Impact of PMU Data Errors on Modal Extraction Using Matrix Pencil Method

Tamara Becejac and Thomas Overbye
Department of the Electrical and Computer Engineering
Texas A&M University
College Station, Texas, USA
tamara.becejac@tamu.edu; overbye@tamu.edu

***Abstract*— Phasor Measurement Units (PMUs) provide information on the dynamic response of power systems that can be used to extract modal information and determine the mode shape of electromagnetic oscillations in real-time. Modal analysis introduces a significant improvement in determining and assessing stability margins and provides insights into the dynamics of the power system for monitoring and preventive control. Performance of modal analysis depends primarily on the availability and quality of the time-domain signals. However, PMU data is often delivered with various data quality issues. This paper examines the impact of different data quality issues on the performance of the Matrix Pencil Method (MPM), which has been used for power system modal analysis. The paper also explores the ability of the MPM to accurately identify dominant system modes in presence of defective data, including noise, outliers, and un-updated data. The results are illustrated through six examples ranging in size from a single signal up to results from a 2000-bus system.**

***Index Terms*— Flawed data, Matrix Pencil Method, sensitivity analysis, singular value threshold, Phasor Measurement Unit.**

I. Introduction

In power systems, small signal stability problems continue to be a concern for power engineers [1], [2]. Power system disturbances are usually followed by low-frequency oscillations that can decay, sustain, or grow. Additionally, due to a variety of reasons many power systems are vulnerable to stability issues including inter-area oscillations. Inter-area oscillations occur when a group of electrically distant generators oscillates against each other [3]. From an operating point of view, oscillations are acceptable as long as they are quickly damped. In well-damped systems these oscillations will be absorbed within a few seconds. However, sometimes they may lead to instability and system collapse [4]. Lightly damped electromechanical oscillations are also of concern in the power industry due to their undesirable effect on economics and operational practices. Additionally, even low-magnitude oscillations, if present over an extended period, can cause substantial damage to power system equipment such as generator shafts [5]. Thus, it is essential to detect these oscillations in the early stages of an event to be able to quickly take appropriate preventive controls.

Modal analysis is a mathematical tool used to study and characterize the small signal stability of a power system in the frequency domain [6]. Modal analysis directly extracts modal information from the system response to a perturbation including the mode's magnitude, phase, frequency, and damping factor. Although oscillations cannot be completely eliminated, advanced controls can be added to the system to improve their damping and modify their frequency, assuming that the modes of the system are correctly identified.

Two main approaches for estimating modal content exist in the power system, 1) the traditional method based on the linearization of power system equations, and 2) the identification approach based on computational techniques that extract the modal information from time domain data. With the deployment of the modern devices that are capable of capturing the dynamic responses of the system, such as PMUs and Digital Fault Recorders (DFRs), an oscillatory response of the system can be measured in real-time. The main idea behind the identification methods is to decompose the system dynamic response into a sum of exponentially damped modes. A significant advantage of the identification methods is that their feasibility does not depend on having a system model [7]. Moreover, as the power grid changes, such as due to load perturbations or transmission switching actions, the modes of the power system also change.

If accurately estimated, system modes may help improve system modeling and the implementation of the control schemes, with a result of increased system reliability. However, PMU data is often delivered with various data quality issues. Data issues in the synchrophasor system can arise for a wide variety of reasons including topology errors, meter failures, and improper configurations. Furthermore, the communication infrastructure brings additional vulnerabilities to the power system, such as allowing intruders to launch various types of cyber-attacks and jeopardize the reliability of the electric grid.

There exist several identification techniques from the signal processing theory for estimating the modes from time-varying waveforms including Prony, Hankel Total Least Square, Matrix Pencil. Prony's method was the first method used for power system small signal stability analyses [8]. Nonetheless, several studies have shown some advantages from using the Matrix Pencil Method for the modal extraction from noisy signals over the traditional Prony method [9]. We are building on these studies to explore the advantages of MPM under not only the presence of noise, but also other issues that can arise in the

978-1-7281-1607-5/19/$31.00 ©2019 IEEE

PMU data measurement set, such as outliers and un-updated data.

The rest of this paper is organized as follows: Section II gives the historical overview of the identification–type modal analyses, Section III reviews the theoretical background of the Matrix Pencil Method, and Section IV documents six case studies. Finally, concluding remarks are given in Section V.

II. Modal Analysis Background

Modal analysis can be formulated as a fitting problem of the pre-specified waveform to an actual time-varying waveform, such that the difference between the proposed and the actual waveform is minimized when estimating the magnitude, phase, frequency, and damping parameters of the fitting function. Estimated coefficients of the modeled waveform determine the modal characteristic of the linear system given as:

$$y(t) \approx \sum_{i=1}^{M} A_i e^{(\sigma_i t)} \cos(\omega_i t + \theta_i) + n(t);\ 0 \le t \le T \tag{1}$$

where the following nomenclature applies:

$y(t)$	= observed signal
$n(t)$	= system noise
M	= number of desired modes
$R_i = A_i e^{i\theta_i}$	= complex amplitude
σ_i	= damping factor
ω_i	= angular frequency

