Multiple Event Detection and Recognition through Sparse Unmixing for High-Resolution Situational Awareness in Power Grid

Wei Wang, Student Member, IEEE, Li He, Student Member, IEEE, Penn Markham, Student Member, IEEE, Hairong Qi, Senior Member, IEEE, Yilu Liu, Fellow, IEEE, Qing Cao, Member, IEEE, Leon Tolbert, Fellow, IEEE

Abstract—A situational awareness system is essential to provide accurate understanding of power system dynamics, such that proper actions can be taken in real time in response to system disturbances and to avoid cascading blackouts. Event analysis has been an important component in any situational awareness system. However, most state-of-the-art techniques can only handle single event analysis. This paper tackles the challenging problem of multiple event detection and recognition. We propose a new conceptual framework, referred to as event unmixing, where we consider real-world events mixtures of more than one constituent root event. This concept is a key enabler for analysis of events to go beyond what are immediately detectable in a system, providing high-resolution data understanding at a finer scale. We interpret the event formation process from a linear mixing perspective and propose an innovative nonnegative sparse event unmixing (NSEU) algorithm for multiple event separation and temporal localization. The proposed framework has been evaluated using both PSS/E simulated cases and real event cases collected from the frequency disturbance recorders (FDRs) of the Frequency Monitoring Network (FNET). The experimental results demonstrate that the framework is reliable to detect and recognize multiple cascading events as well as their time of occurrence with high accuracy.

Index Terms—Wide-area situational awareness, Linear unmixing, Nonnegative sparsity constraint, Event detection and recognition, Power grid.

I. INTRODUCTION

It has become essential that wide-area situational awareness (WASA) systems can enable the monitoring of bulk power systems and provide critical information for understanding and responding to system disturbances and cascading blackouts. Work first began in the 1980s on closely synchronized measurements that would allow direct measurement of voltage phase angle at transmission level. As a result, Phasor Measurement Units (PMU) [1], [2] have been gradually installed in substations that measure phasors at high voltage levels. As a member of the PMU family, the Frequency Disturbance Recorder (FDR) was developed at Virginia Tech in 2003. The FDR collects instantaneous voltage phasor and frequency measurements at low-voltage distribution level using ordinary 120-V wall outlets. It then transmits the measured frequency data remotely via the Internet. Based on these low-cost FDRs, a US-wide Frequency Monitoring Network (FNET) has been implemented [3]–[5]. FNET serves the entire North American power grid through advanced situational awareness techniques, including, for example, real-time event alert, accurate event location estimation, animated event visualization, and post event analysis [6].

When an event occurs in a power grid, a sudden imbalance between generation and load consumption causes frequency changes within the system that can be used as an indicator for event disturbance. There have been a couple of works reported so far conducting event analysis using data collected from FNET [6]–[11]. Although successful, these state-of-the-art techniques can only handle disturbances caused by a single event. If multiple cascading events are involved, existing techniques can only detect frequency disturbances caused by the initial event, and the frequency disturbances from successive events might be overshadowed by continued frequency fluctuation from the initial event. The system disturbance reports from the North American Electric Reliability Corporation (NERC) [12] have made it obvious that major disturbances typically involve a number of unlikely, unplanned events. When multiple events occur in cascading fashion, the electromechanical waves generated will interfere with each other, and the measurement taken at a FDR would more than likely be a “mixture” of multiple frequency signals. Thus, how to determine the number of events that occurred and identify the types of events that occurred with precise estimation of occurring time using simply the observed mixture is a very challenging problem.

In this paper, we focus on the problem of multiple events detection, recognition, and temporal localization. The main contributions of the paper are threefold: 1) the formulation of the multiple event analysis problem using a linear mixing model, 2) the construction of the signature dictionary that incorporates temporal information to reflect the event dynamics in power grid, and 3) the validation of the proposed approaches using extensive simulations and real case studies.

The rest of the paper is organized as follows. Section II formulates the event unmixing problem. Section III explains the rationale behind the linear mixing model. Section IV elaborates on the construction of the root event signature dictionary. Section V describes the nonnegative sparsity constrained unmixing algorithm and detection strategies. Sections VI and VII conduct performance evaluation using both simulated cases produced by PSS/E and real multiple event data recorded from FNET. Finally, we conclude the paper and discuss future work in Section VIII.
II. PROBLEM FORMULATION

Mixed measurements are frequently encountered in real-world applications. Because of the resolution issue associated with discrete sampling and the effect of unknown sources, the measurements can rarely be pure. The existence of mixed measurements has brought the decomposition or unmixing techniques to a wide array of applications. For example, in the field of remote sensing, due to the large footprint, a single pixel usually covers more than one type of ground constituent. Thus, the measured spectrum at a single pixel is a mixture of several ground cover spectra, where the pixel unmixing technique has been applied to subpixel object quantification [13], mineral identification [14], area estimation [15], etc. Another emerging application is in biological microscopy where multispectral fluorescence microscopy is analyzed to discriminate different co-localized fluorescent molecules with highly overlapping spectra [16]. Using common microscopy methods, the number of molecules that can be detected simultaneously is limited by both spectral and spatial overlap. These issues can be tackled by applying linear unmixing, which extends the possibilities to distinguish multiple proteins, organelles or functions within a single cell [17].

