A Novel Method for Despiking Spectra from Synchrophasor Measurements

Marcelo de Castro
Rensselaer Polytechnic Institute
Troy, NY, USA
Email: decasm3@rpi.edu

Chetan Mishra
Dominion Energy
Richmond, VA, USA

Luigi Vanfretti
Rensselaer Polytechnic Institute
Troy, NY, USA

Kevin D. Jones
Dominion Energy
Richmond, VA, USA

***Abstract*—Periodic corrections to phase angle estimates are performed inside PMUs to compensate for the internal clock drift with respect to the GPS clock that PMUs experience. In the frequency domain, this results in narrow bandwidth peaks referred to as spikes. The amplitude of these spikes depends on the amount of correction introduced and, consequently, the spikes' amplitude can sometimes be larger than the actual power system's dynamic system response under ambient excitation. These spikes result in spurious poorly damped modes in data-driven small signal analysis. To address this issue, this work proposes a novel approach that removes these spikes from the phase angle signal. The proposed approach is based on a technique for baseline estimation combined with the Whitaker-Hayes method for despiking spectra. The effectiveness of the proposed approach is demonstrated on PMU data from Dominion Energy's substations with significant internal clock issues.**

I. Introduction

Synchrophasor measurement technology has a wide variety of applications in the power industry [1]. In fact, one of the most consolidated ones is the near real-time monitoring of small-signal electromechanical oscillations [2]. Electromechanical modes are usually studied using angle or frequency data. Furthermore, the study of oscillations using PMU data relies in signal processing techniques and different examples of such applications can be found in the literature [3]–[5].

The synchronized measurements from PMUs are not immune to data quality issues [2] [6]. In fact, it is common to implement procedures that replace defective data points to increase the quality of the data set to be analyzed. Different equipment issues inherent to PMUs can result in spurious measurement data, for example, frequency estimates produced by PMUs are susceptible to any failure that occurs in the angle data [2].

Because the system operates largely under "ambient conditions" [5], which are characterized by small, random perturbations around the equilibrium point, frequency domain techniques for small signal stability analysis have been exploited for PMU applications. This is because they offer the benefit of being robust to random noise while capable of identifying small-amplitude periodic behavior. However, applying these techniques becomes challenging when errors in the PMU data exhibit some sort of periodicity. One such situation arises from the periodic phase angle correction inside the synchrophasor device to compensate for the drift between the internal oscillator and the local GPS clock [6], [7]. This introduces spikes in the frequency domain that can dominate the ambient dynamics, as it will be illustrated in this paper.

Methods for removing spectral spikes coming from forced oscillations have already been presented in the literature [8], [9]. However, spikes coming from forced oscillations due to natural disturbances are wider and, therefore, constitute a different type of spike if compared to the ones studied here. On the other hand, methods addressing narrower spectral spikes are found in other fields of research. For instance, Raman Spectroscopy is a technique used to determine vibration modes of molecules and Raman spectral estimates are often corrupted by spectral spikes resulting from cosmic rays. Therefore, the identification and removal of spikes is a well studied problem in Chemistry. The existing methods can be classified into multi-scan or single-scan procedures. Multi-scan methods operate by comparing results between consecutive scans based on the assumption that the probability of same pixel experiencing a cosmic spike is sufficiently low. Among a vast number of approaches, the upper-bound spectrum method [10], second difference comparison [11], and time domain comparison [12] have received significant attention. However, in this paper the spectral spikes resulting from phase angle correction are not random, unlike the ones from cosmic rays. Hence, multi-scan techniques cannot be applied to our case. Nevertheless, single scan methods can be implemented effectively because they use filters and smoothing to remove spikes. Katsumoto *et al* [13] proposed the use of weighted moving window filtering, and in [14] a wavelet transform was used to identify and remove spikes. Meanwhile, [15] proposed fitting spikes with predefined profiles. Whitaker Hayes Filter [16] identifies spikes as outliers in the gradient of the spectrum owing to the spikes having narrow bandwidth followed by flattening them by averaging.

Overall, the above mentioned approaches are fairly effective when dealing with flat baseline. However, for power system phase angle spectrum, this is not the case, which results in scenarios where a spike is masked by a steepness in the spectral baseline thus rendering these approaches ineffective. Therefore, this paper proposes a novel method to despike PMU phase angle spectra by accounting for the non-flat baseline. In the proposed approach, the baseline is corrected and, then, the spikes are removed using the method presented in [16].

