A Data Driven Approach to Robust Event Detection in Smart Grids Based on Random Matrix Theory and Kalman Filtering †

Fujia Han [1,*], Phillip M. Ashton [2], Maozhen Li [1], Ioana Pisica [3], Gareth Taylor [3], Barry Rawn [3] and Yi Ding [4]

[1] Department of Electronic and Computer Engineering, Brunel University London, London UB8 3PH, UK; maozhen.li@brunel.ac.uk
[2] Network Operations, National Grid, Wokingham RG41 5BN, UK; phil.ashton@nationalgrid.com
[3] Brunel Institute of Power Systems, Brunel University London, London UB8 3PH, UK; ioana.pisica@brunel.ac.uk (I.P.); gareth.taylor@brunel.ac.uk (G.T.); barry.rawn@brunel.ac.uk (B.R.)
[4] College of Electrical Engineering, Zhejiang University, Hangzhou 310000, China; yiding@zju.edu.cn
* Correspondence: fujia.han2@brunel.ac.uk
† This paper is an extended version of our paper published in 2018 IEEE Power & Energy Society General Meeting (PESGM), Portland, OR, USA, 5–10 August 2018.

Abstract: Increasing levels of complexity, due to growing volumes of renewable generation with an associated influx of power electronics, are placing increased demands on the reliable operation of modern power systems. Consequently, phasor measurement units (PMUs) are being rapidly deployed in order to further enhance situational awareness for power system operators. This paper presents a novel data-driven event detection approach based on random matrix theory (RMT) and Kalman filtering. A dynamic Kalman filtering technique is proposed to condition PMU data. Both simulated and real PMU data from the transmission system of Great Britain (GB) are utilized in order to validate the proposed event detection approach and the results show that the proposed approach is much more robust with regard to event detection when applied in practical situations.

Keywords: event detection; Kalman filtering; phasor measurement units (PMUs); random matrix theory (RMT); situational awareness

1. Introduction

The development and increasing complexity of modern power systems is leading to growing concerns over power system operational security, with network management and analysis becoming increasingly challenging. Traditional monitoring devices, employed in supervisory control and data acquisition (SCADA) systems, have become inefficient for this challenge [1]. As a result, phasor measurement units (PMUs) are being deployed to measure phasors of bus voltages and currents, enabling real-time state monitoring in power systems [2,3]. Due to PMUs' higher sampling rates and accurate time synchronization, a huge amount of synchronous data is now being collected from power systems globally [4,5], facilitating the detection of transient events.

At present, event detection is carried out following either a model-driven approach (supervised) or a data-driven approach (unsupervised). Model-driven approaches include neural networks [6,7], decision trees [8,9], and principal component analysis (PCA) [10–14]. However, for supervised learning, it is argued that if the training dataset is selected inappropriately or insufficiently, it will have a significantly negative effect on the performance of the approach when it is applied online. Therefore, much attention has been paid to the employment of unsupervised learning in this field [15–18], including wavelet-based approaches, energy function-based approaches, and frequency difference-based approaches.

More recently, as a data-driven approach, random matrix theory (RMT) was introduced and applied to early event detection in power systems [19]. However, PMU data are normally collected with significant noise levels and plenty of missing measurements,

due to communication issues between PMUs and phasor data concentrators (PDCs) [20,21] and further issues in the measurement processes. When the above approach is faced with significantly noisy and missing data, it will be unable to accurately detect abnormal events.

A large amount of research has been presented in the field of event detection for power systems. In [7], Vasilic and Kezunovic presented an approach for transmission line fault classification building on an adaptive resonance theory (ART) neural network and fuzzy logic. This work applies a fuzzy decision rule to the output of the ART neural network to improve selectivity for a variety of real events. In [9], Adhikari et al. presented a method to effectively classify traditional power system contingencies and cyber-attacks in real time. It combines Hoeffding adaptive trees with a drift detection method and adaptive windowing to update the model continuously.

In recent years, PCA has been widely applied for early event detection in power systems. For example, Xie et al. [10] analyzed the dimensionality of PMU data under both normal and abnormal conditions and employed PCA to detect abnormal events based on the change of core subspaces of PMU data at the occurrence of an anomaly. Meanwhile, this model has an adaptive training mechanism that updates the model when an event happens. A similar work was carried out in [11] that uses a sliding window PCA to detect and classify multiple events, including islanding, loss of load, and loss of generation. However, different from the work presented in [10], this method uses a sliding window to update the PCA model in real time to adapt to the time varying behaviors of power systems. Another similar event detection method was developed by Guo et al. [12], which focuses on islanding detection for distributed generation systems. In this work, a recursive PCA algorithm is employed to analyze the time-varying behaviors of power systems for the purpose of reducing false alarms. Although a wide range of machine learning techniques have been applied to event detection and classification, these methods are mainly based on model-driven supervised learning. As mentioned above, when the sample space is chosen improperly or insufficiently, an ill-trained model may be obtained.

In order to address the limitations of supervised learning techniques, a few works have considered unsupervised learning techniques. For instance, Kim et al. designed a wavelet-based detection method using PMU data [15]. The key idea is to monitor the energy of coefficients within a sliding window in real time and to determine whether the energy value exceeds the threshold. Apart from this work, another data-driven analytics method, which relies on the rules created by PMU data, was proposed for power system fault detection in [18]. However, it should be noted that these unsupervised approaches need to pre-define an appropriate threshold in detection of events, which is a challenging issue. The RMT-based event detection approach presented in [19] does not need to pre-define a threshold, but this approach does not take into account the quality aspects of PMU data in dealing with noisy and missing PMU measurements.

This paper presents an event detection approach that improves the performance of the work presented in [19], making it much more robust when dealing with corrupted and significantly noisy samples of PMU data. To be specific, a robust event detection approach is developed by combining RMT and Kalman filtering. Furthermore, a dynamic Kalman filter is developed to condition PMU data. Finally, the proposed event detection approach is tested and validated on the standard IEEE 118-bus system and the transmission system of Great Britain (GB). The results show that the proposed approach is much more robust in practical situations that include significant levels of noisy or missing PMU measurements.

The remainder of this paper is organized as follows. Section 2 briefly introduces random matrix theory and formulates a dynamic quadratic prediction model based Kalman filtering technique. Section 3 combines RMT with Kalman filtering for event detection. Section 4 evaluates the performance of the proposed event detection approach on both simulated and real PMU data. Finally, Section 5 concludes the paper and points out some future work.

2. Random Matrix Theory and Kalman Filtering

Random matrix theory is first introduced briefly and then a dynamic quadratic prediction model based Kalman filtering technique is formulated.

2.1. Random Matrix Theory

RMT has proven to be effective for power system analysis, and a wide range of laws and theorems, such as Ring Law and Marchenko–Pastur Law, are included in RMT, which have been applied in power systems for different purposes [19,22,23]. In this paper, for the purpose of event detection, Ring Law, as the theoretical basis of data processing of event detection, and mean spectral radius (MSR), as an event indicator, will be briefly introduced.

