Irregularity Detection in Output Power of Distributed Energy Resources Using PMU Data Analytics in Smart Grids

Younes Seyedi, Houshang Karimi, *Senior Member, IEEE*, and Santiago Grijalva, *Senior Member, IEEE*

***Abstract*—The output power of distributed energy resources (DERs) may experience irregular fluctuations due to variations of renewable sources, which need to be monitored in order to reliably control the grid. This paper proposes a novel approach for centralized detection of such irregularities based on the time-series analysis of the data reported by phasor measurement units (PMUs). In this approach, a network controller constructs datasets of time-aligned real/reactive powers for different zones. The datasets are transformed into sequences of short-time local outlier probability (ST-LOP) that are analyzed to identify the DER events. The network controller estimates features such as the average duration and the similarity degree that is a measure of spatio-temporal correlation between the DER events. As a use case, event-triggered control of solar photovoltaic (PV) systems with energy storage devices is investigated. The simulation results for the IEEE 123-bus network corroborate the effectiveness of the developed analytics for detection and mitigation of ramp-rate solar power fluctuations. Smart microgrids and active distribution networks can employ the developed analytics to improve a range of diagnostic and control functionalities.**

***Index Terms*—Energy storage, microgrids, photovoltaic (PV) systems, power control, renewable energy sources, smart grids.**

I. Introduction

THE utilization of distributed energy resources (DERs) in distribution grids and microgrids has been steadily increasing [1]. DERs may include demand-response, energy storage, and renewable sources such as solar photovoltaic (PV) and wind energy [2]. As the penetration of DERs increases, the various electric quantities in the grid experience new behavior at new temporal and spatial scales. In particular, changes in energy production must be followed more carefully by grid control and coordination modules, in order to maintain a reliable system [3]. Advanced sensors such as smart meters [4] and phasor measurement units (PMUs) [5] significantly contribute to the development of modern grids. These specialized sensors rely on communication's infrastructure to deliver accurate and high-resolution data to control and monitoring systems [6], [7].

It is known that the output power of DERs may experience rapid and irregular fluctuations due to the variability of renewable sources. Hereinafter, these irregularities in the power generation are called DER events. The DER events are of particular importance since, if not properly handled, they may jeopardize the power quality and stability [8]. Moreover, they may lead to significant voltage distortions [9], which in turn result in disruptions and potentially outages. Hierarchical control applications in smart microgrids may also require knowledge about such DER events. Power fluctuations of DER systems in a microgrid that operates in the islanded mode may result in loss of voltage/frequency stability. It is also known that DER events have impacts on coordination of protective devices [10] and can affect the performance of protection schemes [11]. In almost all cases, the aforementioned DER events should be detected with low latency.

DER event detection in smart grids has been investigated by several researchers [12]–[17]. A data-driven method for monitoring of PV systems is proposed in [12], which verifies connection to grid of PV systems based on a change-point detection approach. Ge *et al.* [13] address online identification of real power events based on principal component analysis (PCA) of synchrophasor data. Rafferty *et al.* [14] propose a moving window PCA method under time-varying behavior of the power system. Detection of real/reactive power events based on the wavelet analysis of voltage and frequency data is studied in [15]. Negi *et al.* [16] propose an event detection method based on the total energy content of the wavelet coefficients of the PMU data. The wavelet-based methods can identify events such as generator trips that result in abrupt voltage/frequency transients. In [17], the authors study time-series data analysis methods for model-less event detection in distribution networks.

The previously discussed methods have the following limitations. First, they aim at identifying critical events that can create abrupt disturbances in voltage or frequency. They do not consider moderate events such as power irregularities. Second, the existing methods do not estimate the duration of the DER events. Third, they do not provide a measure of the spatio-temporal correlation between concurrent events that emanates from renewable resource variability.

Manuscript received July 15, 2018; revised August 10, 2018; accepted August 12, 2018. Date of publication August 17, 2018; date of current version April 3, 2019. Paper no. TII-18-1826. *(Corresponding author: Houshang Karimi.)*

Y. Seyedi and H. Karimi are with the Department of Electrical Engineering, Polytechnique Montreal, Montreal, QC H3T 1J4, Canada (e-mail: younes.seyedi@polymtl.ca; houshang.karimi@polymtl.ca).

