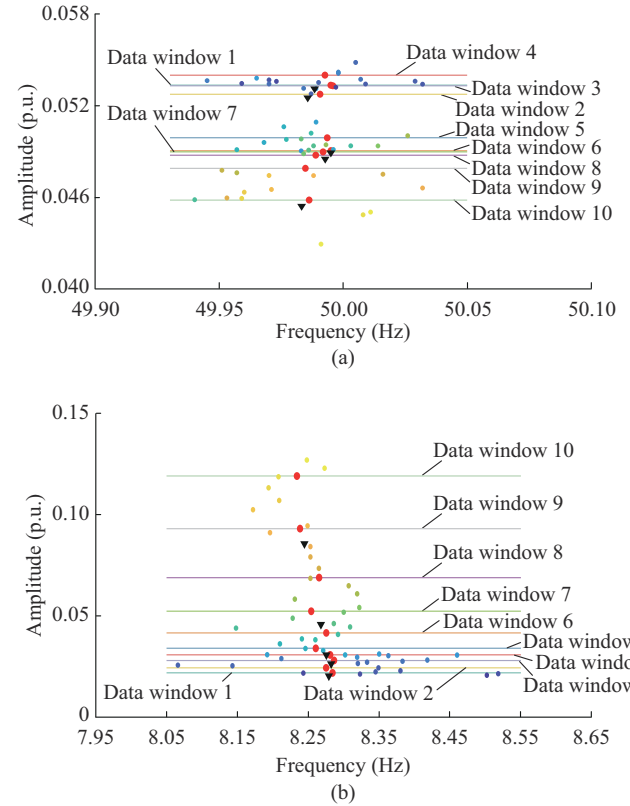

Fig. 9. Identified results of frequency and amplitude for actual measured PMU data case. (a) Fundamental component. (b) Subsynchronous component.

2) Oscillation Component

The results of the supersynchronous component have similar characteristics to the subsynchronous one, so only the results of the subsynchronous component are analyzed here. In the simulated PMU data case, the results of f_{sub} obtained by the improved MPM fluctuate at approximately 8.13 Hz to 8.25 Hz, and those obtained by the existing MPM and interpolated DFT are around 8.19 Hz; the results of x_{sub} obtained by the improved MPM fluctuate at approximately 2.95 p.u. to 3.04 p.u., and those obtained by the existing MPM and IDF are around 3.0 p.u.. The proposed improved MPM can identify the frequency and amplitude of oscillation component accurately.

In the actual measured PMU data case, the results of f_{sub} obtained by the improved MPM fluctuate between approximately 8.14 Hz to 8.52 Hz, and those obtained by the existing MPM and interpolated DFT fluctuate around 8.24 Hz to 8.28 Hz; the variation range of the results of x_{sub} obtained by the improved MPM are similar to those obtained by the existing MPM and interpolated DFT, but f_{sub} and x_{sub} obtained by improved MPM change rapidly with time. Since the identified results of the existing MPM and interpolated DFT are the average of dynamics in 1-s and 2-s data windows and cannot reflect the actual characteristics of SSO at a particular moment, especially during the fast-changing stage of SSO, owing to the 200-ms data window, the improved MPM can intuitively monitor the dynamics of the frequency and amplitude of each component in the process of SSO.

C. Verification III: Performance of Identification and Recovery of Bad Data

This subsection verifies the performance of identification and recovery of bad data with synthetic PMU data. The parameters of fundamental, subsynchronous, and supersynchronous components are $(f_0, x_0, \phi_0) = (50, 50, \pi/3)$, $(f_{sub}, x_{sub}, \phi_{sub}) = (10, 100, \pi/6)$, and $(f_{sup}, x_{sup}, \phi_{sup}) = (90, 100, \pi/4)$, respectively; the other parameters are the same as those in Section V-A; and the noise with 45 dB SNR is added.

As the improved MPM is performed on 200-ms data windows and contains 21 synchrophasor data for each calculation, the constructed Hankel matrix is an 11×11 symmetric matrix in this paper, and the 4th, 7th, and 10th points of the synchrophasor are selected as bad data to perform three tests.

To analyze the effect of bad data, the total vector error (TVE) ε_{TVE} defined by IEEE standard [24] is used to generate multiple bad data with ε_{TVE} ranging from 0% to 500% as:

$$\varepsilon_{TVE} = \frac{\left\| \hat{\dot{X}} - \dot{X} \right\|_2}{\left\| \dot{X} \right\|_2} \times 100\% \qquad (34)$$

where $\hat{\dot{X}}$ and $\dot{X}$ are the bad data and initial phasor data, respectively.