2.1.1. Ring Law

Let $\mathbf{X}_s \in \mathbf{C}^{N \times T}$ be a standard non-Hermitian random matrix, whose entries are independent and identically distributed (i.i.d.) variables with

$$\mu(\mathbf{x}_{s,i}) = 0, \sigma^2(\mathbf{x}_{s,i}) = 1 \ (i = 1, 2, \ldots, N) \tag{1}$$

where $\mathbf{x}_{s,i}$ is the ith row vector of $\mathbf{X}_s$. For L standard non-Hermitian random matrices $\mathbf{X}_{s,i}(i = 1, 2, \ldots, L)$, a matrix product is defined as

$$\mathbf{Z} = \prod_{i=1}^{L} \mathbf{X}_{u,i} \tag{2}$$

where $\mathbf{X}_{u,i}$ is the singular value equivalent of $\mathbf{X}_{s,i}$. The matrix product $\mathbf{Z}$ can be transformed to the standard matrix product $\mathbf{Z}_s$, whose $\sigma^2(\mathbf{z}_{s,i}) = 1/N$ in each row. Thus, the empirical spectral density (ESD) of $\mathbf{Z}_s$ converges almost surely to the limit given by

$$f(\lambda_{\mathbf{Z}_s}) = \begin{cases} \frac{1}{\pi c L} |\lambda_{\mathbf{Z}_s}|^{\frac{2}{L}-2} & (1-c)^{\frac{L}{2}} \leq |\lambda_{\mathbf{Z}_s}| \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

As $N, T \to \infty$ with a constant ratio $c = N/T \in (0,1]$, N is the number of variables, and T is the time length.

2.1.2. MSR

Linear eigenvalue statistics indicate the statistical features of random matrices. A linear eigenvalue statistic of a random matrix $\mathbf{X}$ is defined as

$$S_n(\phi) = \sum_{i=1}^{n} \phi(\lambda_i) \tag{4}$$

where $\lambda_i(i = 1, 2, \ldots, n)$ are eigenvalues of $\mathbf{X}$, and $\phi(\cdot)$ is a test function.

MSR is used to indicate the eigenvalue distribution of a random matrix. For the standard matrix product $\mathbf{Z}_s$ (as mentioned above in *Ring Law*), MSR is formulated as

$$\text{MSR} = \frac{1}{N} \sum_{i=1}^{N} |\lambda_{\mathbf{Z}_s,i}| \tag{5}$$

where $\lambda_{\mathbf{Z}_s,i}(i = 1, 2, \ldots, N)$ are eigenvalues of $\mathbf{Z}_s$, and $|\lambda_{\mathbf{Z}_s,i}|$ is the radius of $\lambda_{\mathbf{Z}_s,i}$ on the complex plane.

2.1.3. Data Processing of Ring Law

Within a raw data source $\mathbf{\Omega}$, a raw random matrix $\mathbf{X} \in \mathbf{C}^{N \times T}$ can be formed through a sliding window. Then, $\mathbf{X}$ can be transformed to a standard non-Hermitian matrix $\mathbf{X}_{\mathrm{s}} \in \mathbf{C}^{N \times T}$ row by row as follows:

$$\mathbf{x}_{\mathrm{s},i} = (\mathbf{x}_i - \mu(\mathbf{x}_i)) \frac{\sigma(\mathbf{x}_{\mathrm{s},i})}{\sigma(\mathbf{x}_i)} + \mu(\mathbf{x}_{\mathrm{s},i}) \ (i = 1, 2, \dots, N) \tag{6}$$

where $\mathbf{x}_i$ and $\mathbf{x}_{\mathrm{s},i}$ are the ith row vectors of $\mathbf{X}$ and $\mathbf{X}_{\mathrm{s}}$, respectively, and $\mu(\mathbf{x}_{\mathrm{s},i}) = 0$, $\sigma^2(\mathbf{x}_{\mathrm{s},i}) = 1$.

Afterward, the matrix $\mathbf{X}_{\mathrm{u}} \in \mathbf{C}^{N \times N}$ is introduced as the singular value equivalent of $\mathbf{X}_{\mathrm{s}}$ by

$$\mathbf{X}_{\mathrm{u}} = \sqrt{\mathbf{X}_{\mathrm{s}}\mathbf{X}_{\mathrm{s}}{}^{\mathrm{T}}} \mathbf{U} \tag{7}$$

where $\mathbf{X}_{\mathrm{s}}{}^{\mathrm{T}}$ is the transpose of $\mathbf{X}_{\mathrm{s}}$, and $\mathbf{U} \in \mathbf{C}^{N \times N}$ is a Haar unitary matrix, $\mathbf{X}_{\mathrm{u}}\mathbf{X}_{\mathrm{u}}{}^{\mathrm{T}} \equiv \mathbf{X}_{\mathrm{s}}\mathbf{X}_{\mathrm{s}}{}^{\mathrm{T}}$.

However, for a series of arbitrary non-Hermitian random matrices $\mathbf{X}_i (i = 1, 2, \dots, L)$ from the raw data source $\mathbf{\Omega}$, the matrix product $\mathbf{Z} = \prod_{i=1}^{L} \mathbf{X}_{\mathrm{u},i} \in \mathbf{C}^{N \times N}$ is obtained. Then, it is converted to the standard matrix product $\mathbf{Z}_{\mathrm{s}} \in \mathbf{C}^{N \times N}$ through the following formula:

$$\mathbf{z}_{\mathrm{s},i} = \frac{\mathbf{z}_i}{\sqrt{N}\sigma(\mathbf{z}_i)} \ (i = 1, 2, \dots, N) \tag{8}$$

where $\mathbf{z}_i$ and $\mathbf{z}_{\mathrm{s},i}$ are the ith row vectors of $\mathbf{Z}$ and $\mathbf{Z}_{\mathrm{s}}$, respectively.

2.2. Dynamic Kalman Filtering Technique

As a state estimation technique, Kalman filtering is a popular method for conditioning PMU data in power systems [24–26]. Jones et al. [27] proposed a quadratic prediction model-based Kalman filtering method to cleanse and recover PMU data easily. However, within this method, the measurement noise covariance matrix $\mathbf{R}$ is constant, thus leading to inaccurate PMU data conditioning in some situations. As a result, this paper makes an improvement by dynamically adapting the measurement noise covariance matrix $\mathbf{R}$ to the conditioning process.

Following [27], the classic model of Kalman filtering is defined as follows:

$$\begin{cases} \mathbf{x}(t+1) = \mathbf{\Phi}(t+1,t)\mathbf{x}(t) + \mathbf{\Gamma}(t+1,t)\mathbf{w}(t) \\ \mathbf{z}(t+1) = \mathbf{H}(t+1)\mathbf{x}(t+1) + \mathbf{v}(t+1) \end{cases} \tag{9}$$

For the purpose of optimization, the classic model of Kalman filtering can be expressed in the recursive way:

$$\begin{cases} \hat{\mathbf{x}}(t+1|t+1) = \mathbf{\Phi}(t+1,t)\hat{\mathbf{x}}(t|t) \\ \qquad\qquad\qquad + \mathbf{K}(t+1)(\mathbf{z}(t+1) - \hat{\mathbf{z}}(t+1|t)) \\ \hat{\mathbf{z}}(t+1|t) = \mathbf{H}(t+1)\hat{\mathbf{x}}(t+1|t) \end{cases} \tag{10}$$

According to [28], the quadratic prediction model is formulated by

$$\hat{x}(t+1|t) = 3\hat{x}(t|t) - 3\hat{x}(t-1|t-1) + \hat{x}(t-2|t-2) \tag{11}$$

In addition, each state vector is formed by a sliding window, which contains three successive snapshots of the system states and moves forward only one snapshot at a time. Thus, in order to estimate the next state, the state vectors $\hat{\mathbf{x}}(t|t)$ and $\hat{\mathbf{x}}(t+1|t)$ are respectively expressed as follows [27]:

$$\hat{\mathbf{x}}(t|t) = [\ \hat{x}(t|t) \quad \hat{x}(t-1|t-1) \quad \hat{x}(t-2|t-2)\ ]^{\mathrm{T}} \tag{12}$$

$$\hat{\mathbf{x}}(t+1|t) = [\ \hat{x}(t+1|t) \quad \hat{x}(t|t) \quad \hat{x}(t-1|t-1)\ ]^{\mathrm{T}} \tag{13}$$

where $\hat{x}(t+1|t)$, $\hat{x}(t|t)$, $\hat{x}(t-1|t-1)$, and $\hat{x}(t-2|t-2)$ are all complex values, including magnitude and phase angle.