S. Grijalva is with the Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: sgrijalva@ece.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2018.2865765

This paper presents a novel approach for centralized detection of DER events that overcomes the mentioned limitations. The network is divided into different detection zones, where each detection zone is associated with a DER system. The network controller employs a set of PMUs to construct local datasets of real/reactive powers for a relatively short interval. The data samples that are most likely affected by the events exhibit a higher degree of outlierness with respect to their local datasets. Based on this premise, the network controller identifies the DER events by analyzing the sequences of short-time local outlier probability (ST-LOP).

The main contributions of this paper are as follows.

1) By adopting the proposed approach, the DER events can be detected irrespective of the detailed network model, and the prior knowledge about the loads is not required.
2) Upon event detection, the signature sequences are analyzed to estimate the duration of events.
3) A meaningful parameter, namely, the interzone similarity degree, is introduced with the aim of measuring the spatio-temporal correlations between concurrent DER events.
4) The proposed detection approach is combined with an event-triggered control scheme to smooth out the PV power events due to ramp-rate solar irradiance variations.

II. Data Analytics for Event Detection

Basically, data-driven methods rely on the analysis of a set of measurements provided by sensors. Therefore, it is important to have a proper mechanism for PMU data collection.

A. Data Collection

In general, the network can have several DER systems, and it is necessary to avoid ambiguity in the event localization with a limited number of PMUs. To this aim, the network is divided into different detection zones such that each zone includes only one DER. However, several dispersed loads may exist in the zones and the PMUs are not necessarily installed at the point of common coupling of the DERs. Suppose that the network is divided into N_z detection zones. The detection zones are denoted by z_n, where the subscript n shows the index of the zone

$$n \in \{1, 2, \ldots, N_z\}.$$

As illustrated in Fig. 1, $P_n(t)$ and $Q_n(t)$ denote the total instantaneous real and reactive powers that enter the nth zone at the time instant t, respectively. $P_n(t)$ and $Q_n(t)$ are basically the algebraic sum of real and reactive powers generated by the DER system and consumed by the dispersed loads in the zone z_n. These quantities are measured based on the synchrophasors of voltage and current signals, as follows:

$$P_n(t) = V_n(t) I_n(t) \cos\left(\phi_V(t) - \phi_I(t)\right) \quad (1)$$

$$Q_n(t) = V_n(t) I_n(t) \sin\left(\phi_V(t) - \phi_I(t)\right) \quad (2)$$

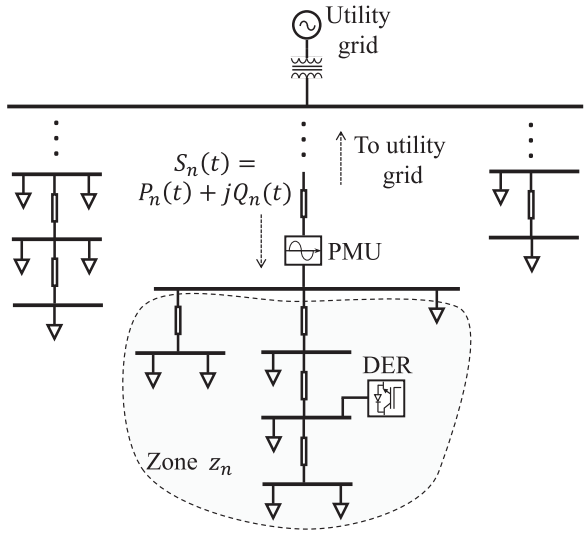


Fig. 1. Generic illustration of the detection zones.

where $V_n(t)$ and $I_n(t)$ are the instantaneous phasor magnitude of the voltage and current signals, respectively. $\phi_V(t)$ and $\phi_I(t)$ are the instantaneous phase-angles (measured at the fundamental frequency) of the voltage and current signals, respectively.

The PMUs usually transmit the synchrophasors of voltage/current signals at a high rate (e.g., 30 samples/s) in order to support multiple applications at the same time [7], [18]. It is therefore plausible to utilize the high-rate PMU data for event detection applications. Applications that are pertinent to faults, e.g., networked protection schemes, are highly time-critical in the sense that they require very fast data sampling [19]. Applications such as centralized DER monitoring are deemed less stringent in terms of the sampling rate [3], [20]. Moreover, the fluctuations in the output power of the DER systems can be tolerated over relatively longer time. This observation implies that the DER events can be detected at lower data rates without compromising the network performance. To this aim, the phasor data concentrator (PDC) uses a decimator to downsample the received synchrophasors to a lower rate and then reports the real/reactive power measurements.