Similar to the ubiquitous existence of mixed measurements, events might not occur in a pure and isolated fashion, especially in power grid. Taking the major U.S. western blackout in 1996 as an example, at the very beginning of the blackout, two parallel lines were tripped due to a fault and mis-operation of the protective equipment, and consequently some generation was tripped as a special protection system (SPS) response. Then, a third line was disconnected due to bad connectors in a distance relay. After about 20 seconds of these events, the last straw of the collapse occurred when the Mill Creek -Antelope line tripped due to an undesired operation of a protective relay. After this line trip, the system collapsed within three seconds. Therefore, to be able to provide high-resolution understanding of the power system dynamics, it is very essential to perform multiple event analysis.

Typical power system disturbances fall into one of the four categories, including generator trip (GT), load shedding (LS), line trip (LT), and oscillation (OS), which we refer to as the “root events” or “pure events”. See Fig. 1 for the frequency variation patterns when each of the root events occurs. Denote the frequency signal collected at a FDR as \( x \) (also referred to as the measured signal) and the frequency variation pattern of each root event as \( s \) (also referred to as the source signal). We propose a new conceptual framework for the study of multiple event analysis, where we consider the disturbances sensed at each FDR, \( x \), as a linear mixture of a limited number of root events \( s \). The formulation can thus be expressed as

\[
x = S\alpha + \epsilon
\]  

where \( x \in \mathbb{R}^l \) is the measured mixture and \( l \) is the number of measurements corresponding to the time over which an event is measured. \( S \in \mathbb{R}^{l \times c} \) is the root event signature dictionary whose columns, \( \{s_j\}^c_{j=1} \in \mathbb{R}^l \), correspond to different pattern profiles of root events, and \( \alpha \) is the mixing coefficient vector. The possible error and noise are taken into account by the \( l \)-dimensional column vector \( \epsilon \). This concept is the key to accurate event analysis that goes beyond immediately detectable information in a system and uncovers the true cause(s) of multiple cascading events. To realize this framework, however, we have to answer the following questions:

1) Is it valid to use a linear mixing model to formulate the mixture sensed at a FDR?
2) How can one obtain the profiles of root events given the complexity and dynamic nature of the power grid?
3) How can one incorporate the different starting times of cascading root events in the construction of \( S \)? and
4) How can one solve \( \alpha \) for event detection and recognition purposes in an on-line fashion?

We defer the discussions of the first issue to Section III and the second issue to Section IV. To resolve the other two issues, we construct an overcomplete signature dictionary to incorporate pattern profiles of various types of root events as well as temporal information on how events cascade. In this way we can simultaneously detect event type and event starting time in one unmixing procedure. In addition, since the number of events is much less than the number of pattern profiles in the dictionary, the “sparsity” enforced on the coefficient vector \( \alpha \) becomes an appropriate constraint that not only reduces the solution space, making the problem well-posed, but it also helps in the identification of event type and precise starting time. The sparse representation of the measured mixture is achieved through solving an \( \ell^1 \)-regularized least squares problem, which can be done efficiently through convex optimization.

To further improve the robustness, we also enforce the “non-negativity” constraint on the coefficient vector \( \alpha \). There are two reasons for adding this constraint: first, from the energy perspective, physical laws govern the power flow in a grid where the mixture of electromechanical waves generated by multiple disturbances should be only an additive combination of the constituent components; second, since generator trips generally cause a decrease in frequency while load sheddings generally cause an increase in frequency, the nonnegativity constraint on \( \alpha \) is especially helpful to eliminate error cases where a load shedding event is mistakenly recognized as a generator trip in the reverse way.

![Fig. 1. Four categories of typical root events in power grid.](image-url)
Based on the discussion above, we mathematically formulate the event unmixing problem as below:

$$\min \| \alpha \|_0 \text{ s.t. } \| x - S\alpha \|_2^2 \leq \epsilon, \; \alpha \geq 0$$

(2)

where $\| \alpha \|_0$ represents $\ell_0$-norm, which is defined as the number of non-zero entries in vector $\alpha$. The proposed cost function consists of two components with one measuring approximation error of the linear mixing model (i.e., $\| x - S\alpha \|_2^2$) and the other measuring sparsity of the coefficient vector, $\alpha$. We refer to the proposed algorithm as nonnegative sparse event unmixing (NSEU) for detection, recognition, and temporal localization of multiple cascading events in power grid. In the following sections, we will discuss solutions to the four questions raised in this section.

III. LINEAR MIXING MODEL

In a power grid, since the total frequency change is proportional to the total generation/load loss [18], it is reasonable to approximate the mixture of frequency signals as a linear addition of frequency signals caused by each individual event.

We validate the linearity of the mixing process through two experiments carried out using the PSS/E simulation. PSS/E is a commercial software for simulating, analyzing, and optimizing power system performance based on concrete system configuration, thus it can precisely reflect the dynamics of a real system. In the first experiment, we conduct two simulations, Simu1 and Simu2. In Simu1, a generator producing 110 MW power is tripped (GT1) at the $10^{th}$ second, producing the frequency signal $x_1$ with about $2.3 \times 10^{-3}$ Hz’s frequency drop. In Simu2, a generator producing 778 MW power is tripped (GT2) at the $30^{th}$ second, producing the frequency signal $x_2$ with about $15 \times 10^{-3}$ Hz’s frequency drop. In the second experiment, we conduct only one simulation, Simu3, where the two generator trips are programmed as two cascading events occurred at the $10^{th}$ and the $30^{th}$ second, sequentially, producing the frequency signal $x_3$. The validation is to see if $x_1 + x_2$ approximates $x_3$. As shown in Fig. 2, the top sub-figure displays measured signals from the two single events, $x_1$ and $x_2$, and the bottom sub-figure compares measured signal $x_3$ and summation of $x_1 + x_2$. It is easily observed that $x_3$ accurately approximates $x_1 + x_2$ with the root mean squared error (RMSE) being only $2.08 \times 10^{-4}$.