1) Ratio r

The ratio of the maximum difference to the average difference r with bad data under different ε_{TVE} is illustrated in Fig. 10. Note that when bad data is at the 10th point, as the ε_{TVE} of bad data increases, r gradually increases from approximately 1 to 5. Meanwhile, the trend of r at the 4th and 7th

points are similar to that of at the 10th point. Only for points at different locations, due to the fact that the magnitudes of the initial phase data are different, r changes at different rates. Thus, the value of r can reflect the significance of bad data.
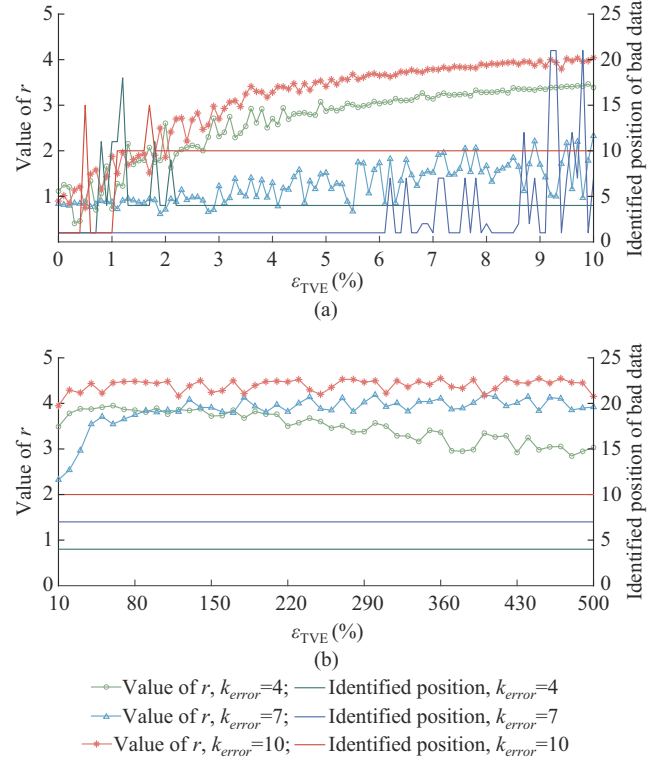

Fig. 10. Value of r and identified position of bad data under different ε_{TVE}. (a) $\varepsilon_{TVE} \in [0, 10\%]$. (b) $\varepsilon_{TVE} \in [10\%, 500\%]$.

2) Identification of Bad Data

The position of bad data identified by locating $\max(\varepsilon(k))$ with bad data under different ε_{TVE} is also shown in Fig. 10. For the test with bad data at the 4th point, bad data can be identified when ε_{TVE} is greater than 2.2%, and the corresponding value of r is approximately 2.0. For the test with bad data at the 7th point, bad data can be identified when ε_{TVE} is greater than 10%, and the corresponding value of r is approximately 2.3. Also, for the test with bad data at the 10th point, bad data can be identified when ε_{TVE} is greater than 1.8%, and the corresponding value of r is approximately 1.9. It is feasible to evaluate the significance of bad data by comparing the calculated r with a threshold value, and 2.5 can be chosen as the threshold value by integrating the result of the three tests.

3) Recovery of Bad Data

The maximum relative errors of the identified results based on the improved MPM before and after the recovery of bad data are illustrated in Fig. 11. For the three tests, before the recovery of bad data, the relative errors are below 4% when r is less than the threshold value of 2.5. In contrast, the relative error reaches more than 400% when r is much greater than 2.5. Then, the recovery of bad data is implemented for those whose r is more significant than 2.5. The recovery accuracy criterion is set to be $r < 2$ or

$\max(\varepsilon(k)) < 0.5$ for the balance between the recovery accuracy and the convergence of algorithms. After the recovery of bad data, the relative errors are all below 5%, which demonstrates that the threshold value of 2.5 is reasonable in the above three tests, and the bad data can be solved by noise reduction when r is smaller than the threshold value and can be identified and recovered when r is greater than the threshold value.
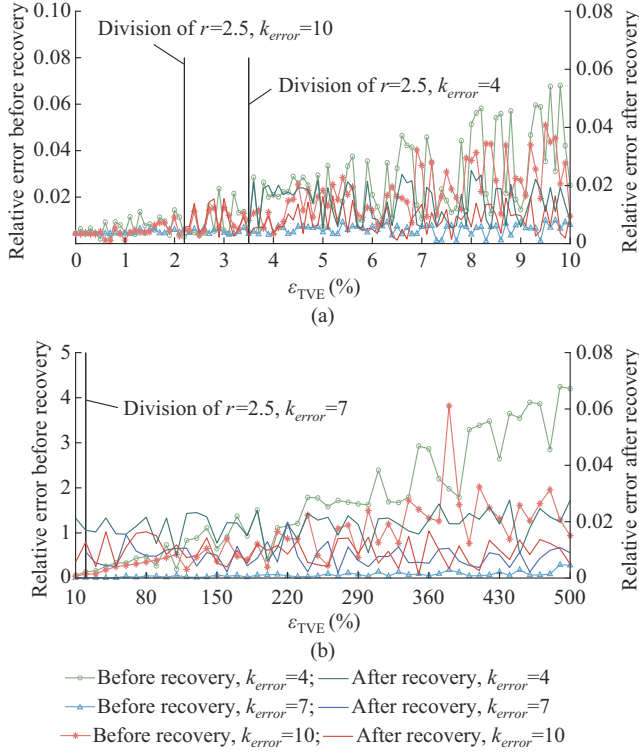