The paper is organized as follows: Section II describes the information about the synchrophasor data and what is the main issue related to angle data. Section III discusses the proposed approach while Section IV enumerates the results for a case study. Finally, Section V states the concluding remarks of this paper.

II. Synchrophasor Data and Spectral Spikes

A. Fundamentals

The synchrophasor can be defined as an estimation of an electrical quantity's phasor (e.g. voltage or current) that is associated with a time tag [1]. The phasor estimates provided by a PMU are calculated using signal processing techniques, which are applied to three-phase sampled-waveform measurements. After the data acquisition and estimation stage, the estimated phasors are given a time-tag related to a time-synchronization signal coming from a GPS-enabled apparatus [1], such as a substation clock. Because GPS time-synchronization signals are distributed at 1 sample per sec., the PMU needs to provide sub-second time through an internal oscillator (i.e. an internal clock). This clock can drift from the "true" GPS time due to several factors discussed in [6], [7].

B. Issues with Angle Data

Angle data coming from synchrophasor measurements provided by PMU is prone to present different types of quality issues, the most common being outliers, which are random in nature. However, spurious data points that present a certain periodicity can be associated with time-synchronization issues [6], which are observed as small repeated variations. This can also be observed even after the angle data has been filtered and detrended for use in spectral analysis.

To illustrate these issues, Figs. 1 and 2, show angle data from a PMU in Dominion Energy's service territory. In Fig. 1, highlighted in light blue, a defective data point in the filtered and detrended angle data is shown. While the outlier removal process can usually deal with these kind of the deviations in the measurement set [2], it will fail in this case. This is because, as shown in Figure 1, the deviation is small with respect to the other measurements in the data set. If these spurious data points were only present sporadically, they would not be of major concern because they deviate little from the other data points. Note that conversely, large outliers can introduce artificial transient behavior in filtered data [2]. However, as shown in Fig. 2, these defective points are periodic and occur at many different time-ranges in the measured data sets.

Different methods to determine the frequency spectrum in PMU data, one common approach is to compute the signal's Power Spectral Density (PSD). To compute a PSD estimate, Welch's method for PSD estimation [5], [17] can be used for the identification of oscillatory modes in a power system [5]. Therefore, it is also important to determine the impact of the spurious data points shown in Figure 2 when the data is transformed to the frequency domain. Indeed, because these spurious data points are recurrent and have no relation to any natural dynamic behavior of the system, they introduce steep spikes in the signal's PSD. This fact is illustrated in Figure 3. Note that the sharp spikes are present at different frequencies precluding the identification of other natural modes that might be underneath. Furthermore, techniques such as Yule-Walker that are used to estimate the frequency and damping of natural modes will have their performance severely deteriorated [18] if the spikes are not removed. Hence, the need for a technique to remove the artificial spikes from data is evident.

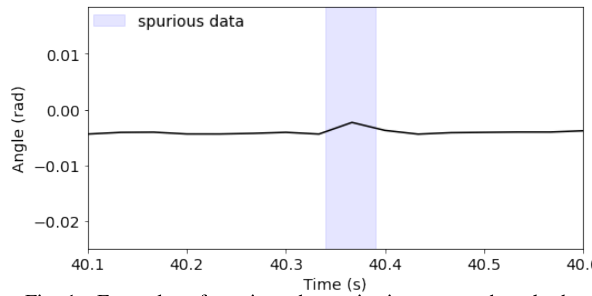


Fig. 1. Examples of spurious data point in processed angle data.

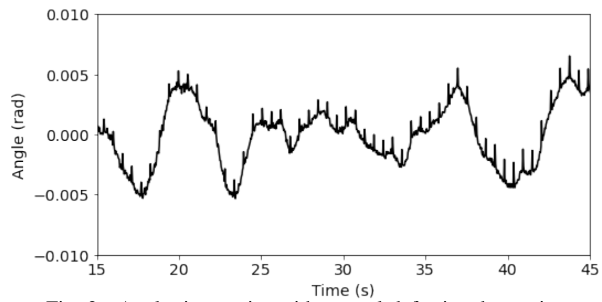


Fig. 2. Angle time series with several defective data points.