As for $\mathbf{\Phi}(t+1, t)$ and $\mathbf{H}(t+1)$, they can be expressed as constant matrices shown in (14) and (15):

$$\mathbf{\Phi}(t+1,t) = \begin{bmatrix} 3 & -3 & 1; & 1 & 0 & 0; & 0 & 1 & 0 \end{bmatrix} \tag{14}$$

$$\mathbf{H}(t+1) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \tag{15}$$

$\mathbf{K}(t+1)$ is specified by the following formulas:

$$\begin{cases} \mathbf{K}(t+1) = \mathbf{P}(t+1|t)\mathbf{H}^{\mathrm{T}}(t+1) \\ \qquad \times [\mathbf{H}(t+1)\mathbf{P}(t+1|t)\mathbf{H}^{\mathrm{T}}(t+1) + \mathbf{R}(t+1)]^{-1} \\ \mathbf{P}(t+1|t) = \mathbf{\Phi}(t+1,t)\mathbf{P}(t|t)\mathbf{\Phi}^{\mathrm{T}}(t+1,t) \\ \qquad + \mathbf{\Gamma}(t+1,t)\mathbf{Q}(t)\mathbf{\Gamma}^{\mathrm{T}}(t+1,t) \\ \mathbf{P}(t+1|t+1) = [\mathbf{I} - \mathbf{K}(t+1)\mathbf{H}(t+1)]\mathbf{P}(t+1|t) \end{cases} \tag{16}$$

where $\mathbf{R}$ and $\mathbf{Q}$ are constant scalar values in this method. Furthermore, it is assumed that the Kalman filtering process starts at $t = 3$, $\hat{x}(3|3)$, and $\mathbf{P}(3|3)$ are the matrixes with 1 s and 0 s, respectively. This is to make the Kalman filter track the optimal estimate fast within a short period.

As the measurement noise covariance matrix $\mathbf{R}$ has a significant influence on the performance of Kalman filtering, this parameter will be adjusted dynamically. First of all, a residue between the a posteriori estimate and the actual measurement is calculated at each time t as follows [29]:

$$r(t) = \hat{x}(t|t) - z(t) \tag{17}$$

where $r(t)$ is the residue at time t, and $z(t) = \mathbf{z}(t)$. For a period, a residue vector is formed as:

$$\mathbf{r}(t+1) = \begin{bmatrix} r(t-W+1) & r(t-W+2) & \cdots & r(t-1) & r(t) \end{bmatrix} \tag{18}$$

where $\mathbf{r}(t+1)$ is the residue vector at time $t+1$, and W is the time length. Then, the variance of the residue vector is calculated as the measurement noise covariance matrix shown below:

$$\mathbf{R}(t+1) = \mathrm{Var}(\mathbf{r}(t+1)) \tag{19}$$

where $\mathbf{R}(t+1)$ is the measurement noise covariance matrix at time $t+1$. Here, it is noted that $\mathrm{Var}(\mathbf{r}(t+1))$ is the variance of a series of complex values. According to (18) and (19), it can be easily seen that $\mathbf{R}(t+1)$ is obtained on the basis of W previous residues—namely, from time $t-W+1$ to time t.

However, when the noise is weak, $\mathbf{R}(t+1)$ approaches 0. As a result, when $\hat{x}(t+1|t+1)$ is calculated, $z(t+1)$ is almost trusted fully, while $\hat{x}(t+1|t)$ is hardly trusted, thus failing to reduce noise effectively. To solve this problem, $R_{\min}$ is introduced as the lower bound for $\mathbf{R}(t+1)$:

$$R_{\min=\alpha} \tag{20}$$

where α is a small value. When $\mathbf{R}(t+1)$ is less than $R_{\min}$, $\mathbf{R}(t+1)$ is set to $R_{\min}$ by default. Additionally, it is supposed that $\mathbf{R}(t+1) = R_{\min}$ within the initial period W.

Similarly, when missing data exist, which is processed as 0 in this paper, $R_{\max}$ is defined as the upper bound for $\mathbf{R}(t+1)$ to enhance the trust in $\hat{x}(t+1|t)$:

$$R_{\max=\beta} \tag{21}$$

where β is a large value. Specifically, in the presence of missing data at time $t+1$, $\mathbf{R}(t+1)$ is set to $R_{\max}$ to recover the missing data. This practice is based on the fact that $z(t+1)$ is far from the real value at time $t+1$ and can be hardly trusted.

3. Real-Time Event Detection

To perform real-time analysis, an $N \times T$ real-time sliding window is utilized to obtain the raw matrix $\mathbf{X}$ from the raw PMU data source $\mathbf{\Omega}$. Let N denote the number of voltage magnitudes and phase angles and T denote the time period. At time t, the raw matrix $\mathbf{X}(t)$ is formed as follows:

$$\mathbf{X}(t) = \begin{bmatrix} \mathbf{x}(t - T + 1) & \mathbf{x}(t - T + 2) & \cdots & \mathbf{x}(t-1) & \mathbf{x}(t) \end{bmatrix} \quad (22)$$

where $\mathbf{x}(t) = \begin{bmatrix} x_1(t) & x_2(t) & \cdots & x_N(t) \end{bmatrix}^{\mathrm{T}}$ is the measurement data at time t.

In order to add white Gaussian noise into the original data source $\mathbf{O}$ to obtain the raw data source $\mathbf{\Omega}$, the signal-to-noise ratio (SNR) is defined as

$$\mathrm{SNR} = 10\mathrm{lg}(\frac{\mathrm{Tr}\left(\mathbf{OO}^{\mathrm{T}}\right)}{\mathrm{Tr}\left(\mathbf{GG}^{\mathrm{T}}\right) \times m^2}) \quad (23)$$

where $\mathrm{Tr}(\cdot)$ denotes the matrix trace, $\mathbf{G}$ is the white Gaussian noise matrix, whose entries follow the standard normal distribution, m is the magnitude of white Gaussian noise, and $\mathbf{G}^{\mathrm{T}}$ and $\mathbf{O}^{\mathrm{T}}$ are the transposes of $\mathbf{G}$ and $\mathbf{O}$, respectively.

During a period $T_{\max}$, the raw data source $\mathbf{\Omega}$ is obtained through the following formulas:

$$\mathbf{\Omega} = \mathbf{O} + m\mathbf{G} \quad (24)$$

$$m = \sqrt{\frac{\mathrm{Tr}\left(\mathbf{OO}^{\mathrm{T}}\right)}{\mathrm{Tr}\left(\mathbf{GG}^{\mathrm{T}}\right) \times 10^{\mathrm{SNR}/10}}} \quad (25)$$

where $\mathbf{\Omega}$, $\mathbf{O}$, and $\mathbf{G}$ are all $N \times T_{\max}$ matrixes.

As the event indicator, MSR is calculated at each time interval and finally visualized to detect events. For simplicity, we set $L = 1$ when forming the matrix product $\mathbf{Z}$.