The key idea is to model the uniformly sampled data as a linear combination of the exponentially dumped functions. There are several well-known methods used for modal identification: Prony's method, Hankel Total Least Square (HTLS), and Matrix Pencil Method. From a historical perspective, the polynomial type of method (i.e., Prony's method) is much older and is still widely used for the stability studies. Baron de Prony first developed this method in 1795 [10]. HTLS [11] is a special case of the ESPRIT algorithm [12]. Hua and Sarkar introduced Matrix Pencil Method (MPM) for extracting poles from antennas' transient responses [13]-[15]. MPM was derived from the pencil-of-functions approach [16], and today it is widely applied in the system identification and spectrum estimation areas. As an extension of the MPM, the Iterative Matrix Pencil Method was recently introduced for the modal analyses of large-scale systems [17]. The Iterative MPM combines the MPM and the cost functions to decrease the computational burden by reducing the number of signals that are processed during the modal extraction. Other methods include variable projection method [18], eigensystem realization method [19], and the dynamic mode decomposition [20].

Unlike the polynomial methods, MPM shows some advantages when dealing with signals with unknown characteristics [21]. Moreover, it has excellent performance even when the input signal has noise, which is usually the case with PMU data [22]. Typically, up to 20-25 dB of signal-to-noise ratio (SNR) can be handled adequately by this technique [15]. The Matrix Pencil approach has a lower variance of estimates of the parameters of interest than a polynomial-type method and is computationally more efficient. While the MPM extracts signal poles directly from the eigenvalues, the Prony method requires a two-step process [3].

In this paper, the MPM is considered. This technique belongs to the group of identification methods and hence it can be applied to time-series data (PMU frequency). One of the main limitations of the identification method is the type of modes that can be computed (i.e., only modes that are present in the input data), and the inability to extract some well-damped modes if the input signals are noisy. Regardless, their practical value has been demonstrated in many studies, such as stability analysis, model validation, and control design [7].

III. Matrix Pencil Method

With the MPM signal poles can be found directly from the eigenvalues of a single developed matrix. This method directly estimates all mode parameters by fitting the signal model given in (2) to an observed measurement (1) that consists of N evenly spaced samples.

$$\hat{y}(k) = \sum_{i=1}^{M} R_i z_i^k + n(k); \qquad k = 0, \dots, N-1 \tag{2}$$

where $z_i = e^{(\sigma_i + j\omega_i)\Delta t}$ are the poles of the system.

Parameters A_i and θ_i are calculated from the complex residues, whereas σ_i and ω_i are estimated from the eigenvalues of the system poles. Since the poles z_i are found as the solution of a generalized eigenvalue problem, the limitation on the number of modes M that can be extracted is removed. To calculate the eigenvalues and eigenvectors of the signal, from the sampling sequence given in (2), the first step is to form a $(N\text{-}L) \times (L+1)$ Hankel matrix:

$$[Y] = \begin{bmatrix} y(0) & y(1) & \cdots & y(L) \\ y(1) & y(2) & \cdots & y(L+1) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-L-1) & y(N-L) & \cdots & y(N-1) \end{bmatrix}_{(N-L)\times(L+1)} \tag{3}$$

where L is called the pencil parameter, representing the free-moving window length. It has been shown that the right choice of the pencil parameter helps when dealing with noisy signals [15]. Performance of the method is maximized if L is selected to be $L=N/2$, making the MPM perform close to the optimal Cramer-Rao bound [13] (the Cramer-Rao bound represents the best results that one can achieve in a noisy environment).

Furthermore, two additional matrices $[Y_1]$ and $[Y_2]$ are constructed from $[Y]$ by deleting its last and first columns, respectively.

$$[Y_1] = \begin{bmatrix} y(0) & y(1) & \cdots & y(L-1) \\ y(1) & y(2) & \cdots & y(L) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-L-1) & y(N-L) & \cdots & y(N-2) \end{bmatrix}_{(N-L)\times L} \tag{4}$$

$$[Y_2]=\begin{bmatrix} y(1) & y(2) & \cdots & y(L) \\ y(2) & y(3) & \cdots & y(L+1) \\ \vdots & \vdots & \ddots & \vdots \\ y(N-L) & y(N-L+1) & \cdots & y(N-1) \end{bmatrix}_{(N-L)\times L} \tag{5}$$

Now, from the equations (2), (4) and (5) we can write:

$$[Y_1]=[Z_L][R][Z_R] \tag{6}$$

$$[Y_2]=[Z_L][R][Z_0][Z_R] \tag{7}$$

where

$$[R]=diag[R_1,R_2,...,R_M] \tag{8}$$

$$[Z_0]=diag[z_1,z_2,...,z_M] \tag{9}$$

$$[Z_L]=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_M \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{(N-L-1)} & z_2^{(N-L-1)} & \cdots & z_M^{(N-L-1)} \end{bmatrix}_{(N-L)\times M} \tag{10}$$

$$[Z_R]=\begin{bmatrix} 1 & z_1 & \cdots & z_1^{(L-1)} \\ 1 & z_2 & \cdots & z_2^{(L-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_M & \cdots & z_M^{(L-1)} \end{bmatrix}_{M\times L} \tag{11}$$

diag[] represents the diagonal matrix.