The overall procedure for event detection and centralized control of DERs is depicted in Fig. 2. The network controller receives the decimated data from the PDC, and a triggering signal is issued upon detection of a DER event. The issued signal triggers a centralized control application to improve the performance of the entire network. The control application makes decision about the proper commands that are sent to the DERs via communication links. The control commands can be switching actions, change of DER parameters, real/reactive power setpoints, etc. The communications aspects for centralized control applications in smart grids are well addressed in the literature, e.g., [21].

In real-life networks, the PMUs may send low-quality or bad synchrophasor data due to timing errors or loss of synchronization. Recently, elaborate methods have been developed for

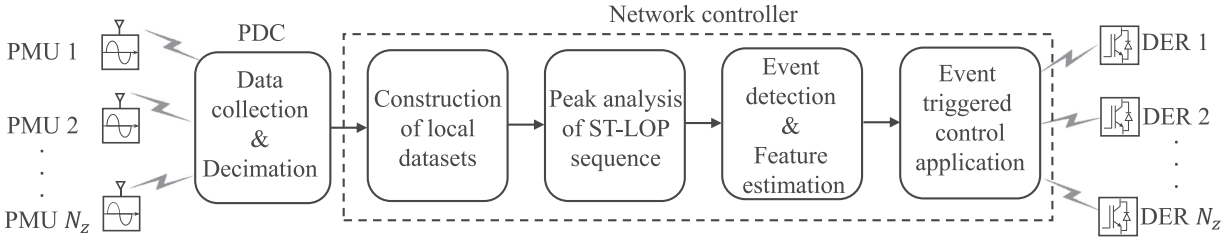


Fig. 2. Overview of the proposed approach for event detection and centralized control of DER systems.

smart grids, which are capable of screening PMU data in real-time with negligible processing delay [22]. Such methods can be employed to build a preprocessing subsystem that can be augmented to the overall system shown in Fig. 2. This new subsystem can be placed before the PDC to detect and compensate for bad PMU data. This increases the robustness of the network controller and improves the reliability of event detection in the presence of bad PMU data.

B. Short-Time Local Outlier Probabilities

Suppose that T_r denotes the reporting time (after the down-sampling) of the power measurements. The power data reported from the nth zone can be represented as discrete-time vectors

$$\mathbf{v}_n(k) = \begin{bmatrix} P_n(kT_r) \\ Q_n(kT_r) \end{bmatrix} + \begin{bmatrix} \varepsilon_{n,P}(kT_r) \\ \varepsilon_{n,Q}(kT_r) \end{bmatrix} \tag{3}$$

where k shows the discrete-time index

$$k \in \{0, 1, 2, 3, ...\}.$$

In (3), $P_n(kT_r)$ and $Q_n(kT_r)$ are the samples of $P_n(t)$ and $Q_n(t)$, which are obtained at the time instant $t = kT_r$, respectively. $\varepsilon_{n,P}(.)$ and $\varepsilon_{n,Q}(.)$ denote the additive noise in the real and reactive power data, respectively.

Equations (1) and (2) reveal that data noises $\varepsilon_{n,P}(.)$ and $\varepsilon_{n,Q}(.)$ are essentially correlated with each other. Therefore, $\varepsilon_{n,P}(.)$ and $\varepsilon_{n,Q}(.)$ can be modeled as jointly Gaussian random variables with zero mean and covariance matrix $\mathbf{\Sigma}_n$

$$\begin{bmatrix} \varepsilon_{n,P}(.) \\ \varepsilon_{n,Q}(.) \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_n).$$

The covariance matrix of the data noise depends on the variance of noise in measuring the synchrophasors of voltage and current signals.