Therefore, it is reasonable to formulate the measured mixture at a FDR using a linear mixing model. The effectiveness of the model will be further demonstrated using both simulated and real case studies, as discussed in Sections VI and VII.

IV. SIGNATURE DICTIONARY CONSTRUCTION

Both the second and the third questions raised in Sec. II regarding the formulation of event unmixing are related to construction of the root event signature dictionary $S$. The dictionary construction involves two steps. First, the root event patterns are learned from the training data previously collected from FDRs recording disturbances experienced during single events. Second, temporal information (i.e., event starting time) is incorporated into dictionary $S$ by augmentation and padding with the learned root event patterns.

A. Learning Root Event Patterns

As mentioned in Sec. II, typical power system disturbances (events) fall into one of four categories, including generator trip (GT), load shedding (LS, or load drop), line trip (LT), and oscillation (OS). As described in [8], [19], [20], events of the same category generally share similar characteristics. For example, a generation trip always starts with a rapid frequency drop while a load shedding always starts with a frequency increase. Meanwhile, events of the same category also contain a certain degree of differences because of the different setups of intrinsic parameters, including power flow output, consumption, etc. It is impractical to include every single event pattern in the dictionary which would affect the performance of online processing, especially for large-scale systems, like the power grid. We, instead, resort to machine learning approaches to find a representative set of root event patterns for each event category. Hereinafter, we refer to these root event patterns as “root-patterns”.

There are various methods that can be adopted for learning the root-patterns, including, for example, K-means clustering and dictionary learning [21], [22]. The K-means clustering is chosen in order to preserve the physical meaning of each element (or column) of the dictionary such that it represents a signature profile of an actual root-pattern, facilitating the event detection and recognition process. Given a set of single event observations $(v_1, v_2, \ldots, v_n)$ of the same event category collected from either recordings of a real system (e.g., FNET) or simulation, K-means clustering aims to partition the $n$ observations into $K$ subsets, so as to minimize the within-cluster sum-of-squares error:

$$\arg \min_{\Phi_E} \sum_{i=1}^{K} \sum_{v_j \in E_i} \| v_j - e_i \|^2$$

(3)

where $\Phi_E = \{ e_1, e_2, \ldots, e_K \}$ refers to the set of learned root-patterns and $e_i$ is the centroid of cluster $E_i$. Notice that before performing K-means clustering on $(v_1, v_2, \ldots, v_n)$, each vector should be normalized to have unit $\ell^2$-norm, such that the scale ambiguity [23], [24] in learning root-patterns is eliminated. In addition, when performing K-means clustering, instead of directly using the mean vectors as cluster centroids, we select the nearest observation $v_i$ as the cluster centroid.
for the $i$th cluster in the final loop. In this way, the learned root-patterns are closer to real data.

In the context of this paper, we expect to detect and recognize three different types of events, i.e., GT, LS, and LT. If we set $K=6$ for K-means clustering, then the number of root-patterns will be $3 \times 6 = 18$.

B. Incorporating the Temporal Information

After learning the set of root-patterns from each category, the next step is to incorporate the temporal information into construction of the root event signature dictionary, $S$. We refer to each column of $S$ that already embeds temporal information as the "temporal root-pattern". The reason for doing this is that any event, be it single or consisting of cascading events, can start at any time and can last for a different period of time. To ensure that the unmixing algorithm always captures the entire duration of the root event for better recognition accuracy, the dictionary needs to contain the root-patterns starting at all possible sample points from the first one to the $t_s \times \omega$. For instance, for a root-pattern $\alpha$ and use $\omega$ for event detection – coefficients of non-zero mixing coefficients in $\alpha$, reducing the false-alarm rate. In the following, we detail the construction process of the temporal root-patterns.

Suppose the root-patterns last for at most $t_s$ seconds and the sampling frequency is $\omega$, then each root-pattern would have $t_s \times \omega$ samples. For an observation vector $x$ acquired from a FDR that lasts for $t_c$ seconds, it can have $t_c \times \omega$ samples starting from the pre-event steady state, like 60 Hz, to the post-event steady state. The dimension of $t_c \times \omega$ is always larger than or equal to that of $t_s \times \omega$ since $t_c \geq t_s$ is always true, i.e., a multiple-event observation always consists of at least one root event.

From the temporal perspective, each root-pattern can start at all possible sample points from the first one to the $(t_c - t_s) \times \omega$th during the time period $t_c$. To obtain the temporal root-patterns, we construct $(t_c - t_s) \times \omega$ number of profiles for each root-pattern by shifting it from the first sample to the $(t_c - t_s) \times \omega$th sample. For instance, denote a root-pattern as $\mathcal{R}(k), k = 1, \ldots, t_s \times \omega$. If we want to construct one temporal root-pattern $\mathcal{T}(k), k = 1, \ldots, t_c \times \omega$ occurring at the $i$th sample point for $\mathcal{R}(k)$, then $\mathcal{T}(k)$ can be calculated as:

$$\mathcal{T}(k) = \begin{cases} \mathcal{R}(k) \otimes \delta(k-i) & \text{for } 1 \leq k < i + t_s \times \omega \\ \mathcal{R}(t_s \times \omega) & \text{for } i + t_s \times \omega \leq k \leq t_c \times \omega \end{cases}$$

(4)

where $\delta(k-i), k = 1, \ldots, t_c \times \omega$ is the Dirac delta function and $\otimes$ denotes the convolution process. Because we only analyze frequency fluctuation of the signal for event unmixing purpose, the starting value of $\mathcal{R}(k)$ is 0 Hz (by removing the base frequency 60 Hz). Therefore, the function of (4) actually pads the samples before the selected $i$th sample with zero and the samples beyond $t_s \times \omega + i$ with tail stable value of $\mathcal{R}(k)$ to form a temporal root-pattern. The whole formation of a group of temporal root-patterns from one root-pattern is illustrated in Fig. 3.