Next, we defined the matrix pencil as given in (12):

$$[Y_2]-\lambda[Y_1]=[Z_L][R]\{[Z_0]-\lambda[I]\}[Z_R] \tag{12}$$

where [*I*] is an identity matrix.

If the pencil parameter L satisfies $M\le L\le N-M$, then the rank of $[Y_2]-\lambda[Y_1]$ is equal to M [13]. On the other hand, if $\lambda=z_i$, the ith row of $[Y_2]-\lambda[Y_1]$ is equal to zero, and the rank decreases to $M-1$. Therefore z_i is the generalized eigenvalue of the matrix pencil pair $\{[Y_2], [Y_1]\}$. To find poles, one should find the eigenvalues of the:

$$[Y_1]^{+}[Y_2]=\left\{[Y_1]^{H}[Y_1]\right\}^{-1}[Y_1]^{H}[Y_2] \tag{13}$$

The superscript "+" denotes the Moore-Pensore inverse or pseudoinverse matrix, superscript "H" stand for Hermitian (conjugate) transpose, and "-1" denotes inverse.

After solving (13) and obtaining z_i and M, the residues R_i can be calculated from the least square problem as given in (14);

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix}=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_M \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{(N-1)} & z_2^{(N-1)} & \cdots & z_M^{(N-1)} \end{bmatrix}_{N\times M}\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_M \end{bmatrix} \tag{14}$$

Finally, we can get the amplitudes, phases, frequencies, and damping factors as follows:

$$A_i=|R_i| \tag{15}$$

$$\theta_i=\arctan\left[\mathrm{Im}(R_i)/\mathrm{Re}(R_i)\right] \tag{16}$$

$$\omega_i=\arctan\left[\mathrm{Im}(z_i)/\mathrm{Re}(z_i)\right]/\Delta t \tag{17}$$

$$\sigma_i=\ln|z_i|/\Delta t \tag{18}$$

The damping ratio can be calculated using the following equation:

$$\zeta_i=\frac{1}{\sqrt{1+\left(\frac{2\pi}{\sigma_i}\right)^2}}\frac{\omega_i}{2\pi}\times 100\% \tag{19}$$

In the case of data that contains noise before estimating the desired number of poles, the sampled sequence should be conditioned to reduce the impact of the input noise. This can be achieved by applying the Singular Value Decomposition (SVD) of the matrix in (3). The SVD is a method for factorization of matrices of the form:

$$[Y]=[U][\Sigma][V]^{H} \tag{20}$$

where $[\Sigma]$ is the rectangular diagonal matrix. The diagonal matrix elements σ_i, are the singular values of the matrix [*Y*]. The columns of [*U*] and [*V*] are unitary matrices containing the left-singular vectors and right-singular vectors of [*Y*], respectively. The SVD reduces the dimension of the data set while preserving the primary relationships within the set and ignoring the variations below some predefined threshold.

The Matrix Pencil decomposes [*Y*] with the SVD with the result being the poles and residuals. In this case, the SVD is also used to determine the number of modes (model order) M, in the observed signal. In the noiseless case, M is equal to the number of non-zero singular values. However, in the presence of noise, singular values that were previously zero are perturbed and become non-zero. Now, the best way to approximate M is to look into the ratio of singular values of [*Y*], σ_i, to the largest singular value, σ_{max}, and compare to predefined Singular Value Threshold (SVT), defined as $SVT=10^{-p}$, where p represents the number of significant decimal digits. In this case, all singular values with the ratio below the threshold are considered noise and discarded. This step is reflected as a built-in pre-filtering process [15].

$$\frac{\sigma_i}{\sigma_{\max}}\approx SVT \tag{21}$$

Next, we can use the first M significant singular values to reconstruct the original matrix [*Y*], and the pair $\{[Y_1], [Y_2]\}$:

$$\begin{aligned} [V_s]&=[v_1,v_2,...,v_M] \\ [Y_1]&=[U_s][\Sigma_s][V_{1s}]^{H} \\ [Y_2]&=[U_s][\Sigma_s][V_{2s}]^{H} \end{aligned} \tag{22}$$

where $[\Sigma_s]$ is the first M columns of $[\Sigma]$, and $[V_{1s}]$ and $[V_{2s}]$ are obtained from $[V_s]$ by deleting its last and first rows respectively. Using the set of equations from (13)-(18), all desired shapes and parameters can be calculated.

IV. CASE STUDY

The impact of the synchrophasors errors on the performance of the MPM is demonstrated through two examples:

1) Artificially created single data sequence with known modal parameters.
2) The transient stability simulation of a 2000-bus synthetic grid [23].

In both examples, we explored the sensitivity of the algorithm to noise, outliers present in the data sequence, and the effect of un-updated data. The numerical results verify the effectiveness of this approach.