The local datasets are analyzed by the network controller every τ_D second, where τ_D is the length of the detection interval. Without loss of generality, it is assumed that the detection interval is an integer multiple of the data reporting time, i.e.,

$$\tau_D = N_D T_r, \qquad N_D \gg 1. \tag{4}$$

Hence, the ith local dataset for the nth zone contains N_D data samples and is constructed as

$$\mathcal{D}_n(i) = \Big\{ \mathbf{v}_n\big(iN_D - N_D\big), \mathbf{v}_n\big(iN_D - N_D + 1\big), \ldots, \mathbf{v}_n \times \big(iN_D - 1\big) \Big\}. \tag{5}$$

The length of the detection interval affects the amount of information that the network controller can extract from the local datasets. A short detection interval can degrade the statistical inference in each time window. Basically, the value of τ_D should be greater than the average duration of the DER events. The reporting time controls the sampling resolution of the events and should be determined based on the ramp-rate of the power fluctuations. A typical value of T_r can be equal to the time interval during which the output power varies by at least 10% of the DERs' capacity. The ramp-rate of the power fluctuations is governed by the variation rate of the energy resources and the inertia of the DER systems [23].

If a DER event occurs within the detection interval, then some data samples exhibit a higher degree of outlierness with respect to the local datasets. Therefore, once the local datasets are determined, the network controller evaluates the degree of outlierness for every data sample in the local datasets. Algorithm 1 explains the procedure of finding the sequences of ST-LOP, where $(.)^{\mathrm{T}}$ denotes the transpose operator.

It should be noted that the network controller is capable of receiving information from other monitoring systems based on the IEC 61850 protocols [24]. For instance, supervisory control and data acquisition, protective relays, and intelligent electronic devices can share information with the network controller. These systems send alarm messages when an abnormal condition is observed in the network, e.g., fault, outage, overloading, etc. By processing the messages sent from such systems, the network controller can discriminate between outliers due to DER events and outliers that stem from malfunctioning of other elements.

In step 1, the network controller constructs the context sets for the data samples. In step 2, the standard distance is found as the square root of the mean squared distance. The standard distance of a data sample shows the density around that point based on its context set. According to (9) in step 3, the estimated densities are centered and normalized by the average of the densities around all data samples that belong to the context set of a data sample. Step 4 calculates the standard probabilistic outlier factor by normalizing the short-term probabilistic local outlier factors.

Algorithm 1: Outlier Analysis of Local Datasets.

For $1 \le n \le N_z$, $i \ge 1$, and $iN_D - N_D \le k \le iN_D - 1$:
1) Construct the context set for every data sample in the local dataset $\mathcal{D}_n(i)$:

$$\mathcal{C}_n(i,k'') = \Big\{ \mathbf{v}_n(k') \,|\, \mathbf{v}_n(k') \in \mathcal{D}_n(i), k' \neq k'' \quad iN_D - N_D \le k', k'' \le iN_D - 1 \Big\} \tag{6}$$

2) Calculate the standard distance for every data sample in the local dataset $\mathcal{D}_n(i)$:

$$\bar{d}_n(i,k'') = \sqrt{\frac{1}{N_D - 1} \sum_{k'} \Big(d_n^2(k',k'') \,|\, \mathbf{v}_n(k') \in \mathcal{C}_n(i,k'') \Big)} \tag{7}$$

where the distance between two data samples that belong to the same local dataset is given by

$$d_n(k',k'') \triangleq \sqrt{\Big(\mathbf{v}_n(k') - \mathbf{v}_n(k'')\Big)^{\mathrm{T}} \Big(\mathbf{v}_n(k') - \mathbf{v}_n(k'')\Big)} \tag{8}$$

3) Find the short-term probabilistic local outlier factor for every data sample in $\mathcal{D}_n(i)$:

$$\gamma_n(i,k'') = \frac{(N_D - 1)\bar{d}_n(i,k'')}{\sum_{k'} \Big(\bar{d}_n(i,k') \,|\, \mathbf{v}_n(k') \in \mathcal{C}_n(i,k'') \Big)} - 1 \tag{9}$$

4) Calculate the standard probabilistic local outlier factor:

$$\alpha_n(i) = \sqrt{\frac{1}{N_D} \sum_{k''=iN_D - N_D}^{iN_D - 1} \Big(\gamma_n^2(k'') \,|\, \mathbf{v}_n(k'') \in \mathcal{D}_n(i) \Big)} \tag{10}$$

5) Find the ST-LOPs for the data samples:

$$\beta_n(i,k) = \max\Big\{0, \mathrm{erf}\Big(\frac{\gamma_n(i,k)}{\sqrt{2}\lambda\alpha_n(i)}\Big)\Big\} \tag{11}$$