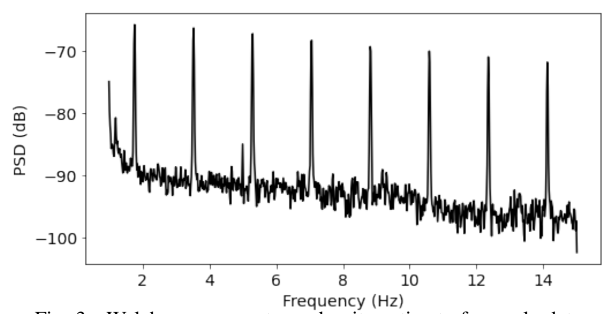


Fig. 3. Welch power spectrum density estimate for angle data.

III. Proposed Approach

Raman Spectrum gives the the scattering intensity as a function of frequency shifts. For our case i.e. the phase angle, the Raman Spectrum can be replaced by the PSD estimate. There is however one major challenge when directly using the spike removal techniques on the phase angle spectrum. These techniques rely on the spectral slope estimate for detecting the spikes where they are marked by steep slopes. Now, the spectral slope is not only a function of the spike's slope but also depends on the slope of the spectral baseline. For phase angle spectrum, the baseline is especially steep for lower frequencies due to the influence of various generator controls. This can make it difficult to decide a threshold on the calculated slope for distinguishing the normal part from a spike. Therefore, our approach first processes the spectrum by estimating and consequently flattening the baseline followed by spike removal. The two key steps viz. baseline correction and spike removal are done using standard techniques, which are discussed next as applied to the phase angle data. The PMU data received goes through standard pre-processing based on the final application. Let this data be denoted by the uniformly sampled time series $\mathcal{X} \in \mathbb{R}^n : \{x_i\}, i = 1 \dots n$.

A. Spectrum Baseline Estimation and Correction

The spectral baseline $\mathbf{b}$ can be defined as a smooth, relatively flat, under approximation of the original spectrum. While in the current application, baseline is being estimated to make it easier to detect spikes, baseline removal is also important for effectively comparing multiple spectra. An asymmetric least squares based approach has been proposed in [19], [20], which is employed in this work with slight modifications. Let the PSD (in dB) of $\mathcal{X}$ be denoted by $\Phi \in \mathbb{R}^n : \{\phi_i\}$. The estimation of baseline $\mathbf{b}$ for the PSD is modeled as the following optimization problem,

$$\mathbf{b} = \arg\min_{\mathbf{b}} \left\{ \sum_i w_i(\phi_i - b_i)^2 + \lambda \sum_i (\nabla^2 b_i)^2 \right\}, \quad (1)$$

where ∇ is the difference operator. The weight w_i assumes the asymmetric form $w_i = p$ if $\phi_i > b_i$ else $w_i = 1 - p$, and λ is a parameter penalising the curvature. The effect of parameter p is such that lower the value, more is the baseline $\mathbf{b}$ forced to be smaller than the original spectrum. The minimization problem give by (1) leads to

$$(\mathbf{W} + \lambda \mathbf{D}^\intercal \mathbf{D})\, \mathbf{b} = \mathbf{W}\mathbf{\Phi} \quad (2)$$

where $\mathbf{W} = diag(w_1, w_2, \cdots, w_i, \cdots, w_n)$ and $\mathbf{D} = \nabla^{\mathbf{2}}$ is the second order difference matrix [21]. Equation (2) can be solved for $\mathbf{b}$ through iteration. This is then removed from the PSD to obtain the flat baselined PSD $\Phi^b : \{\phi_i^b = \phi_i - b_i\}$. The typical values used for λ are from 10^2 to 10^9 and p are from 0.0001 to 0.1.

B. Whitaker-Hayes Method for Despiking Spectra

For spike removal, the *Whitaker-Hayes Method* (WHM) [16] is used. The WHM starts by estimating the spectral slope using a detrended difference series $\nabla \Psi^b$ as

$$\nabla \psi_i^b = \psi_i^b - \psi_{i-1}^b, \quad (3)$$

for $i = 2, 3, ..., n$. Here, $\Psi^b : \{\psi_i^b = antilog\left(\phi_i^b\right)\}$. Spikes correspond to large values can be easily identified. To automate this process, each index i in the spectrum value is given a score as follows