To mitigate against false positives, the dynamic Kalman filter also serves as the data conditioner. Thus, after data conditioning, the conditioned raw matrix $\hat{\mathbf{X}}(t)$ is formed as follows:

$$\hat{\mathbf{X}}(t) = \begin{bmatrix} \hat{\mathbf{x}}(t - T + 1) & \hat{\mathbf{x}}(t - T + 2) & \cdots & \hat{\mathbf{x}}(t-1) & \hat{\mathbf{x}}(t) \end{bmatrix} \quad (26)$$

where $\hat{\mathbf{x}}(t) = \begin{bmatrix} \hat{x}_1(t|t) & \hat{x}_2(t|t) & \cdots & \hat{x}_N(t|t) \end{bmatrix}^{\mathrm{T}}$ is the conditioned data at time t. As a result, the procedure of real-time event detection based on random matrix theory and Kalman filtering is shown in Table 1.

Table 1. Procedure of real-time event detection based on random matrix theory and Kalman filtering [30]. (Reproduced with permission from Fujia Han, IEEE Proceedings; published by IEEE 2018.).

Kalman Filter	1: Initialization
	2: Calculate $\hat{x}_i(t\|t)$ using (10), $1 \leq i \leq N$
	3: Update $\mathbf{R}_i(t)$ using (19)–(21)
	4: Update $\mathbf{K}_i(t)$ using (16)
Event Detector	5: Form $\hat{\mathbf{X}}(t)$ using (26)
	6: Transform $\hat{\mathbf{X}}(t)$ to $\mathbf{X}_{\mathrm{s}}(t)$ using (6)
	7: Transform $\mathbf{X}_{\mathrm{s}}(t)$ to $\mathbf{X}_{\mathrm{u}}(t)$ using (7)
	8: Form $\mathbf{z}(t)$ using (2)
	9: Transform $\mathbf{z}(t)$ to $\mathbf{z}_{\mathrm{s}}(t)$ using (8)
	10: Calculate MSR using (5)
	11: If $t < T_{\max}$, repeat 2–10 at the next time $t + 1$; otherwise, go to 12
	12: Visualize MSR

4. Experimental Results and Discussion

In this section, the improved event detection approach is tested and compared with the work [19] using both simulated and real PMU data. The simulated data were generated based on the standard IEEE 118-bus system using MATPOWER [31], while the real PMU data were collected from the GB transmission system. Besides, voltage magnitude (in p.u.) and phase angle (in rad) were used, respectively, to validate the feasibility and effectiveness of the improved event detection approach. The SNRs of voltage magnitude and phase angle were kept the same in the tests. All the experiments were conducted on a common desktop using the matlab programming language.

For all the cases, we set $\Gamma = 0.3\mathbf{I}$ ($\mathbf{I}$ is a 3×3 matrix of 1 s), $\mathbf{Q} = 1 \times 10^{-2}$, $\alpha = 1 \times 10^{-2}$, $\beta = 1 \times 10^6$, $W = 50$, $N = 118$, and $T = 240$.

4.1. Data Conditioning

4.1.1. Noise Reduction

A comparison was performed between the dynamic and original Kalman filters. During the whole period $T_{max} = 1500$, there was a discrete event (an increase of active power demand at bus 60) happening at $t = 601$. In this case, the constant $\mathbf{R}$ of the original Kalman filter was 1×10^{-2}. The results are shown in Table 2.

Table 2. Performance comparison of noise reduction between the dynamic and original Kalman filters at bus 60 ($T_{max} = 1500$).

SNR	RMSE of Voltage Magnitude			SNR	RMSE of Voltage Phase Angle		
	Raw Values	Original KF	Dynamic KF		Raw Values	Original KF	Dynamic KF
5	17.519	8.6245	3.7744	5	6.6478	4.7623	3.8785
15	5.6129	2.7867	2.3817	10	3.5212	2.0497	1.5254
25	1.7739	0.8944	0.8906	15	2.0960	1.0613	0.9556
35	0.5534	0.2807	0.2765	20	1.1541	0.5932	0.5699
45	0.1738	0.0886	0.0877	25	0.6867	0.3389	0.3274
55	0.0550	0.0278	0.0276	30	0.3650	0.1867	0.1762
65	0.0174	0.0090	0.0088	35	0.2182	0.1118	0.1020
75	0.0055	0.0029	0.0027	40	0.1178	0.0663	0.0655
85	0.0018	0.0010	0.0009	45	0.0670	0.0424	0.0409
95	0.0006	0.0006	0.0005	50	0.0357	0.0326	0.0295

In order to eliminate the effect of the error, root mean squared error (RMSE) is calculated during $t = 501$–1500. In Table 2, it is clearly shown that the RMSE value of voltage magnitude of the dynamic Kalman filter is almost the same as that of the original Kalman filter when SNR is greater than or equal to 25, while the RMSE value of voltage magnitude of the dynamic Kalman filter is less than that of the original Kalman filter when SNR is less than or equal to 15. The RMSE value of voltage phase angle follows similar to that of voltage magnitude.

4.1.2. Missing Data Recovery

A further comparison was performed between the dynamic and original Kalman filters in terms of missing data recovery, as shown in Table 3. In this case, the same discrete event happened to bus 60, and RMSE was still calculated during $t = 501$–1500, and the constant $\mathbf{R}$ of the original Kalman filter remained 1×10^{-2}.

In Table 3, for voltage magnitude, it can be seen that the RMSE of the dynamic Kalman filter is significantly lower than that of the original Kalman filter. It is worth noting that as the percentage of missing data goes up, the RMSE of the dynamic Kalman filter only increases slightly from 6×10^{-4} to 13×10^{-4}, while the RMSE of the original Kalman filter increases dramatically from 3.08 to 16.03. However, for voltage phase angle, it is shown clearly in Table 3 that the RMSE of the dynamic Kalman filter is only slightly lower than that of the original Kalman filter. Moreover, when the percentage of missing data rises, both cases go up rather slowly.

Table 3. Performance comparison of missing data recovery between the dynamic and original Kalman filters at bus 60 (T_{max} = 1500).

Percentage of Missing Data	RMSE of Voltage Magnitude (SNR = 95)			RMSE of Voltage Phase Angle (SNR = 35)		
	Raw Values	Original KF	Dynamic KF ($\times 10^{-4}$)	Raw Values	Original KF ($\times 10^{-2}$)	Dynamic KF ($\times 10^{-2}$)
5%	6.96	3.08	6	2.63	18.22	16.82
10%	10.02	4.62	7	3.71	18.27	16.84
15%	12.08	6.17	8	4.62	18.46	17.40
20%	14.12	7.50	9	5.33	18.63	17.87
25%	15.63	8.84	9	5.95	18.78	18.29
30%	17.17	10.37	10	6.52	18.90	18.79
35%	18.46	11.74	11	7.05	19.52	19.11
40%	19.85	12.96	12	7.53	19.88	19.68
45%	21.18	14.41	13	8.01	20.30	20.01
50%	22.11	16.03	13	8.42	20.82	20.71

4.2. Event Detection Using Voltage Magnitudes

Event detection was conducted in situations with heavy noise and missing data using voltage magnitudes. Different signals were used to test the performance of the improved and original event detection approaches. The defined signals are shown in Table 4.

Table 4. Defined signals for event detection.