V. NONNEGATIVE SPARSE LINEAR UNMIXING

Once we have constructed the overcomplete dictionary $S$, we can apply it to estimate the coefficient vector $\alpha$ and use $\omega$ for event detection – coefficients of non-zero value indicate that the corresponding temporal root-pattern in the dictionary should be used to reconstruct the observation mixture, $x$. Since the temporal root-pattern indicates event’s pattern occurred at a certain time, by deriving $\alpha$, we can detect existence of multiple cascading events as well as their temporal locations (or starting time). Given the overcomplete dictionary, traditional methods for coefficient estimation such as the Fully Constrained Least Squares (FCLS) [25] or the Nonnegatively Constrained Least Squares (NCLS) [26] would not work, since the estimated coefficient by FCLS or NCLS may have non-zero values on each root-pattern which would not serve as suitable methods for detection purposes. Recall that the number of constituent events involved in a cascading incident is generally small.

On the other hand, recent developments on sparse coding techniques [27]–[29] would provide good solution to the proposed NSEU algorithm in (2). The sparse coefficient vector produced by sparse coding is confined by the sparsity constraint that only a few elements of the vector are non-zero,
indicating that only a few temporal root-patterns would be utilized to reconstruct the original signal \( x \). In the application of event detection, it is equivalent to say that only a limited number of events have occurred. Although the sparse optimization problem in (2) is NP-hard in general, Donoho [30] suggested that as long as the desired coefficient vector \( \alpha \) is sufficiently sparse, it can be efficiently recovered by instead minimizing the \( L_1 \)-norm, as follows:

\[
\min \| \alpha \|_1 \text{ s.t. } \| S\alpha - x \|_2^2 \leq \epsilon, \alpha \geq 0
\]

(5)

Eq. (5) can also be expressed as

\[
\alpha = \arg \min_{\alpha} \| x - S\alpha \|_2^2 + \lambda \| \alpha \|_1, \alpha \geq 0
\]

(6)

where \( \lambda \) balances the sparsity of the solution and the fidelity of the approximation to \( x \). The “feature-sign search algorithm” in [29] is revised to solve this sparse coding problem with the nonnegativity constraint added.

Event detection is mainly based on studying the non-zero coefficients in the sparse coefficient vector, \( \alpha \), as each coefficient corresponds to the weight of each atom (or column vector in the dictionary) in forming the observation vector \( x \). The larger the coefficient, the more contribution the root event has in making the mixture. We apply an empirically-determined threshold on \( \alpha \) where coefficients above the threshold would correspond to the temporal root-patterns that occurred at a certain time.

VI. EVALUATION WITH SIMULATED DATA

This section conducts a series of experiments to demonstrate the effectiveness of the proposed NSEU algorithm using simulated data. The simulations are done on a small synthetic power grid model “savnw” using the software “Power System Simulator for Engineering (PSS/E)” [31]. The grid model “savnw” is an example bench case supplied with PSS/E, whose configuration is illustrated in Fig. 4. As shown in the figure, the model includes 6 generators, 17 branch lines, 7 loads, and 21 buses. Each type of single event will cause the system to arrive at a new steady state within 30 seconds. We thus use 30-second frequency fluctuation of single event for learning root-patterns. Because this power grid model is quite small, we simply select 4 generator trips (GT) and 1 load shedding (LS) as root-patterns. As for line trips (LT), we apply the K-means method to extract 5 root-patterns using the training data collected from 17 lines. Since we have not collected data with oscillation, we omit the oscillation event for simplicity.

The dictionary \( S \) is thus built based on the 10 extracted root-patterns. Suppose a multiple event case would last for 40 seconds and the sampling rate is set to 60 Hz, the dictionary will then include \((40 - 30) \times 60 \text{ Hz} \times 10 = 6,000 \) temporal root-patterns. We also can improve the on-line performance by decreasing the number of temporal root-patterns in \( S \) by shifting at every two or more samples, at the sacrifice of losing a bit of temporal localization accuracy. In each test, the regularization parameter of NSEU is selected as \( \lambda = 0.1 \) and the threshold is selected as 0.034.

Based on the grid model “savnw”, we have conducted experiments on 8 cases of single event (S1C), 10 cases of multiple event with two constituent components (M2C), and 5 cases of multiple event with three constituent components (M3C). We use four metrics to measure performance, including detection accuracy, false alarm rate, root-pattern recognition rate (R-P Recog), and occurrence time deviation from the ground truth (OT-Deviation), as shown in Table I.

Both detection accuracy and false alarm rate are used to evaluate detection performance of the NSEU algorithm. The detection accuracy calculates the ratio between the number of correctly detected root events and the number of total root events according to the ground truth. The false alarm rate calculates the ratio between the number of detected root events not really happen and the total number of root events according to the ground truth. As indicated in Table I, we can find that the proposed NSEU algorithm can detect all constituent events with 100% accuracy while with very low rate of false alarms.