$$Z_i = \frac{0.6745 \left(\nabla \psi_i^b - M\right)}{M_{ad}}, \quad (4)$$

where M is the median of the series $\{\nabla \psi_i^b\}$ and M_{ad} is the median of the absolute deviation, i.e. the median of the series $\{|\nabla \psi_i^b - M|\}$. More details can be found in [16]. Subsequently, a threshold τ is defined such that the set of indices on the spectrum corresponding to spikes is given as $\mathcal{K} = \{k | Z_k > \tau\}$. Let, $\mathcal{Y} \in \mathbb{C}^n : \{y_i\}$ denote the FFT coefficients of the original phase angle time series $\mathcal{X}$. Finally, the spikes are flattened in the frequency domain using the spectrum values at neighboring frequencies to obtain the de-spiked FFT coefficients $\mathcal{Y}^* : \{y_i^*\}$ as follows

$$y_i^* = \begin{cases} mean\left([y_j \mid j \notin \mathcal{K}, j \in \mathcal{M}_i]\right), & \text{if } i \in \mathcal{K} \\ y_i, & \text{otherwise} \end{cases} \quad (5)$$

where $\mathcal{M}_i = [i - \frac{m-1}{2}, i + \frac{m-1}{2}]$. Finally, the corresponding de-spiked time series $\mathcal{X}^* : \{x_i^*\}$ can then be obtained by inverse FFT of $\mathcal{Y}^*$.

C. Overall Approach

The overall proposed approach is summarized in Figure 4. There are basically four parameters to be set. The parameters

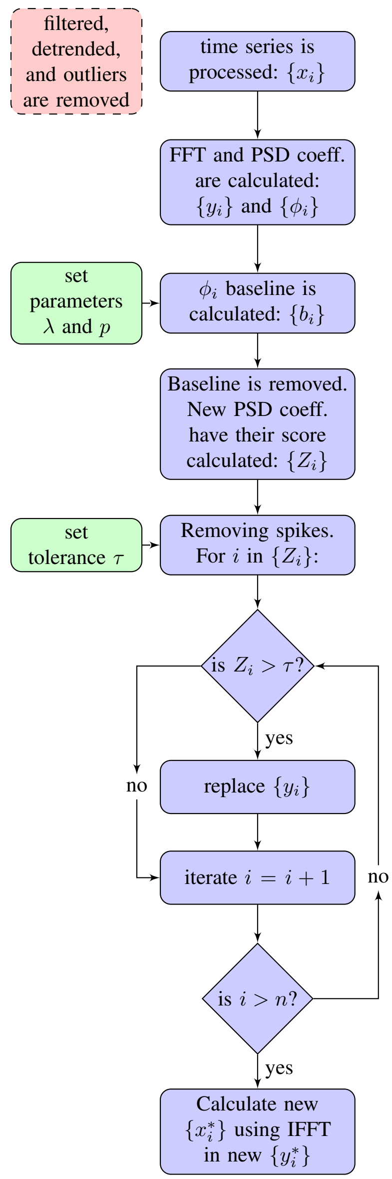


Fig. 4. Flowchart describing the proposed algorithm.

λ and p for the baseline calculation, the parameter τ, which is the tolerance used to flag data considered to be part of a spike, and the parameter m, which is the number of neighboring spectrum points for flattening each spike.

IV. Case Study Results

The data used in this case study was measured on July 5th, 2019 in one of Dominion Energy's substations. The data was sampled with a frequency of 30 Hz during an interval of 20 minutes. The data was then pre-processed as recommended in [2], i.e., the data was filtered, re-sampled, detrended and all the outliers were removed. The filter was set to have a low-pass frequency of 15 and a high-pass frequency of 0.1 Hz.

A Welch estimate of the PSD is calculated using the entire time interval and the baseline is computed using this estimate. The parameters for the asymmetric baseline calculation were set to be $p = 0.01$ and $\lambda = 10^7$. Note that the baseline is not used to remove the modes from the PSD estimate, but to eliminate its trends as it is shown in Figure 5. Therefore, the baseline does not need to be closely aligned with the spectrum, but it should depict a trend instead. In addition, observe that, in this particular case, the baseline removal results in minor adjustments if compared to the original PSD estimate.