Signal	Bus Number	Sampling Time	Active Power Demand (MW)
swell and sag signals	60	t = 1–600	fluctuation around 80
		t = 601–700	fluctuation around 120
		t = 701–1050	fluctuation around 80
		t = 1051–1150	fluctuation around 40
		t = 1151–1500	fluctuation around 80
	others	t = 1–1500	no change
multiple signals	60	t = 1–600	fluctuation around 120
		t = 601–700	fluctuation around 150
		t = 701–800	fluctuation around 180
		t = 801–900	fluctuation around 210
		t = 901–1000	fluctuation around 240
		t = 1001–1100	fluctuation around 270
		t = 1101–1200	fluctuation around 300
		t = 1201–1300	fluctuation around 330
		t = 1301–2000	fluctuation around 360
	99	t = 1–600	fluctuation around 60
		t= 601–700	fluctuation around 100
		t = 701–1200	fluctuation around 60
		t = 1201–1300	fluctuation around 20
		t = 1301–2000	fluctuation around 60
	others	t = 1–2000	no change

4.2.1. Heavy Noise

Figure 1 shows the MSR-*t* curves of the improved and original event detectors. It is noted that event detection starts after a short period because the sliding window needs to be filled up with samples and the dynamic Kalman filter takes a short time to track the real values.

In Figure 1a, it can be observed from the MSR-*t* curve of the improved event detector that signals can be detected easily. Specifically, during t = 260–600, MSR remains stable around 0.77, which means that the system state is normal without any signals. Then, at t = 600, MSR starts to decrease gradually until t = 770 (from 0.7733 to 0.7462). After that, MSR begins to increase to the original stable level (0.7674) until t = 940, then MSR remains

almost constant. During this period, a U-shaped curve is observed, which indicates that there are signals changing the system state during t = 601–940. In addition, the effect of a signal on MSR extends T extra instants, due to historical data in the sliding window. As a result, the actual duration of the signals can be calculated as 940 − 601 + 1 − T = 100 instants, and the occurrence time of the signals is the starting time of the U-shaped curve, which is t = 601. That corresponds exactly to the fact that there is an increase of active power demand at bus 60 during t = 601–700. As a result, event detection is conducted in this way. Likewise, there is a second U-shaped curve from t = 1051 to t = 1390, which indicates that the signals occur at t = 1051, and the duration of the signals is 100 instants. However, for the original event detector, signals cannot be detected under the same situation. In other words, there is no clear U-shaped curve during the whole period because MSR changes slightly around 0.86 during t = 600–1390.

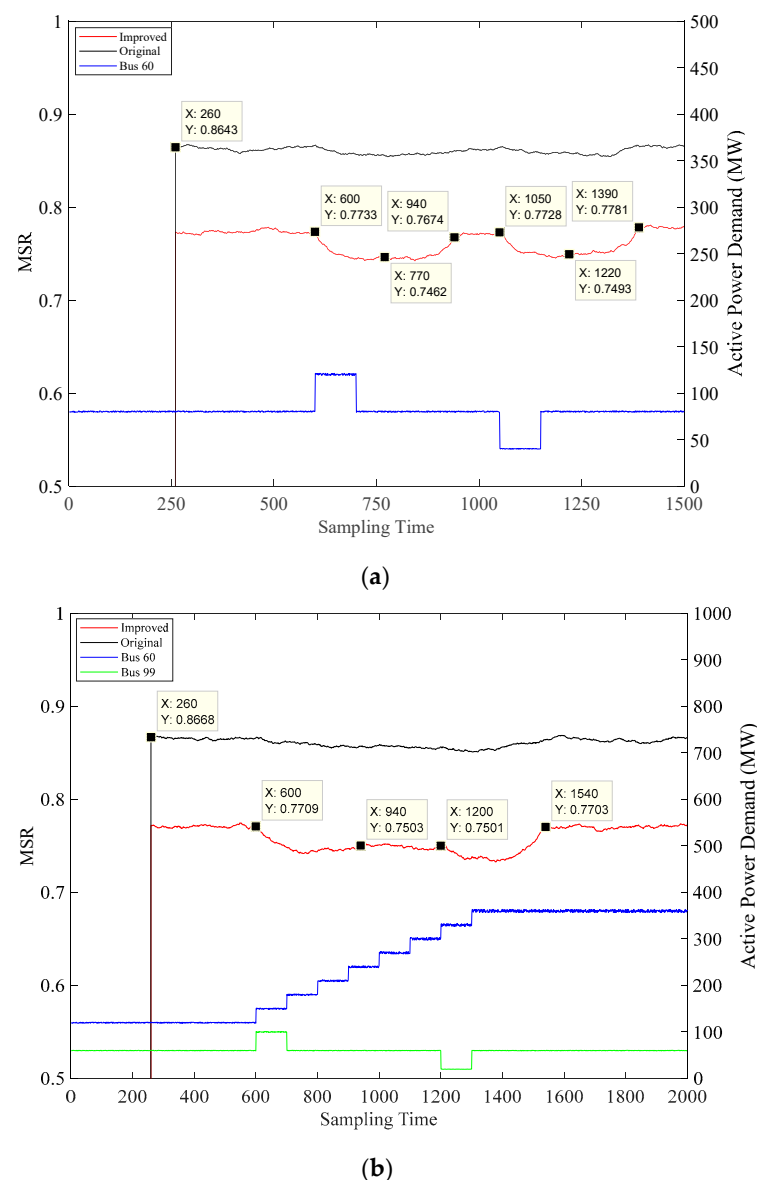**Figure 1.** (**a**) Swell and sag signals (signal-to-noise ratio (SNR) = 75), (**b**) Multiple signals (SNR = 70), mean spectral radius (MSR)-*t* curves of the improved and original event detectors in the situation with heavy noise [30]. (Reproduced with permission from Fujia Han, IEEE Proceedings; published by IEEE, 2018.).

Apart from single signals, multiple signals can be also discovered by the improved event detector in a situation with heavy noise, as shown in Figure 1b. According to the MSR-*t* curve

of the improved event detector, two types of signals can be detected. Firstly, two U-shaped curves are found during t = 601–940 and t = 1201–1540, and it indicates continuous signals during t = 601–700 and t = 1201–1300, respectively. Secondly, a third U-shaped curve is found during t = 601–1540, which indicates continuous signals during t = 601–1300. However, the same conclusion cannot be drawn from the MSR-t curve of the original event detector.

4.2.2. Missing Data

Figure 2 depicts the MSR-t curves of the improved and original event detectors when the percentage of missing data is 30%. As the dynamic Kalman filter also needs to take a short time to track the real values in the situation with missing data, event detection starts after a short period.

(**a**)