The root-pattern recognition rate (R-P Recog) is used to evaluate the recognition performance. It calculates the ratio between the number of correctly identified events (i.e., events with correct type of root-pattern) and the number of correctly detected events. Table I indicates the averaged root-pattern recognition rates are very high.

The deviation between the detected occurring time and the ground truth (OT-Deviation) is used to evaluate the temporal localization performance. As shown in Table I, the averaged OT-Deviation values are all very small, indicating the high accuracy of temporal localization.

In addition, the results also show that the NSEU algorithm has strong capability of signal reconstruction based on the linear mixing model, where the averaged reconstruction MSE of the 23 example cases is \( 1.798 \times 10^{-5} \).

We also visually show one simulated example case of three cascading events: LT 154-3008, LT 151-201, and GT 3018 occurring at the 1\textsuperscript{st}, 8\textsuperscript{th}, and 15\textsuperscript{th} second, respectively, with a recording length of 50 s. The frequency recording from a randomly selected bus is used to simulate data collected from an FDR. The unmixing results using the proposed NSEU are shown in Fig. 5. From the top sub-figure, we clearly see that there are three events above the detection threshold that occurred around 1.03 s, 8.04 s, and 15.05 s, respectively. The reconstructed signal shown in the middle sub-figure resembles the original input signal with just very small deviation. In the bottom sub-figure, we sum up all the coefficients that belong to the same root-pattern for recognition purpose, in order to identify what type of event has occurred. The indices for root-pattern 1 \( \sim 4, 5 \sim 9, 10 \) represent generator trip, line trip, and load shedding, respectively, and the indices with black square is the ground truth. From this detection result we know that the first two events are line trips and the third event is a generator trip, which is exactly what happened according to the ground truth.

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Detection</th>
<th>False Alarm</th>
<th>R-P Recog</th>
<th>OT-Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1C</td>
<td>100%</td>
<td>0%</td>
<td>100%</td>
<td>0.057%</td>
</tr>
<tr>
<td>M2C</td>
<td>100%</td>
<td>5%</td>
<td>90%</td>
<td>0.1331s</td>
</tr>
<tr>
<td>M3C</td>
<td>100%</td>
<td>6.5%</td>
<td>86.7%</td>
<td>0.1056s</td>
</tr>
</tbody>
</table>
VII. EVALUATION WITH REAL EVENT DATA

For evaluation with real event cases, we take the same strategies as used for experiments on simulated cases. However, the root-patterns utilized are learned from real data collected from both US-wide FNET (generator trip and load shedding events) and PSS/E (the line trip events).

A. Event Data and Preprocessing

Three real event cases are used for evaluation.

Case 1: A single event (generator trip) happened at a power plant on May 30, 2006, as shown in Fig. 6 (Left). We use the first 50-second data (which is provided at 0.1s interval for a total of 500 samples). 10 FDRs at different locations are used for checking the detection repeatability.

Case 2: A multiple-event case (a generator trip and a line trip) happened at Surry, VA on February 2, 2011, as shown in Fig. 6 (Middle). We use the first 50-second data and 18 FDRs are used for checking the detection repeatability.

Case 3: A multiple-event case that happened on August 4, 2007, comprised of multiple single-line-to-ground faults on a 765-kV line and generator trip at two locations. The Eastern Interconnection frequency thus dropped from 60 to 59.864 Hz, as shown in Fig. 6 (Right). We use the first 60-second data because major events occurred in this period of frequency transition (from stable state 60 Hz to the next stable 59.864 Hz) for event detection purpose. 18 FDRs at different locations are used for checking the detection repeatability.

Before applying the NSEU algorithm to detect each constituent event, it is necessary to eliminate noises in the frequency data. There are, in general, two kinds of noises reside in a frequency signal, including white noise and impulsive noise. Elimination of these two kinds of noises is an intricate process [32]. We adopt the adaptive median filter designed in [7] for noise removal. Fig. 7 shows an example of the filtering effect. We observe that the filter successfully removed white noise and spikes in original frequency signal, preventing possible detection error caused by undesired noises.

B. Root- Pattern Learning

1) Training Case Selection: In order to select appropriate training cases, real data obtained from FNET and simulated
We set $K=6$ and six root-patterns are learned for each of the three event categories. As shown in Fig. 8, each root-pattern has 200 data entries standing for 20 s data with 0.1 s interval. From this result we can confirm that the learned root event patterns of each category indeed share similar characteristics with a certain degree of difference, as explained in Sec. IV-A. Other clustering algorithms, such as expectation-maximization (EM) or Dirichlet Process (DP) may be used for clustering as well, however, those algorithms produce similar results.

C. Detection Result

The learned root-patterns from K-means are used to construct temporal root-patterns to form the dictionary $S$. We set the detection threshold as 0.04. That is, a temporal root-pattern will be selected if the value of its corresponding coefficient is above the threshold.

.data from PSS/E are used for learning. As described in [20], significant time was devoted for selecting the training cases. In the United States, the power system is divided into the Eastern (EI), Western (WECC), and Texas (ERCOT) interconnections, which are not synchronized with each other.

First, data from individual generation trip and load shedding events are retrieved from the FNET database [33]. Since FNET does not currently detect line trips, there is no entries or corresponding data for this event type. PSS/E was instead used to perform time-domain simulations of line trips. A 16,000-bus model of the Eastern Interconnection was used for simulations. Approximately 75 buses corresponding to actual FDRs were selected as measurement points, and lines adjacent to these buses were tripped one at a time. A 20-second simulation was performed in each case, with measurement points being saved at 0.1-second intervals to match the reporting rate used by FDRs. Finally, all the data selected from different sources for training is shown in Table II.