A. Simulated signal decomposition, without noise

First the signal given in (23) is used to generate an example data sequence with known modes:

$$y(t)=e^{-0.075t}\cos\left(2\pi*0.15t\right)+e^{-0.15t}\cos\left(2\pi*0.35t\right)+n(t) \quad (23)$$

where *n(t)* represents any flawed data (i.e., noise, outliers, and un-updated data), *y(t)* contains 300 points, and the signal is sampled at a rate of 30 Hz (a total of 10 seconds of sampled data is used for the analysis). This signal clearly has two modes at 0.15 Hz and 0.35 Hz with damping ratios of 7.96% and 6.82%, respectively. Fig. 1 shows the original clean data curve.

In order to apply the MPM to this data set, the pencil parameter L and the Singular Value Threshold need to be defined first. Values for the pencil parameter between $N/3$ and $2N/3$ are preferred, where N is the number of data points [15]. The fixed value of L in this study is calculated as follows:

$$L=ceil\left(1/2*\left(ceil\left(N/3\right)+floor\left(N/2\right)\right)\right) \quad (24)$$

The MPM is very sensitive to the value of the SVT. For this example, all values in the range 0<SVT<0.692 allow correct extraction of the modes and its damping ratios. However, if the SVT is 0.693 then the MPM detects only one mode with a frequency of 0.147 Hz, which is relatively close to the original value of 0.15 Hz. However, the damping ratio is underestimated, with the values of 4.67% compared to 7.96%. Thus, the SVT plays a critical role in a case that signal is not "clean". In the following analysis, the optimal value of the parameter SVT is determined in the presence of noise and other flawed data.

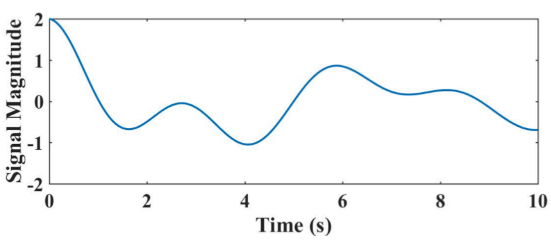

Figure 1. Original simulated signal.

1) Decomposition of the simulated signal with noise

In order to determine the optimal value of the SVT, white Gaussian random noise is added to the original synthetic signal with four values of the signal-to-noise ratio (SNR) ranging from 10-40dB. We considered the number of estimated modes and the mean error between the original waveform without noise and one reconstructed from the modes calculated with the PM algorithm. One hundred simulation runs for each test case were done and Fig. 2 shows the results from these analyses. The "stars" represent the value of the SVT for which the MPM correctly calculated the number of modes present in the original signal. The findings are summarized in Table I.

The SVT max value mostly depends on the signal content while the SVT min value controls the process of removing the noise. With increasing the SVT, more noise is removed from the signal, and the reconstructed waveforms better fit the

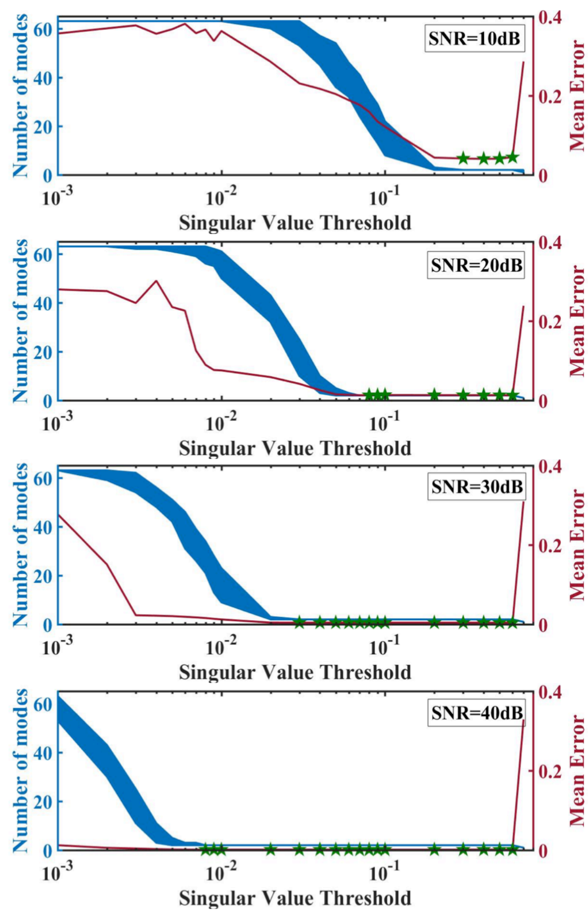

Figure 2. Number of modes and mean error of the reconstructed signal with four different values of additive noise.

TABLE I. LIMITS OF THE SINGULAR VALUE THRESHOLD IN THE PRESENCE OF NOISE

SNR (dB)	Singular Value Threshold	
	min	*max*
10	0.3	0.7
20	0.07	0.7
30	0.03	0.7
40	0.008	0.7

original, giving a highly accurate estimation of the mode parameters. However, a large value for the SVT could remove actual existing modes, so one has to be careful with the choice of the threshold. Typical noise level present in PMU data is around 40 dB [22], which means that parameter SVT can take value as low as 0.008. To accommodate for the uncertainties and possible higher noise, we fixed the threshold value to 0.05.