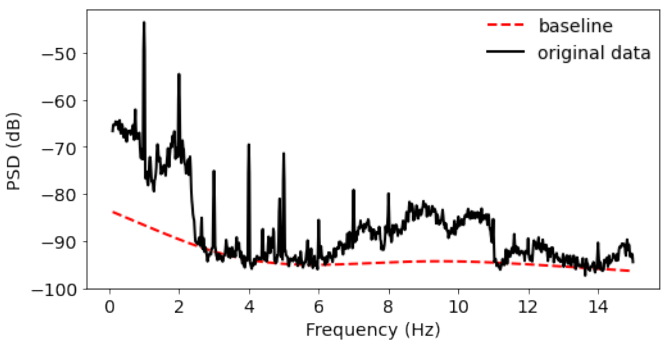


Fig. 5. Comparison of original PSD estimate and calculated baseline for baseline removal procedure.

The score Z_i is then calculated using the PSD data after the baseline has been removed. The results for the score are shown in Figure 6. Note that the spikes attain a high score while other data, e.g. the modes at ≈ 1.9 Hz, are given a lower score. This is important because PMU applications can be affected by the spike [18], and thus, depend on removing spikes accurately. In this step, the tolerance τ should be adjusted. In this case study, the tolerance was set to be $\tau = 13$ and is displayed with a red dashed line in Fig. 6.

Next, the procedure of despiking the data is applied using $m = 50$, where m is the number of neighbor samples for which the average is calculated for FFT coefficient replacement. The results for a 200-second interval of angle time-series is shown in Fig. 7a. Note that the time-series obtained by despiking the PSD estimate preserve the underlying physical behavior contained in the original data, but with a smaller variance. This is confirmed in Figure 7b where a comparison between the Welch PSD estimates from filtered and original data is depicted. Note that many of the sharp spikes have been removed.

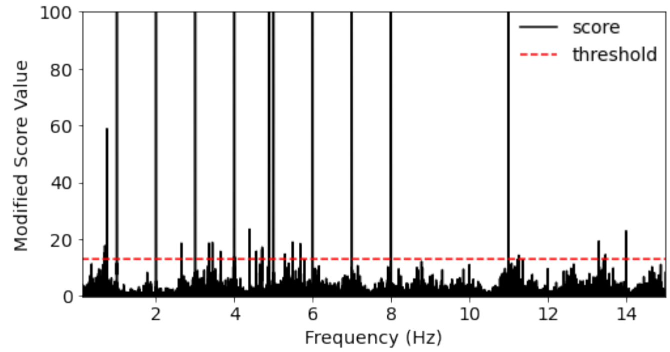
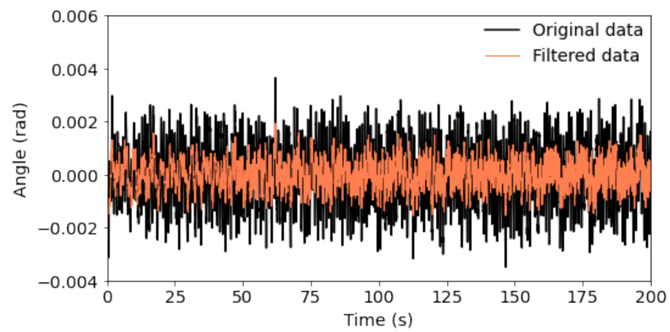
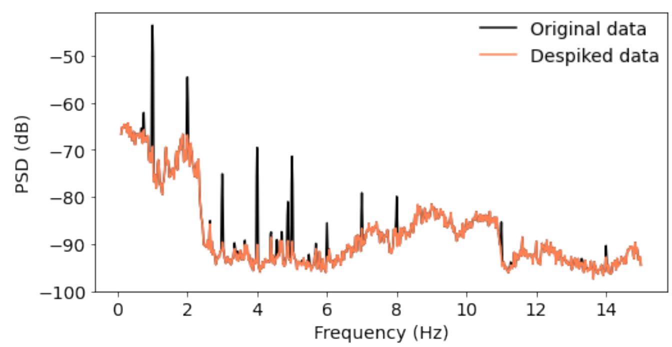


Fig. 6. Modified Z_i scores and comparison with tolerance τ.