(**b**)
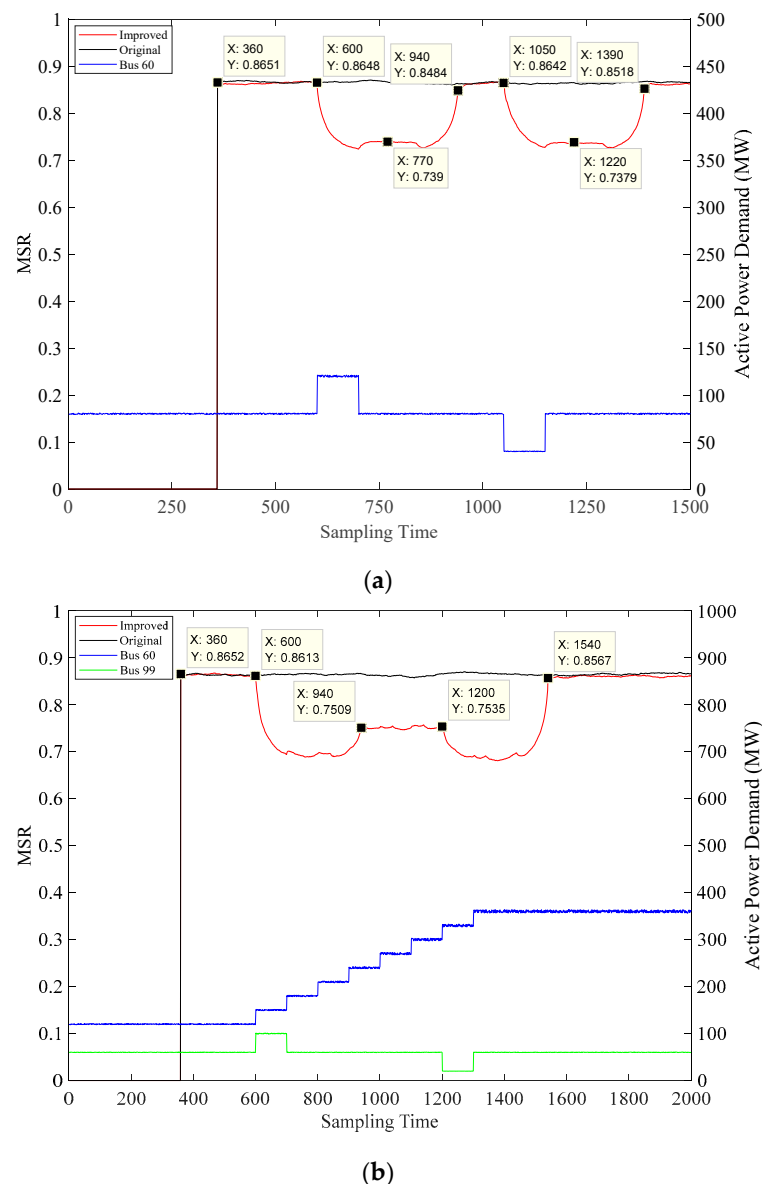
Figure 2. (**a**) Swell and sag signals (SNR = 95), (**b**) multiple signals (SNR = 90), MSR-t curves of the improved and original event detectors in the situation with missing data.

In Figure 2a, according to the MSR-t curve of the improved event detector, signals can be detected in the same way mentioned above. To be specific, two U-shaped curves are observed during t = 601–940 and during t = 1051–1390. However, the original event detector is unable to detect signals under the same situation because U-shaped curves cannot be found during

the signal occurrence period. In addition, as shown in Figure 2b, multiple signals can still be discovered by the improved event detector in the situation with missing data. By contrast, the same signals cannot be discovered by the original event detector.

4.3. Event Detection Using Voltage Phase Angles

Event detection was also conducted using voltage phase angles with respect to the slack bus in situations with heavy noise and missing data. However, as voltage magnitude-based event detection is quite similar to voltage phase angle-based event detection, only swell and sag signals (as shown in Table 4) were employed to test both approaches.

Figure 3a describes the MSR-t curves of the improved and original event detectors when SNR is 20, while Figure 3b shows the MSR-t curves of the improved and original event detectors when SNR is 35 and the percentage of missing data is 30%. In Figure 3a,b, regardless of heavy noise or missing data, a U-shaped curve can be observed from $t = 601$ to $t = 940$ and from $t = 1051$ to $t = 1390$, respectively, for the improved event detector. By contrast, for the original event detector, there is no U-shaped curve found during the whole period in both situations.
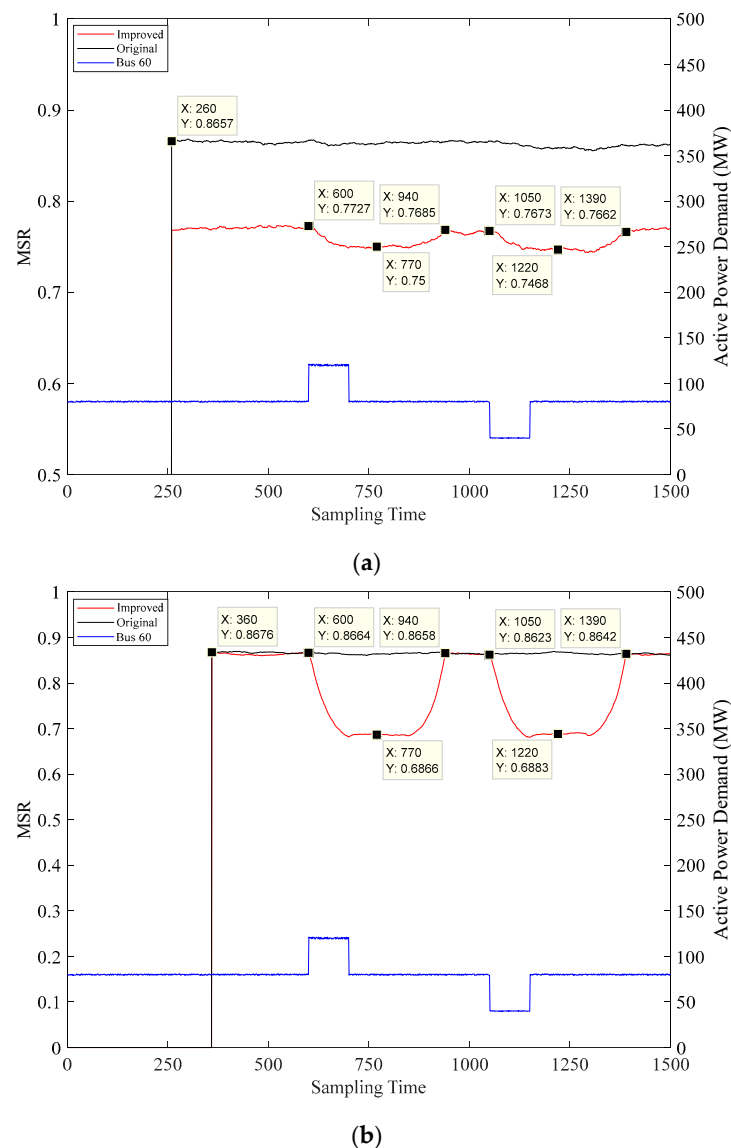

(a)

(b)

Figure 3. (**a**) Swell and sag signals (in the situation with heavy noise), (**b**) swell and sag signals (in the situation with missing data), MSR-t curves of the improved and original event detectors using voltage phase angles [30]. (Reproduced with permission from Fujia Han, IEEE Proceedings; published by IEEE, 2018.).

4.4. Event Detection Using Real PMU Data

Real PMU data were also utilized to further validate the robustness of the proposed event detection approach. The PMU data were collected from 13 PMUs of the GB transmission system [32,33], and the sampling rate of the PMUs was 50 Hz. Besides, the total time length of the PMU data was 1 min, and thus there were $50 \times 60 = 3000$ samples in total for each PMU. What is more, the missing samples in the PMU data accounted for 12.5%. In this case, N was 13 and T was 500. The results are shown in Figures 4 and 5.