1) K-means: Next, K-means clustering algorithm is used for root-pattern learning. We also used the adaptive median filter to preprocess the training data before applying K-means. We set $K=6$ and six root-patterns are learned for each of the three event categories. As shown in Fig. 8, each root-pattern has 200 data entries standing for 20 s data with 0.1 s interval. From this result we can confirm that the learned root event patterns of each category indeed share similar characteristics with a certain degree of difference, as explained in Sec. IV-A. Other clustering algorithms, such as expectation-maximization (EM) or Dirichlet Process (DP) may be used for clustering as well, however, those algorithms produce similar results.

C. Detection Result

The learned root-patterns from K-means are used to construct temporal root-patterns to form the dictionary $S$. We set the detection threshold as 0.04. That is, a temporal root-pattern will be selected if the value of its corresponding coefficient is above the threshold.

![Adaptive Median Filter](Image)

Fig. 7. Result of applying the adaptive median filter on a signal collected from the FDR 14 of Case 3. The filter successfully removed white noise and spikes, especially the large spike around the 32-th second.

<table>
<thead>
<tr>
<th>Table II</th>
<th>Breakdown of training case event types</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EI</td>
</tr>
<tr>
<td>Generation Trip</td>
<td>543</td>
</tr>
<tr>
<td>Load Shedding</td>
<td>160</td>
</tr>
<tr>
<td>Line Trip</td>
<td>257</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table III</th>
<th>Event detection results for Case 1. One generator trip actually happened in this real event. All the FDRs successfully detected the generator trip.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDR</td>
<td>1</td>
</tr>
<tr>
<td>GenTrip1</td>
<td>8.4s</td>
</tr>
<tr>
<td>FDR</td>
<td>6</td>
</tr>
<tr>
<td>GenTrip1</td>
<td>9.0s</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table IV</th>
<th>Event detection results for Case 2. One generator trip and one line trip actually happened in this real event. All the FDRs successfully detected one generator trip (root-pattern 6) and one line trip (root-pattern 8 or 12).</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDR</td>
<td>1</td>
</tr>
<tr>
<td>GenTrip6</td>
<td>12.4s</td>
</tr>
<tr>
<td>LineTrip8</td>
<td>14.2s</td>
</tr>
<tr>
<td>FDR</td>
<td>13</td>
</tr>
<tr>
<td>GenTrip6</td>
<td>12.0s</td>
</tr>
<tr>
<td>LineTrip8</td>
<td>15.2s</td>
</tr>
</tbody>
</table>

1) Case 1: Recall that case 1 is a single event case. The detection results, as shown in Table III, indicate that the NSEU algorithm successfully detected a single generator trip of root-pattern 1 from each signal of 10 FDRs at different locations without any false alarm.

2) Case 2: It is a multiple-event case with one generator trip followed by one line trip. The detection results are shown in Table IV. All of the signals from 18 FDRs successfully detected one generator trip from root-pattern 6 and one line trip, although the detected line trip may be from either root-pattern 8 or 12. This maybe because the line trip’s root-patterns learned based on PSS/E simulations can not reflect very well the behavior of real line trip events.

3) Case 3: Case 3 is another multiple-event case with two generator trips and possibly two or three line trips involved, thus it is more complicated than cases 1 and 2. The detection results are shown in Table V. The results show that the NSEU algorithm successfully detected two generator trips from 16 out of 18 FDRs and two line trips from 17 FDRs without any false alarms. An example of the detected root-patterns and the reconstructed signal from FDR 14 are shown in Fig. 9. From the top sub-figure of sparse coefficients, it is clear that the algorithm detects two generator trips and two line trips correctly. The reconstructed signal in the middle sub-figure is also very close to the original signal.

This result further confirms that the NSEU algorithm can not only correctly detect single event, but it also can handle multiple cascading events with accurate temporal localization. However, line trips in this case are not completely detected since one line trip is missed. As for the processing time, since the sampling cycle of the FDR sensors in FNET is 0.1 second, a multiple-event case from one stable state to another stable state lasted for 60 s will have 600 samples. Then, in our settings, the dictionary derived from 18 root-
### D. Summary of Unmixing Performance

In the three experiments with real event cases, we have analyzed frequency signals from 46 FDRs in total. As shown in Table VI, the proposed NSEU algorithm can detect constituent events with 98.15% averaged accuracy for generator trips and 82.4% averaged accuracy for line trips without any false alarms (FA). Due to the lack of ground truth, we cannot evaluate performance of recognition or temporal localization.

From the experimental results, it is easy to see the advantage of the proposed NSEU algorithm over existing event detection techniques. In Fig. 6, we observe that there is no immediately perceivable difference between frequency signal of a multiple-event case and that of a single-event case, because the nature of the mixing process will occlude or degrade most of the features from different root patterns. Most of the existing techniques based on immediately detectable information can only detect the starting time of the initial event involved in multiple cascading events. In contrast, the NSEU algorithm is able to uncover the constituent root events with high detection accuracy, except for the line trip detection rate in Case 3. The difficulty for line trip detection can be analyzed from three aspects: First, the root-patterns of line trip are learned based on simulation, which might not reflect the dynamics of line trips that occurred in real world. Second, the frequency change of line trips is generally smaller than that of generator trips, probably causing the unmixed coefficients on line trips to be smaller than that on generator trips. Therefore, line trips might not be detected if using the same detection threshold for GTs and LSs. Third, the power imbalance caused by some line trips is quite small that can be easily adjusted by system’s self-resilience.