Fig. 11.   The maximum relative errors before and after recovery of bad data. (a) $\varepsilon_{TVE} \in [0, 10\%]$. (b) $\varepsilon_{TVE} \in [10\%, 500\%]$.

The characteristics of r may vary slightly in other scenarios such as different parameters of PMU data or different positions of bad data, in which misclassification or omission of bad data may occur when r is close to the threshold value. However, as stated in Section V-B, the threshold value does not need to be solved exactly. The threshold value of 2.5 is robust enough to ensure that the bad data can be solved by either noise reduction or recovery when r is close to 2.5 and thus can also be utilized in other scenarios.

Consequently, the proposed identification and recovery of bad data can significantly reduce the impact of bad data and guarantee the performance of the MPM effectively.

VI.   CONCLUSION

An improved SSO parameter identification based on MPM is proposed by utilizing the high computing efficiency and noise-resilient features of MPM to realize the dynamic monitoring of SSO. The MPM is enhanced by the angular frequency fitting equations based on the characteristic polynomial coefficients of the matrix pencil to ensure the accuracy of the identified parameters, since the existing eigenvalue estimation of MPM ignores the angular frequency conjuga-

tion constraints of the two fundamental modes and two oscillation modes. Then, the MPM is enhanced by the identification and recovery of bad data by utilizing the difference in temporal continuity of the synchrophasors before and after noise reduction. The proposed improved MPM is verified with synthetic, simulated, and actual measured PMU data. Compared with the existing MPM, the improved MPM achieves better accuracy for parameter identification of fundamental, subsynchronous, and supersynchronous components with a 200-ms data window. The improvements in the identification and recovery of bad data can significantly reduce the impact of bad data, thus enhancing the practicability of MPM.

REFERENCES

[1] L. Wang, X. Xie, Q. Jiang *et al.*, "Investigation of SSR in practical DFIG-based wind farms connected to a series-compensated power system," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2772-2779, Feb. 2015.

[2] C. Lin, "Simplified modelling of oscillation mode for wind power systems," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 1, pp. 56-67, Jan. 2021.

[3] H. Liu, T. Bi, X. Chang *et al.*, "Impacts of subsynchronous and supersynchronous frequency components on synchrophasor measurements," *Journal of Modern Power Systems and Clean Energy*, vol. 4, no. 3, pp. 362-369, Jul. 2016.

[4] L. Kilgore, D. Ramey, and M. Hall, "Simplified transmission and generation system analysis procedures for subsynchronous resonance problems," *IEEE Transactions on Power Apparatus and systems*, vol. 96, no. 6, pp. 1840-1846, Nov. 1977.

[5] Y. Li, L. Fan, and Z. Miao, "Replicating real-world wind farm SSR events," *IEEE Transactions on Power Delivery*, vol. 35, no. 1, pp. 339-348, Feb. 2020.

[6] X. Xie, Y. Zhan, H. Liu *et al.*, "Wide-area monitoring and early-warning of subsynchronous oscillation in power systems with high-penetration of renewables," *International Journal of Electrical Power & Energy Systems*, vol. 108, pp. 31-39, Jun. 2019.

[7] J. Shair, X. Xie, and G. Yan, "Mitigating subsynchronous control interaction in wind power systems: existing techniques and open challenges," *Renewable and Sustainable Energy Reviews*, vol. 108, pp. 330-346, Apr. 2019.

[8] J. Shair, X. Xie, J. Yang *et al.*, "Adaptive damping control of subsynchronous oscillation in DFIG-based wind farms connected to series-compensated network," *IEEE Transactions on Power Delivery*, vol. 37, no. 2, pp. 1036-1049, Apr. 2022.

[9] H. Khalilinia and V. Venkatasubramanian, "Subsynchronous resonance monitoring using ambient high speed sensor data," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1073-1083, Mar. 2016.