(a)

(b)
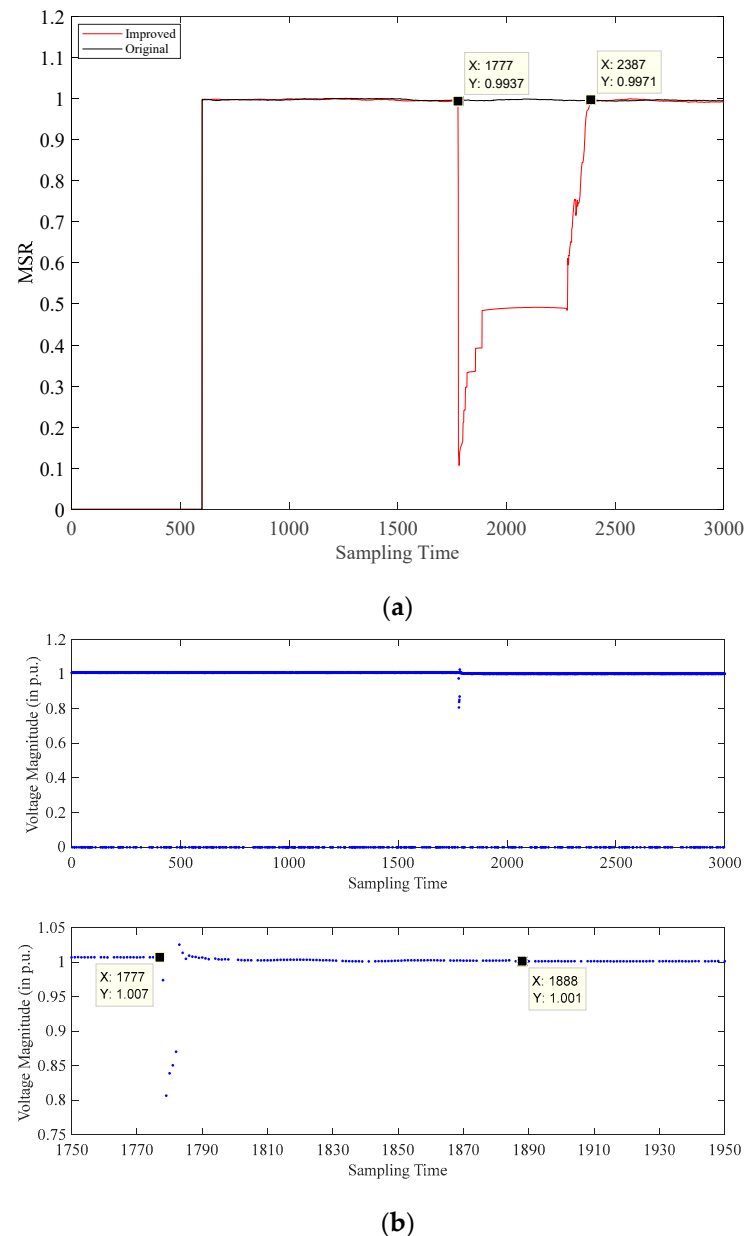
Figure 4. (**a**) MSR-t curves of the improved and original event detectors, (**b**) voltage magnitude of a certain PMU, event detection using voltage magnitudes of real PMU data.

In Figure 4a or Figure 5a, it can be observed that there is a U-shaped curve during $t = 1777$–2387 in the MSR-t curves of the improved event detector. This indicates that there are signals occurring at $t = 1778$ and lasting for $2387 - 1778 + 1 - T = 110$, which is $110/50 = 2.2$ s. According to Figure 4b or Figure 5b, the voltage magnitude and phase angle of a certain PMU change dramatically during $t = 1778$–1887, which confirms the above inference. In fact, there is indeed a tripping event from $t = 1778$ to $t = 1887$ across

the GB transmission system. However, the original event detector fails to detect this event since no U-shaped curve is seen throughout the period.
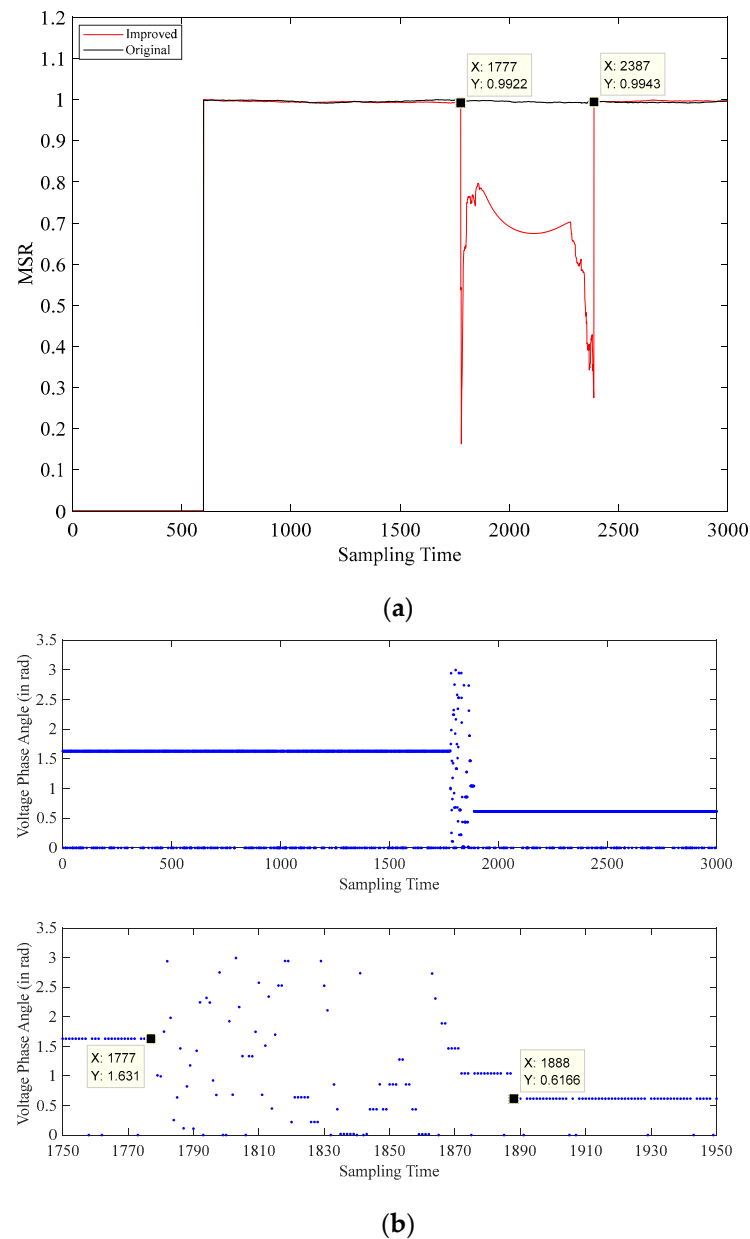

(a)

(b)

Figure 5. (**a**) MSR-t curves of the improved and original event detectors, (**b**) voltage phase angle of a certain PMU, event detection using voltage phase angles of real PMU data.

5. Conclusions

This paper presents a robust event detection approach based on random matrix theory and Kalman filtering. In addition, as the data conditioner, a dynamic Kalman filtering technique was developed through the adjustment of the measurement noise covariance matrix in order to reduce noise and recover missing samples in PMU data. The experimental results show that the dynamic Kalman filter outperforms the original one, in terms of both noise reduction and missing data recovery. Furthermore, the improved event detection approach is much more robust than the original one, especially in practical situations where PMU data exhibit significant noise levels and large numbers of missing samples. The proposed event detection approach is also capable of performing computationally efficient event detection in real time in practice, without pre-training a model.

In this paper, the proposed approach only aimed to identify the occurrence time and the duration period of an event, without locating and classifying the event. Hence, in future research, the authors will focus on the location and classification of events to further develop and demonstrate the proposed robust approach in terms of event detection.