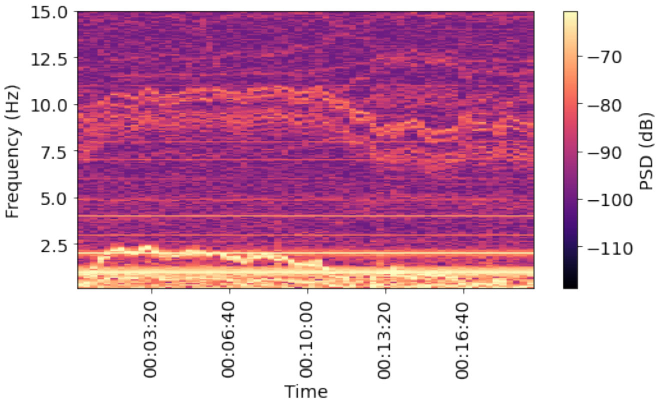
(a) Comparison of original and filtered angle time series.



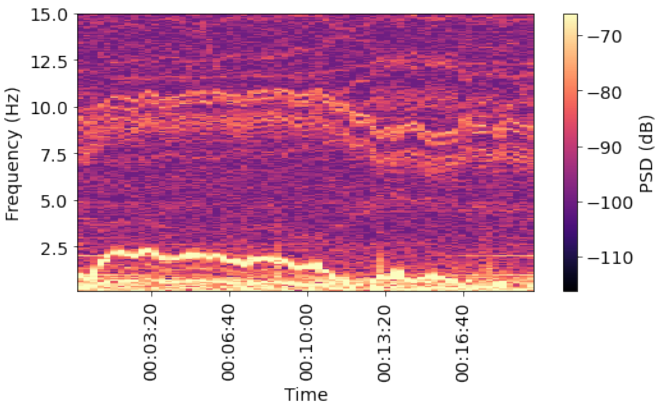
(b) Comparison of original and filtered angle PSD estimate.

Fig. 7. Outcome verification for proposed algorithm.

In addition, the effectiveness of the proposed algorithm can also be verified by using spectrograms [5], which are plots that display the behavior of the PSD estimate over time. In this work, the color scheme was chosen so the brightest color in the spectrogram corresponds to the highest PSD estimated value in dB. Note that the original spectrogram in Figure 8a, depicts horizontal lines that are product of spikes. These represent the artificial oscillations caused by the defective data. On the other hand, observe how in the spectrogram shown in Figure 8b the horizontal lines have been largely removed. This implies that data coming from the despiking procedure also provide spectrograms that allow for better interpretation of the physical phenomena, which is, otherwise, difficult to understand due to the effect of the spikes.



(a) Spectrogram created with original data.



(b) Spectrogram created with despiked data.

Fig. 8. Comparison of spectrograms.

V. Conclusion

The current work proposes a new algorithm to remove spikes from Welch PSD estimates computed from angle data from PMUs. The method proposed in this paper is based on baseline removal and on the calculation of a modified Z score. Both steps require the user to adjust two parameters, respectively. The former requires the setting of p and λ such that the baseline is calculated. The latter requires the user to set a tolerance τ and how many samples m are used for average calculation. All required parameters can be easily tuned but the authors are currently investigating ways of automating this step.

The proposed method was applied to real-world PMU data from a unit at Dominion Energy's network, which is known to have time synchronization issues. The results show the effectiveness of the proposed method in removing the spikes. The results from the PSD and the spectrogram indicate that the despiking routine improves interpretation and might improve mode identification and damping estimation PMU data-based applications. Hence, the proposed algorithm is now being included as an initial step in frequency domain analysis of angle data and has been implemented in the PredictiveGrid PMU data platform at Dominion Energy.

References

[1] J. De La Ree, V. Centeno, J. S. Thorpe and A. G. Phadke, "Synchronized phasor measurement applications in power systems", *IEEE Transactions on Smart Grid*, vol. 1, No. 1, June, pp. 20-27, 2010.

[2] L. Vanfretti, S. Bengtsson, J. O. Gjerde, "Preprocessing synchronized phasor measurement data for spectral analysis of electromechanical oscillations in the Nordic Grid", *International Transactions on Electrical Energy Systems*, vol. 25, No. 2, February, pp. 348-358, 2015.

[3] J. W. Pierre, D. J. Trudnowski, M. K. Donnelly, "Initial results in electromechanical mode identification from ambient data", *IEEE Transactions on Power Systems*, vol. 12, No. 3, August, pp. 1245-1251, 1997.

[4] J. F. Hauer, D. J. Trudnowski, J. G. DeSteese, "A perspective on WAMS analysis tools for tracking of oscillatory dynamics", In Proc. 2007 IEEE Power Engineering Society General Meeting, 2007.