### Table V

Table V: Event detection results for Case 3. Two generator trips and multiple line trips may have happened during this real event. Most of the FDRs detected two generator trips (root-patterns 3&6) and two line trips (root-patterns 8&12). The generator trip root-pattern 3 was not detected by FDR 2&16 and the line trip root-pattern 12 was not detected by FDR 3.

<table>
<thead>
<tr>
<th>FDR</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>GenTrip1</td>
<td>5.4s</td>
<td>4.3s</td>
<td>3.6s</td>
<td>2.6s</td>
<td>3.2s</td>
<td>4.2s</td>
</tr>
<tr>
<td>GenTrip6</td>
<td>6.2s</td>
<td>4.2s</td>
<td>3.4s</td>
<td>3.2s</td>
<td>4.2s</td>
<td>4.2s</td>
</tr>
<tr>
<td>LineTrip8</td>
<td>9.2s</td>
<td>7.2s</td>
<td>7.4s</td>
<td>7.2s</td>
<td>9.0s</td>
<td>6.2s</td>
</tr>
<tr>
<td>LineTrip12</td>
<td>7.4s</td>
<td>6.2s</td>
<td>5.6s</td>
<td>6.4s</td>
<td>6.4s</td>
<td>5.4s</td>
</tr>
<tr>
<td>FDR</td>
<td>7</td>
<td>5.5s</td>
<td>4.4s</td>
<td>4.4s</td>
<td>4.4s</td>
<td>4.4s</td>
</tr>
<tr>
<td>GenTrip3</td>
<td>4.4s</td>
<td>3.3s</td>
<td>6.0s</td>
<td>6.0s</td>
<td>6.0s</td>
<td>5.3s</td>
</tr>
<tr>
<td>GenTrip6</td>
<td>6.4s</td>
<td>3.3s</td>
<td>6.6s</td>
<td>6.6s</td>
<td>6.6s</td>
<td>5.4s</td>
</tr>
<tr>
<td>LineTrip8</td>
<td>9.6s</td>
<td>5.6s</td>
<td>4.6s</td>
<td>4.6s</td>
<td>4.6s</td>
<td>4.6s</td>
</tr>
<tr>
<td>LineTrip12</td>
<td>7.2s</td>
<td>5.8s</td>
<td>7.2s</td>
<td>6.0s</td>
<td>6.6s</td>
<td>7.2s</td>
</tr>
<tr>
<td>FDR</td>
<td>15</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>GenTrip3</td>
<td>3.5s</td>
<td>3.5s</td>
<td>3.5s</td>
<td>3.5s</td>
<td>3.5s</td>
<td>3.5s</td>
</tr>
<tr>
<td>GenTrip6</td>
<td>4.2s</td>
<td>3.0s</td>
<td>3.2s</td>
<td>4.2s</td>
<td>4.2s</td>
<td>4.2s</td>
</tr>
<tr>
<td>LineTrip8</td>
<td>9.0s</td>
<td>7.4s</td>
<td>6.8s</td>
<td>7.2s</td>
<td>7.2s</td>
<td>6.2s</td>
</tr>
<tr>
<td>LineTrip12</td>
<td>7.0s</td>
<td>8.0s</td>
<td>7.2s</td>
<td>7.4s</td>
<td>5.8s</td>
<td>6.4s</td>
</tr>
</tbody>
</table>

---

Table VI: Quantitative evaluation on real event cases

<table>
<thead>
<tr>
<th>FDR Num</th>
<th>GT Det</th>
<th>GT FA</th>
<th>LT Det</th>
<th>LT FA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case1</td>
<td>10</td>
<td>100%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Case2</td>
<td>18</td>
<td>100%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Case3</td>
<td>18</td>
<td>94.4%</td>
<td>0%</td>
<td>64.8%</td>
</tr>
<tr>
<td>Mean</td>
<td>98.15%</td>
<td>0%</td>
<td>82.4%</td>
<td>0%</td>
</tr>
</tbody>
</table>

---

The three experiments with real event cases have analyzed frequency signals from 46 FDRs in total. As shown in Table VI, the proposed NSEU algorithm can detect constituent events with 98.15% averaged accuracy for generator trips and 82.4% averaged accuracy for line trips without any false alarms (FA). Due to the lack of ground truth, we cannot evaluate performance of recognition or temporal localization.

From the experimental results, it is easy to see the advantage of the proposed NSEU algorithm over existing event detection techniques. In Fig. 6, we observe that there is no immediately perceivable difference between frequency signal of a multiple-event case and that of a single-event case, because the nature of the mixing process will occlude or degrade most of the features from different root patterns. Most of the existing techniques based on immediately detectable information can only detect the starting time of the initial event involved in multiple cascading events. In contrast, the NSEU algorithm is able to uncover the constituent root events with high detection accuracy, except for the line trip detection rate in Case 3. The difficulty for line trip detection can be analyzed from three aspects: First, the root-patterns of line trip are learned based on simulation, which might not reflect the dynamics of line trips that occurred in real world. Second, the frequency change of line trips is generally smaller than that of generator trips, probably causing the unmixed coefficients on line trips to be smaller than that on generator trips. Therefore, line trips might not be detected if using the same detection threshold for GTs and LSs. Third, the power imbalance caused by some line trips is quite small that can be easily adjusted by system’s self-resilience.
VIII. CONCLUSION AND FUTURE WORK

This paper presented a novel interpretation of the systematics of frequency signal formation of multiple events in a power grid. Through analysis of the connection between frequency disturbance of multiple events and frequency fluctuation of single event, we developed an effective and promising constrained “event unmixing” algorithm, NSEU, based on a simple linear mixture model for multiple cascading events detection, recognition, and temporal localization. The experimental results with both simulated and real event cases demonstrated the effectiveness of the proposed algorithm. Benefited by the frequency data recorded by FNET, this work provides a new and feasible way to obtain high-resolution situational awareness of the power grid system. The findings from this work would also benefit other applications, particularly for example, microgrids remote monitoring, multiple events localization, and smart grid coordination.