In the last step, (11) yields the ST-LOPs, where erf(.) denotes the error function. The error function can be numerically calculated based on the following expression:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-y^2\right) dy. \tag{12}$$

The parameter λ in (11) represents the significance level in the estimation of the degree of outlierness [25]. It should be noted that the value of λ does not affect ranking of DER events for a given local dataset. Equation (11) maps the estimated probabilistic local outlier factor into a number that lies in the interval $[0, 1]$. Therefore, the degree of outlierness can also be interpreted as the likelihood of being an outlier. If the value of ST-LOP of a data sample is close to 0, then that data sample belongs to a dense local dataset. An advantage of Algorithm 1 is that it decreases the impact of the data noise on the detection mechanism.

Algorithm 2: Peak Analysis of the ST-LOP Sequence.

Given $1 \le n \le N_z$, and $i \ge 1$: Initialize with $\mathcal{M}'_n(i) = \mathcal{M}_n(i) = \emptyset$
1) Update the index set $\mathcal{M}'_n(i)$ as

$$\mathcal{M}'_n(i) = \big\{ k \,|\, \beta_n(i,k) \ge \beta_{\mathrm{th}}, iN_D - N_D \le k \le iN_D - 1 \big\} \tag{13}$$

2) If $|\mathcal{M}'_n(i)| = 0$, then terminate the algorithm. Otherwise, find the index in $\mathcal{M}'_n(i)$ that yields the largest ST-LOP:

$$k_{\max} = \arg \max_{k \in \mathcal{M}'_n(i)} \beta_n(i,k) \tag{14}$$

3) Append $k_{\max}$ to the index set $\mathcal{M}_n(i)$.
4) Update $\mathcal{M}'_n(i)$ by removing all indices that are within τ_s s of the time instant $k_{\max}T_r$, i.e., eliminate all indices that satisfy the following inequality:

$$|k - k_{\max}|T_r \le \tau_s, \qquad k \in \mathcal{M}'_n(i) \tag{15}$$

5) Go to step 2 and repeat the procedure for the largest remaining ST-LOP and iterate until $\mathcal{M}'_n(i)$ runs out of indices.

C. Event Detection: Analysis of the ST-LOP Sequence

Algorithm 1 maps the reported data vectors to a time-series of scalar ST-LOPs. In realistic scenarios, however, the ST-LOP sequences become random processes. The randomness of ST-LOPs emanates from variations of the energy resource, load changes, and the measurement noise. Hence, the ST-LOP sequences should be further analyzed to distinguish between irregular power fluctuations and other random phenomena. Algorithm 2 explains event detection based on the peak analysis of the ST-LOP sequences.

The presence of an irregularity in the output power of a DER system adds a higher degree of outlierness to the data samples. Based on this premise, it is concluded that larger peaks that appear in the ST-LOP sequence are most likely produced by the DER events. Moreover, the powers consumed by the loads follow a regular trend or vary in a predictable manner over a short time window [26], [27]. Therefore, lower scores are assigned to the load changes within the detection interval, and the peak analysis of the ST-LOP sequence distinguishes the DER events. Robustness against the measurement noise is another advantage of this method such that the events can be reliably detected when the synchrophasor data are polluted by noise.

The core of Algorithm 2 is an iterative method that ensures that legitimate events stand out due to their higher significance relative to other points in the sequence of ST-LOP. In step 1, the instantaneous values of the ST-LOP are compared with the detection threshold, denoted by β_{th}. If the PMUs are installed in a noisy place, then the data noise can rapidly change within the detection interval. As a result, some of the local maxima in the

Algorithm 3: Event Signature.