[5] L. Vanfretti, R. García-Valle, K. Uhlen, E. Johansson, D. Trudnowski, J. W. Pierre, J. H. Chow, O. Samuelsson, J. Ostergaard and K. E. Martin, "Estimation of Eastern Denmark's Electromechanical Modes from Ambient Phasor Measurement Data", In Proc. 2010 IEEE Power Engineering Society General Meeting, 2010.

[6] M. S. Almas, L. Vanfretti, R. S. Singh and G. M. Jonsdottir, "Vulnerability of Synchrophasor-Based WAMPAC Applications' to Time Synchronization Spoofing," *IEEE Transactions on Smart Grid*, vol. 9, No. 5, Sept., pp. 4601-4612, 2018.

[7] R. S. Singh, H. Hooshyar and L. Vanfretti, "Laboratory test set-up for the assessment of PMU time synchronization requirements," 2015 IEEE Eindhoven PowerTech, Eindhoven, 2015, pp. 1-5, doi: 10.1109/PTC.2015.7232731.

[8] M. A. Khan and J. W. Pierre, "Separable Estimation of Ambient Noise Spectrum in Synchrophasor Measurements in the Presence of Forced Oscillations," *IEEE Transactions on Power Systems*, vol. 35, No. 1, June, pp. 415-423, 2019.

[9] J. Follum, F. Tuffner and U. Agrawal, "Applications of a new non-parametric estimator of ambient power system spectra for measurements containing forced oscillations," In Proc. 2017 IEEE Power & Energy Society General Meeting, 2017.

[10] D. Zhang, K. N. Jallad, and D. Ben-Amotz, "Stripping of Cosmic Spike Spectral Artifacts Using a New Upper-Bound Spectrum Algorithm" *Applied Spectroscopy*, vol. 55, No. 11, pp. 1523–1531, 2001.

[11] H. G. Schulze, and R. F. Turner, "A Two-Dimensionally Coincident Second Difference Cosmic Ray Spike Removal Method for the Fully Automated Processing of Raman Spectra," *Applied Spectroscopy*, vol. 68, No. 2, pp. 185–191, 2014.

[12] S. Mozharov, A. Nordon, D. Littlejohn, and B. Marguardt, "Automated Cosmic Spike Filter Optimized for Process Raman Spectroscopy," *Applied Spectroscopy*, vol. 66, No. 11, pp. 1326–1333, 2012.

[13] Y. Katsumoto and Y. Ozaki, "Practical Algorithm for Reducing Convex Spike Noises on a Spectrum," *Applied Spectroscopy*, vol. 57, No. 3, pp. 317–322, 2003.

[14] F. Ehrentreich and L. Summchen, "Spike Removal and Denoising of Raman Spectra by Wavelet Transform Methods," *Analytical Chemistry*, vol. 73, No. 17, pp. 4364–4373, 2001.

[15] W. Hill and D. Rogalla, "Spike-Correction of Weak Signals from Charge-Coupled Devices and Its Application to Raman Spectroscopy," *Analytical Chemistry*, vol. 64, No. 21, pp. 2575–2579, 1992.

[16] D. A. Whitaker and K. Hayes, "A simple algorithm for despiking Raman spectra", *Chemometrics and Intelligent Laboratory Systems*, vol. 179, August, pp. 82-84, 2018.

[17] J. Proakis and D. Manolakis *Digital Signal Processing: Principles, Algorithms and Applications*, 4th ed., New Jersey: Prentic Hall, 2006.

[18] L. Vanfretti, S. Bengtsson, V. S. Peric and J. O. Gjerde, "Effects of forced oscillations in power system damping estimation," in Proc. of the IEEE International Workshop on Applied Measurements for Power Systems, Aachen, pp. 1-6, 2012.

[19] P. H. C. Eilers, "A Perfect Smoother", *Analytical Chemistry*, vol. 75, No. 14, July, pp. 3631-3636, 2003.

[20] P. H. C. Eilers, "Parametric Time Warping", *Analytical Chemistry*, vol. 76, No. 2, January, pp. 404-411, 2004.

[21] S. He, W. Zhang, L. Liu, Y. Huang, J. He, W. Xie, P. Wu, C. Du, "Baseline correction for Raman spectra using an improved asymmetric least squares method", *Analytical Methods*, vol. 6, No. 12, pp. 4402-4407, 2014.