2) Decomposition of the simulated signal with arbitrary data injection

Here the tested data sequence is modified to contain data spikes lasting for a total of 0.13s in samples 172 to 175. These spikes could occur in real PMU data due to either time skew in the GPS or a cyber-attack [26]. Once again, we sweep the values of the SVT with the results of the analyses presented in Fig 3. The calculated mode parameters are given in Table II.

When the SVT is equal to 0.05 the number of estimated modes is 13, much higher than the original. Nevertheless, both “real” mode frequencies are calculated with a maximum error of 0.01 Hz. The damping ratio is overestimated for both modes with the maximum error equal to 0.77%. Decreasing the value of SVT to 0.06 gives the exact number of modes and a slightly better estimate of the damping ratio. However, if the SVT decreases to 0.01, the number of modes is 63, and the error is much higher.

3) Decomposition of the simulated signal with un-updated data

Next, we added a 0.5s long sequence of un-updated data from 6.5s to 7s with the results shown in Fig 4. For all values of the SVT higher than 0.05, the MPM was able to correctly estimate the number of modes with the results given in Table III.

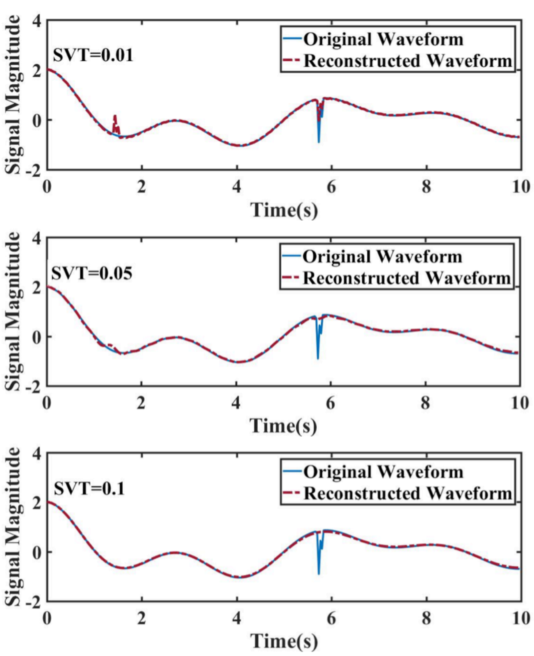

Figure3. Simulated and reconstructed waveforms in the presence of data spikes.

TABLE II. ESTIMATED MODE PARAMETERS IN THE PRESENCE OF ARBITATY DATA INJECTION

Singular Value Threshold=0.05			
Frequency (Hz)	**Damping Ratio (%)**	**Lambda**	**Amplitude**
0.149	**8.32**	**-0.0779**	**0.992**
2.845	0.67	-0.1203	0.016
2.614	0.91	-0.1488	0.019
2.388	1.10	-0.1655	0.002
0.350	**7.59**	**-0.1671**	**1.022**
0.593	4.55	-0.1697	0.021
2.162	1.30	-0.1761	0.018
1.936	1.50	-0.1832	0.014
1.711	1.75	-0.1879	0.012
1.485	2.05	-0.1912	0.019
0.811	3.75	-0.1913	0.011
1.259	2.44	-0.1933	0.001
1.035	2.98	-0.194	0.020
Singular Value Threshold≥0.06			
0.149	8.24	-0.0772	0.988
0.350	7.53	-0.1659	1.0156

TABLE III. LIMITS OF THE SINGULAR VALUE THRESHOLD IN THE PRESENCE OF UN-UPDATED DATA

Singular Value Threshold≥0.05			
Frequency (Hz)	**Damping Ratio (%)**	**Lambda**	**Amplitude**
0.150	6.58	-0.0621	0.978
0.347	7.41	-0.1616	1.033

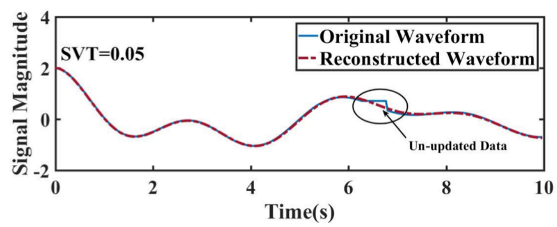

Figure 4. Simulated and reconstructed waveforms with the un-updated data

B. Large Synthetic Grid

A 2000-bus network that is publicly available at [25] is used to generate the synthetic synchrophasor measurements. Figure 5 shows 500kV part of the transmission network along with the 13 synchrophasors located in the substations with the large generators. The waveforms used for this study are from the virtual PMU frequency measurements in response to the loss of generator in the substation “MIAMI”. The sampling frequency is initially set to 0.5 cycles, and data is later re-sampled with the factor four, to get the reporting rate of 30 frames per second. The simulation duration is 20s, producing 600 points per PMU channel. Figure 6 shows the original measurements, where each curve represents the frequency measurement acquired from one of the PMUs in the system. The authors in [24] developed the approach to tune the model dynamics, so that the synthetic data created using this fictitious network has the characteristics of actual networks.