References

1. Mousavi-Seyedi, S.S.; Aminifar, F.; Afsharnia, S. Application of WAMS and SCADA data to online modeling of series-compensated transmission lines. *IEEE Trans. Smart Grid* **2017**, *8*, 1968–1976. [CrossRef]
2. Kekatos, V.; Giannakis, G.B.; Wollenberg, B. Optimal placement of phasor measurement units via convex relaxation. *IEEE Trans. Power Syst.* **2012**, *27*, 1521–1530. [CrossRef]
3. Liu, J.; Tang, J.; Ponci, F.; Monti, A.; Muscas, C.; Pegoraro, P.A. Trade-offs in PMU deployment for state estimation in active distribution grids. *IEEE Trans. Smart Grid* **2012**, *3*, 915–924. [CrossRef]
4. Li, Q.; Cui, T.; Weng, Y.; Negi, R.; Franchetti, F.; Ilic, M.D. An information-theoretic approach to PMU placement in electric power systems. *IEEE Trans. Smart Grid* **2013**, *4*, 446–456. [CrossRef]
5. Maji, T.K.; Acharjee, P. Multiple solutions of optimal PMU placement using exponential binary PSO algorithm for smart grid applications. *IEEE Trans. Ind. Appl.* **2017**, *53*, 2550–2559. [CrossRef]
6. Lin, W.; Yang, C.; Lin, J.; Tsay, M. A fault classification method by RBF neural network with OLS learning procedure. *IEEE Trans. Power Deliv.* **2001**, *16*, 473–477. [CrossRef]
7. Vasilic, S.; Kezunovic, M. Fuzzy ART neural network algorithm for classifying the power system faults. *IEEE Trans. Power Deliv.* **2005**, *20*, 1306–1314. [CrossRef]
8. Nguyen, D.; Barella, R.; Wallace, S.A.; Zhao, X.; Liang, X. Smart grid line event classification using supervised learning over PMU data streams. In Proceedings of the 2015 Sixth International Green and Sustainable Computing Conference (IGSC), Las Vegas, NV, USA, 14–16 December 2015.
9. Adhikari, U.; Morris, T.; Pan, S. Applying Hoeffding adaptive trees for real-time cyber-power event and intrusion classification. *IEEE Trans. Smart Grid* **2018**, *9*, 4049–4060. [CrossRef]
10. Xie, L.; Chen, Y.; Kumar, P.R. Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis. *IEEE Trans. Power Syst.* **2014**, *29*, 2784–2794. [CrossRef]
11. Rafferty, M.; Liu, X.; Laverty, D.M.; McLoone, S. Real-time multiple event detection and classification using moving window PCA. *IEEE Trans. Smart Grid* **2016**, *7*, 2537–2548. [CrossRef]
12. Guo, Y.; Li, K.; Laverty, D.M.; Xue, Y. Synchrophasor-based islanding detection for distributed generation systems using systematic principal component analysis approaches. *IEEE Trans. Power Deliv.* **2015**, *30*, 2544–2552. [CrossRef]
13. Ge, Y.; Flueck, A.J.; Kim, D.; Ahn, J.; Lee, J.; Kwon, D. Power system real-time event detection and associated data archival reduction based on synchrophasors. *IEEE Trans. Smart Grid* **2015**, *6*, 2088–2097. [CrossRef]
14. Liu, X.; Laverty, D.M.; Best, R.J.; Li, K.; Morrow, D.J.; McLoone, S. Principal component analysis of wide-area phasor measurements for island detection-A geometric view. *IEEE Trans. Power Deliv.* **2015**, *30*, 976–985. [CrossRef]
15. Kim, D.; Chun, T.; Yoon, S.; Lee, G.; Shin, Y. Wavelet-based event detection method using PMU data. *IEEE Trans. Smart Grid* **2017**, *8*, 1154–1162. [CrossRef]
16. Kavasseri, R.G.; Cui, Y.; Brahma, S.M. A new approach for event detection based on energy functions. In Proceedings of the IEEE PES General Meeting, National Harbor, MD, USA, 27–31 July 2014; pp. 1–5.
17. Lin, Z.; Xia, T.; Ye, Y.; Zhang, Y.; Chen, L.; Liu, Y.; Tomsovic, K.; Bilke, T.; Wen, F. Application of wide area measurement systems to islanding detection of bulk power systems. *IEEE Trans. Power Syst.* **2013**, *28*, 2006–2015. [CrossRef]
18. Liang, X.; Wallace, S.A.; Nguyen, D. Rule-based data-driven analytics for wide-area fault detection using synchrophasor data. *IEEE Trans. Ind. Appl.* **2017**, *53*, 1789–1798. [CrossRef]
19. He, X.; Chu, L.; Qiu, R.C.; Ai, Q.; Lin, Z. A novel data-driven situation awareness approach for future grids-Using large random matrices for big data modeling. *IEEE Access* **2018**, *6*, 13855–13865. [CrossRef]

20. Zhang, J.; Welch, G.; Bishop, G.; Huang, Z. A two-stage Kalman filter approach for robust and real-time power system state estimation. *IEEE Trans. Sustain. Energy* **2014**, *5*, 629–636. [CrossRef]

21. Gao, P.; Wang, M.; Ghiocel, S.G.; Chow, J.H.; Fardanesh, B.; Stefopoulos, G. Missing data recovery by exploiting low-dimensionality in power system synchrophasor measurements. *IEEE Trans. Power Syst.* **2016**, *31*, 1006–1013. [CrossRef]

22. Xu, X.; He, X.; Ai, Q.; Qiu, R.C. A correlation analysis method for power systems based on random matrix theory. *IEEE Trans. Smart Grid* **2017**, *8*, 1811–1820. [CrossRef]

23. He, X.; Ai, Q.; Qiu, R.C.; Huang, W.; Piao, L.; Liu, H. A big data architecture design for smart grids based on random matrix theory. *IEEE Trans. Smart Grid* **2017**, *8*, 674–686. [CrossRef]

24. Karimipour, H.; Dinavahi, V. Extended Kalman filter-based parallel dynamic state estimation. *IEEE Trans. Smart Grid* **2015**, *6*, 1539–1549. [CrossRef]

25. Netto, M.; Zhao, J.; Mili, L. A robust extended Kalman filter for power system dynamic state estimation using PMU measurements. In Proceedings of the 2016 IEEE Power and Energy Society General Meeting, Boston, MA, USA, 17–21 July 2016.

26. Ghahremani, E.; Kamwa, I. Local and wide-area PMU-based decentralized dynamic state estimation in multi-machine power systems. *IEEE Trans. Power Syst.* **2016**, *31*, 547–562. [CrossRef]

27. Jones, K.D.; Pal, A.; Thorp, J.S. Methodology for performing synchrophasor data conditioning and validation. *IEEE Trans. Power Syst.* **2015**, *30*, 1121–1130. [CrossRef]

28. Gao, F.; Thorp, J.S.; Pal, A.; Gao, S. Dynamic state prediction based on auto-regression model using PMU data. In Proceedings of the IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, USA, 24–25 February 2012; pp. 1–5.

29. Mahmood, F.; Hooshyar, H.; Vanfretti, L. Extracting steady state components from synchrophasor data using Kalman filters. *Energies* **2016**, *9*, 315. [CrossRef]

30. Han, F.; Taylor, G.; Li, M. Towards a data driven robust event detection technique for smart grids. In Proceedings of the 2018 IEEE Power & Energy Society General Meeting (PESGM), Portland, OR, USA, 5–10 August 2018.

31. Zimmerman, R.D.; Murillo-Sánchez, C.E.; Thomas, R.J. MATPOWER: Steady-state operations, planning and analysis tools for power systems research and education. *IEEE Trans. Power Syst.* **2011**, *26*, 12–19. [CrossRef]

32. Ashton, P.M.; Saunders, C.S.; Taylor, G.A.; Carter, A.M.; Bradley, M.E. Inertia estimation of the GB power system using synchrophasor measurements. *IEEE Trans. Power Syst.* **2015**, *30*, 701–709. [CrossRef]

33. Ashton, P.M.; Taylor, G.A.; Irving, M.R.; Pisica, I.; Carter, A.M.; Bradley, M.E. Novel application of detrended fluctuation analysis for state estimation using synchrophasor measurements. *IEEE Trans. Power Syst.* **2013**, *28*, 1930–1938. [CrossRef]