Given $1 \leq n \leq N_z$, and $i \geq 1$: Initialize the signature function as $\mathcal{S}_n(i,k) = 0$, where the time index k maintains $iN_D - N_D \leq k \leq iN_D - 1$. Moreover, let $\mathcal{H} = \mathcal{M}_n(i)$.
1) If $|\mathcal{H}| > 0$, then go to step 2. Otherwise, terminate the algorithm.
2) Choose $k_{\max} \in \mathcal{H}$.
3) Find the smallest integer number u such that

$$\beta_n(i, k_{\max} + m - 1) > \beta_n(i, k_{\max} + m) \tag{16}$$

holds for all $m \in \{1, 2, \ldots, u\}$.
4) Find the smallest integer number l such that

$$\beta_n(i, k_{\max} - m + 1) > \beta_n(i, k_{\max} - m) \tag{17}$$

holds for all $m \in \{1, 2, \ldots, l\}$.
5) Update the signature function as:

$$\mathcal{S}_n(i,k) = \max\Big\{\beta_n(i, k_{\max}), S_n(i,k)\Big\},$$
$$k_{\max} - l < k < k_{\max} + u \tag{18}$$

6) Remove $k_{\max}$ from $\mathcal{H}$ and go to step 1.

ST-LOP sequence may be very close to each other. In step 4, the minimum separation time, denoted by τ_s, is used to filter out spurious peaks, which are likely to be result of noise. If the minimum separation time is very small, then false events may be observed in the vicinity of a legitimate event that has the highest degree of outlierness. The output of Algorithm 2 is the index set $\mathcal{M}_n(i)$. This set consists of the time indices that are most likely associated with legitimate DER events. Therefore, its cardinality, $|\mathcal{M}_n(i)|$, gives the estimated number of DER events within the ith detection interval for the nth zone.

D. Estimation of Event Features

The network controller can obtain more information about spatial and temporal features of the DER events by processing of the ST-LOP sequences. From a network perspective, it is also important to obtain knowledge about any correlations between the DER events. To this end, a signature sequence, denoted by $\mathcal{S}_n$, is derived for the detection intervals. Algorithm 3 describes the procedure of obtaining the signature sequence by means of the index set $\mathcal{M}_n$. If no event is detected during the detection interval, then the signature sequence is an all-zero sequence within that interval. The signature sequence becomes a piecewise constant sequence when at least one event is detected, i.e., when $|\mathcal{M}_n| > 0$.

Suppose that at least one DER event is observed within the ith detection interval. In Algorithm 3, $k_{\max}$ shows a time index, which is selected from the set $\mathcal{M}_n(i)$. Basically, the ST-LOP sequence has a local maxima at $k = k_{\max}$.

In step 3, the network controller finds the time index at which the DER event ends. The end of an event is equivalent to consecutive decrements of the ST-LOPs after $k_{\max}$. In step 4, the network controller finds the time index at which the DER event begins. The beginning of an event is equivalent to consecutive increments of the ST-LOPs before $k_{\max}$. Accordingly, the estimated time interval, which is probabilistically associated with the event, is $[(k_{\max} - l)t, (k_{\max} + u)t]$, where u and l are found in steps 3 and 4, respectively. The signature sequence is updated in step 5. The value of the signature sequence is equal to the ST-LOP at the time index $k = k_{\max}$. If the intervals of two events overlap, then the largest ST-LOP is used to update the signature sequence at the common indices. This procedure is repeated until all of the time indices in $\mathcal{M}_n(i)$ are processed.

In some cases, there may be multiple events within a detection interval. Under such circumstances, the average duration of events, denoted by $\tau_n(i)$, is estimated by the normalized sum of durations of the subintervals in which the event signature is nonzero, i.e.,

$$\tau_n(i) = \frac{T_r}{|\mathcal{M}_n(i)|} \sum_{k=iN_D-1}^{iN_D-N_D} \operatorname{sgn}\Big(\mathcal{S}_n(i,k)\Big) \tag{19}$$

where sgn(.) stands for the sign function. The value of τ_n helps the network controller with estimating the amount of energy fluctuations. If the DER systems are driven by renewable resources, then monitoring and analysis of τ_n can facilitate design of energy storage devices for counteracting resource fluctuations.

Utilization of renewable energy resources results in a random degree of spatio-temporal correlation between DER systems [28]. Under such circumstances, DER events with similar features can be observed in two or more zones. The interzone similarity degree is a qualified metric that measures the similarity between the DER events that occur in different zones. The similarity degree for the zones z_n and z_m can be described by the following parameter:

$$\Gamma_{n,m}(i) = \frac{2\sum_{k=iN_D-1}^{iN_D-N_D} \mathcal{S}_n(i,k)\mathcal{S}_m(i,k)}{\sum_{k=iN_D-1}^{iN_D-N_D} \mathcal{S}_n^2(i,k) + \mathcal{S}_m^2(i,k)}. \tag{20}$$

The similarity degree lies in the interval $[0, 1]$. A small similarity degree implies low spatio-temporal correlation between DER events.