To apply the MPM on the synthetic PMU data set, the following parameters for the algorithm are used: pencil

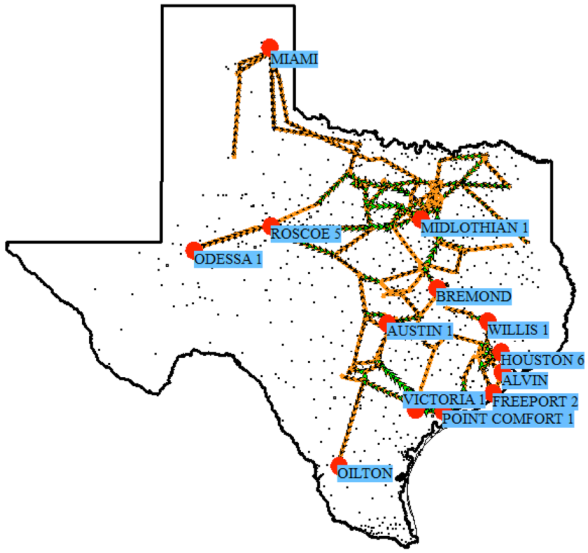

Figure 5. 2000 bus system with 13 virtual PMUs.

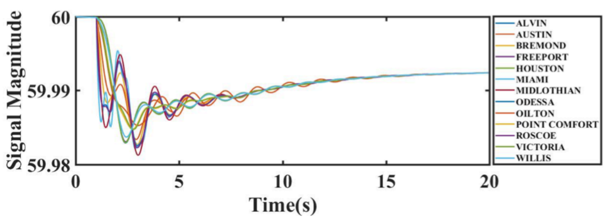

Figure 6. Original clean frequency measurements from 13 PMUs.

parameter of L=250 is calculated using the (23) and the SVT is initialized to 0.05. Furthermore, the data set is pre-conditioned. That is, from each PMU channel, we first subtracted the trend and then scaled it by its respective standard deviation.

In general, the MPM is designed to process single signals. However, with small adjustments it can be used on the multi-channel data set. First, the Henkel matrix as given in (3) for each signal is constructed and at the end, all matrices are vertically concatenated to form one single matrix. Once the modes are estimated, (14)-(19) are used to evaluate the mode shapes and frequencies in each channel separately.

To verify the finding from the first set of case studies regarding the optimal value of SVT, additional sensitivity analysis was conducted, now on the synchrophasor frequency measurements, and the best results were achieved with the value of SVT=0.05. Figure 7 shows the solution of the MPM when the SVT is 0.05 and the estimated parameters are given in Table IV, including dominant modes frequencies, damping ratios, and lambdas. The tested case has seven dominant

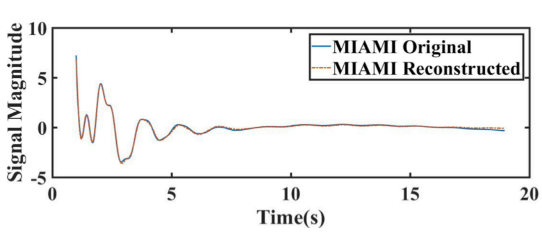

Figure 7. Clean frequency measurement and reconstructed signal.

TABLE IV. MODAL PARAMETERS ESTIMATED FROM THE 13 PMU FREQUENCY MEASUREMENTS

Singular Value Threshold≥0.05		
Frequency (Hz)	**Damping Ratio (%)**	**Lambda**
0.061	24.55	-0.094
0.223	35.30	-0.496
0.329	34.36	-0.716
0.633	9.04	-0.360
0.660	7.96	-0.330
0.942	6.86	-0.407
1.828	8.78	-1.022

modes, and these results were utilized as the reference point for the following analyses. In the next three case studies low-quality measurements were introduced demonstrate the effectiveness of the MPM. In each case, one type of flawed PMU data has been randomly inserted into the data set.

1) Synchrophasor data with noise

First, Gaussian random noise with SNRs ranging from 40 to 20 dB were added to the original clean PMU data set. Figure 8 shows one data curve with the applied noise with SNR=20 dB and the reconstructed waveform using the results from the MPM algorithm. Detail results are presented in Table V.

As expected, the MPM could effectively extract all the modes with high accuracy despite the additive noise. It can be seen that this algorithm shows great resistance to the noise and successfully identifies all the dominant modes. The maximum frequency error is 0.003 Hz and the damping ratio error is 1.2%.