In order to further improve the performance of the NSEU algorithm and extend its potential for situational awareness, we plan to conduct the following investigations in the future. First, to make the algorithm more robust to oscillations, we will consider adding the oscillation root-events in the overcomplete source signature matrix, S. Second, to improve performance of line trip detection, we will design a “weighted” dictionary such that the unmixed coefficients of different root events would be at comparable scale. Third, as observed from Tables III and IV, different FDRs would report different starting times for the same event. This can be further utilized for event localization purpose as described in [9], [11].

ACKNOWLEDGMENT

This work was supported in part by the Engineering Research Center Program of the National Science Foundation and the Department of Energy under NSF Award Number EEC-1041877 and the CURENT Industry Partnership Program. The work was also supported in part by NSF-CNS-1239478.

REFERENCES

Wei Wang (S’12) is currently a Ph.D. student in the Department of Electrical Engineering and Computer Science at the University of Tennessee, Knoxville. He received his B.S. and M.S. degrees from University of Science and Technology Beijing in 2006 and Beihang University in 2009, respectively. His research interests include signal/image analysis, pattern recognition, machine learning and smart grid secure monitoring.

Li He (S’11) received his Ph.D. from the University of Tennessee, Knoxville in 2013, M.S. degree from Beijing Jiaotong University in 2008, and B.S. degree from China University of Geoscience, Beijing in 2005. He is currently an Algorithm Engineer at KLA-Tencor Corporation. His research interests include machine learning, image processing, linear unmixing and color emotion.

Penn Markham (S’03) received his B.S.E.E from Virginia Polytechnic Institute and State University in 2006 and his Ph.D. from the University of Tennessee, Knoxville, in 2012. He is now a research engineer in the Grid Operations and Planning group at the Electric Power Research Institute (EPRI) in Knoxville, Tennessee.

Hairong Qi (S’97-M’00-SM’05) received the B.S. and M.S. degrees in computer science from Northern JiaoTong University, Beijing, China, in 1992 and 1995, respectively, and the Ph.D. degree in computer engineering from North Carolina State University, Raleigh, in 1999. She is now a Professor in the Department of Electrical Engineering and Computer Science at the University of Tennessee, Knoxville. Her current research interests are in advanced imaging and robust collaborative processing in resource-constraint distributed environment and information unmixing. She is the recipient of the NSF CAREER Award. She also received the Best Paper Awards at the 10th International Conference on Pattern Recognition and the third ACM/IEEE International Conference on Distributed Smart Cameras. She recently receives the Highest Impact Paper from the IEEE Geoscience and Remote Sensing Society.

Yifu Liu (S’88-M’89-SM’99-F’04) is currently a Governor Chair Professor at the University of Tennessee, Knoxville and Oak Ridge National Laboratory. Prior to joining UTK/ORNL, she was a professor at Virginia Tech. Dr. Liu received her M.S. and Ph.D. degrees from the Ohio State University, Columbus, in 1986 and 1989. She received the B.S. degree from XiAn Jiaotong University. She led the effort to create the North American power grid monitoring network (FNET) at Virginia Tech which is now operated at UTK and ORNL as FNET/GridEye. Her current research interests include power system wide-area monitoring and control, large interconnection level dynamic simulations, electromagnetic transient analysis, and power transformer modeling and diagnosis.

Qing Charles Cao (M’10) is currently an assistant professor in the Department of Electrical Engineering and Computer Science at the University of Tennessee. He received his Ph.D. degree from the University of Illinois in October, 2008, and his Master’s degree from the University of Virginia. His research interests include wireless sensor networks, embedded systems, and distributed networks. He is a member of IEEE, ACM, and the IEEE Computer Society. He has published more than 50 papers in various journals and conferences, and has served as technical committee members for more than 30 conferences held by IEEE and ACM.

Leon M. Tolbert (S’88-M’91-SM’98-F’13) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1989, 1991, and 1999, respectively. He worked at Oak Ridge National Laboratory, Oak Ridge, TN, from 1991 until 1999. He was appointed as an Assistant Professor with the Department of Electrical and Computer Engineering, The University of Tennessee, Knoxville, in 1999. He is currently the Min Kao Professor in the Department of Electrical Engineering and Computer Science, The University of Tennessee. He is the UTK Campus Director for the National Science Foundation/Department of Energy Research Center, CURENT (Center for Ultra-wide-area Resilient Electric Energy Transmission Networks). He is also a Senior Research Engineer with the Power Electronics and Electric Machinery Research Center, Oak Ridge National Laboratory. He received an NSF CAREER Award in 2001, the 2001 IEEE Industry Applications Society Outstanding Young Member, and four prize paper awards from the IEEE Industry Applications Society and the IEEE Power Electronics Society.