[10] X. Xie, H. Liu, Y. Wang *et al.*, "Measurement of sub-and supersynchronous phasors in power systems with high penetration of renewables," in *Proceedings of 2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Minneapolis, USA, Dec. 2016 , pp. 1-5.

[11] H. Lin, "Power harmonics and interharmonics measurement using recursive group-harmonic power minimizing algorithm," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 2, pp. 1184-1193, Feb. 2012.

[12] N. Zhou, D. J. Trudnowski, J. W. Pierre *et al.*, "Electromechanical mode online estimation using regularized robust RLS methods," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1670-1680, Nov. 2008.

[13] T. Rauhala, A. M. Gole, and P. Järventausta, "Detection of subsynchronous torsional oscillation frequencies using phasor measurement," *IEEE Transactions on Power Delivery*, vol. 31, no. 1, pp. 11-19, Feb. 2016.

[14] P. Wall, P. Dattaray, Z. Jin *et al.*, "Deployment and demonstration of wide area monitoring system in power system of Great Britain," *Journal of Modern Power Systems and Clean Energy*, vol. 4, no. 3, pp. 506-518, Jul. 2016.

[15] F. Zhang, L. Cheng, W. Gao *et al.*, "Synchrophasors-based identification for subsynchronous oscillations in power systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2224-2233, Mar. 2019.

[16] X. Yang, J. Zhang, X. Xie *et al*., "Interpolated DFT-based identification of sub-synchronous oscillation parameters using synchrophasor data," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2662-2675, May 2020.

[17] J. F. Hauer, C. Demeure, and L. Scharf, "Initial results in prony analysis of power system response signals," *IEEE Transactions on Power Systems*, vol. 5, no. 1, pp. 80-89, Feb. 1990.

[18] M. Netto and L. Mili, "A robust prony method for power system electromechanical modes identification," in *Proceedings of 2017 IEEE PES General Meeting*, Chicago, USA, Jul. 2017, pp. 1-5.

[19] M. L. Crow and A. Singh, "The matrix pencil for power system modal extraction," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 501-502, Feb. 2005.

[20] Y. Wang, X. Jiang, X. Xie *et al*., "Identifying sources of subsynchronous resonance using wide-area phasor measurements," *IEEE Transactions on Power Delivery*, vol. 36, no. 5, pp. 3242-3254, Oct. 2021.

[21] H. Liu, Y. Qi, J. Zhao *et al*., "Data-driven subsynchronous oscillation identification using field synchrophasor measurements," *IEEE Transactions on Power Delivery*, vol. 37, no. 1, pp. 165-175, Feb. 2022.

[22] Y. Zhu, C. Liu, and K. Sun, "Image embedding of pmu data for deep learning towards transient disturbance classification," in *Proceedings of 2018 IEEE International Conference on Energy Internet (ICEI)*, Beijing, China, May 2018, pp. 169-174.

[23] F. Zhang, J. Li, J. Liu *et al*., "An improved interpolated dft-based parameter identification for sub-/super-synchronous oscillations with synchrophasors," *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1741-1727, Mar. 2022.

[24] *IEEE Standard for Synchrophasor Measurements for Power Systems*. IEEE Std C37.118.1-2011, Dec. 2011.

Xiaoxue Zhang received the B.S. and M.S. degrees from China Agricultural University, Beijing, China, in 2014 and 2016, respectively. She is currently pursuing the Ph.D. degree with the Laboratory of Power System Protection and Control Research at Beijing Jiaotong University, Beijing, China. Her research interests include power system dynamic measurement and control based on wide area measurement system (WAMS).

Fang Zhang received the B.S. and Ph.D. degrees from Tsinghua University, Beijing, China, in 2010 and 2015, respectively. Currently, he is working as an Associate Professor with the School of Electrical Engineering, Beijing Jiaotong University, Beijing, China. His research interests include power system dynamic measurement and control based on WAMS, optimal operation and control of multi-energy systems and transportation energy systems.

Wenzhong Gao received the M.S. and Ph.D. degrees from Georgia Institute of Technology, Atlanta, USA, in 1999 and 2002, respectively. He is currently working as a Professor with the Department of Electrical and Computer Engineering, University of Denver, Denver, USA. His current research interests include renewable energy, smart microgrid, and WAMS.

Jinghan He received the B.S. and M.S. degrees from Tianjin University, Tianjin, China, in 1987 and 1994, respectively, and the Ph.D. degree from Beijing Jiaotong University, Beijing, China, in 2007. She is currently a Professor with the School of Electrical Engineering, Beijing Jiaotong University, Beijing, China. Her main research interests include protective relaying, fault distance measurement, and location in power systems.