TABLE V. MODAL PARAMETERS ESTIMATES FROM 13 PMU FREQUENCY MEASUREMENTS WITH NOISE AND RESPECTIVE ERRORS

SNR (dB)	Frequency (Hz)	Frequency Error (Hz)	Damping Ratio (%)	Damping Ratio Error
40	0.061	0	24.48	-0.075
	0.223	-0.0002	35.21	-0.090
	0.329	-0.0003	34.50	0.147
	0.633	0.0002	9.04	-0.002
	0.659	-0.0004	7.96	0.001
	0.943	0.0005	6.84	-0.016
	1.828	-0.0001	8.77	-0.012
30	0.061	0	24.40	-0.148
	0.222	-0.0011	36.50	1.200
	0.331	0.0016	33.81	-0.549
	0.633	0.0000	8.82	-0.216
	0.659	-0.0005	8.19	0.234
	0.945	0.0028	6.80	-0.056
	1.827	-0.0010	8.80	0.014
20	0.061	0.0001	24.90	0.346
	0.222	-0.0014	36.38	1.076
	0.334	0.0044	33.60	-0.762
	0.632	-0.0009	8.52	-0.519
	0.661	0.0015	8.61	0.656
	0.946	0.0033	6.72	-0.139
	1.832	0.0038	8.71	-0.078

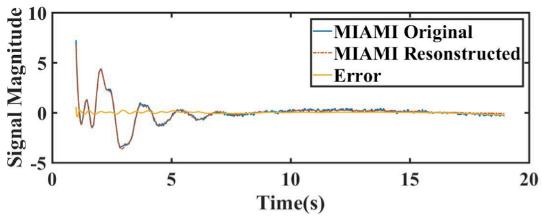

Figure 8. Noisy frequency measurement and the reconstructed signal.

2) Synchrophasor data with outliers

Next, the test data set of the 13 synthetic PMU frequency measurements was modified with small constant values of random "data spikes" added to all of them. The length of each corrupted data segment is 0.13s. Figure 9 shows the reconstructed waveform with the presence of spike; the results are summarized in the Table VI. It can been seen that the MPM is relatively immune to this type of flawed PMU data.

3) Synchrophasor data with un-updated data segments

In this test case a 0.5s duration of the tested sequence was modified to contain an un-updated segment for all 13 waveforms simultaneously. This type of error can arise due to the issues in PDC or the communication traffic. Figure 10 shows one of the original measurement curves along with the reconstructed one. Mode parameters are given in Table VII. Even though all the real modes are detected, in this test case, there is one additional "fake" mode extracted with the low damping ratio. This leads to the conclusion that some pre-screening of PMU data should take place, to avoid false detection.

V. CONCLUSIONS

The applicability of the MPM for the modal extraction from the time-domain data sequence was examined, and results were presented through two case examples. In this study, we found the optimal value of the SVT for the proposed method, for which the MPM was able to accurately calculate the modes parameters in the presence of the flawed data, i.e., noise, outliers, and un-updated data. We also showed how the method could be applied on sets of multiple PMU streams, and we showed its ability to correctly detect and accurately characterize dominant modes after a major disturbance in the large-scale system.

TABLE VI. MODAL PARAMETERS ESTIMATES FROM 13 PMU FREQUENCY MEASUREMENTS WITH SPIKES AND RESPECTIVE ERRORS

Frequency (Hz)	Frequency Error (Hz)	Damping Ratio (%)	Damping Ratio Error
0.061	0.0003	21.12	-3.436
0.202	-0.0207	48.76	**13.456**
0.345	**0.0154**	23.97	-10.386
0.629	-0.0043	9.47	0.435
0.665	0.0053	7.19	-0.766
0.945	0.0025	5.87	-0.988
1.829	0.0010	7.98	-0.806

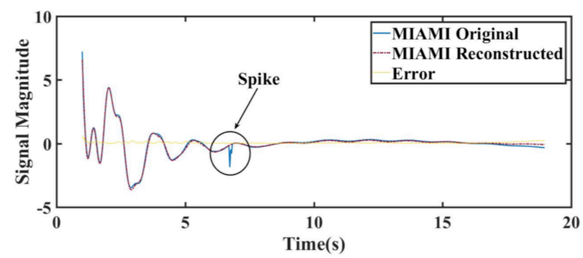

Figure 9. Frequency measurement with the spikes and the reconstructed signal.

TABLE VII. MODAL PARAMETERS ESTIMATES FROM 13 PMU FREQUENCY MEASUREMENTS WITH UN-UPDATED DATA AND RESPECTIVE ERRORS

Frequency (Hz)	Frequency Error (Hz)	Damping Ratio (%)	Damping Ratio Error
0.063	0.0015	25.16	0.613
0.197	-0.0262	53.43	**18.129**
0.369	**0.0394**	25.78	-8.574
0.639	0.0060	12.96	3.923
0.673	0.0134	5.73	-2.231
0.935	-0.0071	5.30	-1.560
1.162	**1.1618**	**2.13**	**2.134**
1.804	-0.0248	8.03	-0.755

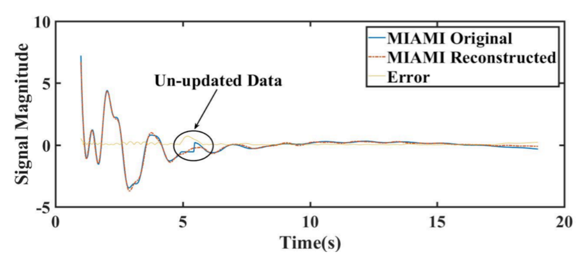

Figure 10. Frequency measurement with un-updated data and the reconstructed signal.

REFERENCES

[1] P. W. Sauer, M. A. Pai, J. H. Chow, "Small□Signal Stability," in *Power System Dynamics and Stability: With Synchrophasor Measurement and Power System Toolbox*, , IEEE, 2017, pp.

[2] G. Liu, J. Quintero and V. M. Venkatasubramanian, "Oscillation monitoring system based on wide area synchrophasors in power systems," *2007 iREP Symposium - Bulk Power System Dynamics and Control - VII. Revitalizing Operational Reliability*, Charleston, SC, 2007, pp. 1-13.

[3] M. Klein, G. J. Rogers, P. Kundur, "A Fundamental Study of Inter-Area Oscillations in Power Systems," *IEEE Trans. Power Systems*, vol. 6, pp. 914-921, Aug. 1991.

[4] V. Venkatasubramanian and Y. Li, "Analysis of 1996 western american electric blackouts", in *Proc. Bulk Power System Phenomena-Stability and Control*, Venice, Italy, 2004.

[5] P. Kundur, "Power System Stability and Control" in New York:McGraw-Hill, Inc., 1994.

[6] F. Dussaud "An application of modal analysis in electric power systems to study inter-area oscillations" in Degree Project In Electric Power System Second Level, STOCKHOLM, Sweden, 2015.

[7] J. Hauer, E. Martinez, A. R. Messina, and J. Sanchez-Gasca. "Identification of Electromechanical Modes in Power Systems." *Special Publication, TP462, IEEE Power & Energy Society, Tech. Rep.* 2012.

[8] J. F. Hauer, C. J. Demeure, L. L. Scharf, “Initial Results in Prony Analysis of Power System Response Signals”, *IEEE Trans. Power Systems*, vol. 5, pp. 80-89, Feb. 1990.

[9] L. L. Grant and M. L. Crow, “Comparison of Matrix Pencil and Prony methods for power system modal analysis of noisy signals,” *2011 North American Power Symposium*, Boston, MA, 2011, pp. 1-7.

[10] G. R. B. Prony, “Essai experimental et analytique sur les lois de la dilatalrlite de fluids elastiques et sur cells de la vapeur de l'alcool a differents tempoeratures,” *Journal de l'Ecole Polytechnique (París)*, vol. 1, pp. 24-76, 1795.

[11] S. Van Huffel, H. Chen, C. Decanniere, and P. Van Hecke, “Algorithm for time-domain NMR data fitting based on total least squares,” *J. Mag. Res*., vol. A110, pp. 228–237, 1994.

[12] R. Roy and T. Kailath, “ESPRIT – estimation of signal parameters via rotational invariance technique,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[13] Y. Hua and T. K. Sarkar, “Matrix pencil method and its performance,” *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*, New York, NY, USA, 1988, pp. 2476-2479 vol.4.

[14] Y. Hua and T. K. Sarkar, “Generalized pencil-of-function method for extracting poles of an EM system from its transient response,” in *IEEE Transactions on Antennas and Propagation*, vol. 37, no. 2, pp. 229-234, Feb. 1989.

[15] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 5, pp. 814-824, May 1990"

[16] V. K. Jain, T. K. Sarkar, “High-performance signal-modeling by Pencil-of-Functions method: band-pass case”, *IEEE Trans. on ASSP*, Aug. 1986.

[17] W. Trinh, K. S. Shetye, I. Idehen, and T. J. Overbye, “Iterative Matrix Pencil Method for Power System Modal Analysis,” 2019 *Hawaii International Conference on System Sciences (HICSS)*, Wailea, HI, January 2019., in press.

[18] A. R. Borden, B. C. Lesieutre, “Variable projection method for power system modal identification,” *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2613-2620, Nov. 2014.

[19] R. A. DeCallafon, B. Moaveni, J. P. Conte, X. He, E. Udd, “General realization algorithm for modal identification of linear dynamic systems”, *ASCE J. Eng. Mech.*, vol. 134, no. 9, pp. 712-722, 2009.

[20] S. Mohaptra, T. J. Overbye, “Fast modal identification monitoring and visualization for large-scale power systems using dynamic mode decomposition,” *Proc. 2016 Power Systems Computation Conference (PSCC)*, pp. 1-7.

[21] A. R. Borden, B. C. Lesieutre and J. Gronquist, “Power system modal analysis tool developed for industry use,”*2013 North American Power Symposium (NAPS)*, Manhattan, KS, 2013, pp. 1-6.

[22] M. Brown, M. Biswal, S. Brahma, S. J. Ranade and H. Cao, “Characterizing and quantifying noise in PMU data,” *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Boston, MA, 2016, pp. 1-5.

[23] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye and T. J. Overbye, "Grid Structural Characteristics as Validation Criteria for Synthetic Networks," in *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3258-3265, July 2017.

[24] T. Xu, A. B. Birchfield and T. J. Overbye, “Modeling, Tuning and Validating System Dynamics in Synthetic Electric Grids,” *in IEEE Transactions on Power Systems.*

[25] https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg2000/

[26] M. Wu and L. Xie, “Online Detection of Low-Quality Synchrophasor Measurements: A Data-Driven Approach,” in *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2817-2827, July 2017.