According to (20), the similarity degree becomes zero in either of the following cases. First, at least one zone does not indicate any event during the detection interval, i.e., when the event signature of the zone is all-zero sequence. Second, the zones indicate events that are temporally independent, i.e., when the event signatures do not overlap. The similarity degree becomes unity when the two zones indicate DER events with identical signature sequences. An advantage of the signature sequences is that they filter out fluctuations due to load changes and measurement noise. Hence, $\Gamma_{n,m}$ can be interpreted as a meaningful measure of the spatio-temporal correlation between DER events that occur in the zones z_n and z_m.

III. Use Case: Event-Triggered Control of PV-Based DERs

The event detection approach proposed in Section II provides a basis for event-triggered and centralized control of DERs in smart distribution grids. An example of event-triggered control

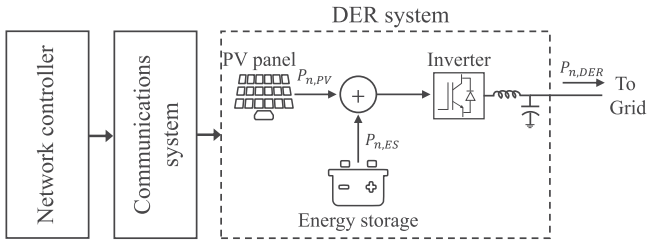


Fig. 3. Conceptual diagram of event-triggered control of DER systems that utilize solar PV sources and storage devices.

of PV-based DER systems is illustrated in Fig. 3. The DER systems are equipped with energy storage devices that compensate for shortage/surplus of energy during unexpected solar power events. In this control mechanism, the network controller communicates with the DER systems whenever an event is detected. The advantage of this event-triggered control scheme is that the network controller can effectively mitigate voltage fluctuations with minimal communications cost.

As shown in Fig. 3, the output power of the DER system in the zone z_n is given by [23]

$$P_{n,\mathrm{DER}}(k) = \eta_n \Big(P_{n,\mathrm{PV}}(k) + P_{n,\mathrm{ES}}(k) \Big) \tag{21}$$

where η_n shows the efficiency of the inverter, and $P_{n,\mathrm{PV}}$ is the power generated by the solar PV panel. $P_{n,\mathrm{ES}}$ denotes the compensation power that is injected by the energy storage device in the DER system. In order to mitigate the ramp-rate of the solar PV power, the compensation power can be determined by a moving average filtering method [29], [30].

Let $W_n(i)$ denote the size of the moving average window for the ith detection interval. The instantaneous compensation power is thus given by

$$P_{n,\mathrm{ES}}(k) = -P_{n,\mathrm{PV}}(k) + \sum_{j=0}^{W_n(i)-1} \frac{P_{n,\mathrm{PV}}(k-j)}{W_n(i)} \tag{22}$$

where $iN_D - N_D \le k \le iN_D - 1$. The window size $W_n(i)$ controls the smoothing effect in the ith detection interval. Equation (22) points out that the network controller can decrease the power fluctuations by sending the proper values of $W_n(i)$ to the DER systems. Algorithm 4 explains the procedure for updating the parameters W_n based on the estimated features of DER events, knowing that $\lceil a \rceil$ shows the nearest integer number greater than or equal to a.

The control mechanism begins with the initial window size $W_n(1)$ and the fixed step size ΔW. If no event is detected, then the network controller does not change the window size for the next detection interval. Once at lease one DER event is detected within the current detection interval, the network controller increases the window size for the next interval to further smooth the PV power fluctuations.

When a given zone indicates events, the increase in the window size is proportional to the number of events within the current detection interval, according to (24). If the given zone does not indicate any event but the other zones indicate events, then the history of spatio-temporal correlations between DER events is used for updating the window size. In such cases, the increase in the window size is determined based on the sum of the similarity degrees over the past detection intervals, as given by (25). An advantage of using (25) is that the network controller accounts for the DER events that are not directly observed, and therefore, more resource fluctuations can be suppressed in the zone of interest.

Algorithm 4: Procedure of Updating Moving Average Window.

Given $W_n(1)$, step size ΔW, and $i \ge 1$:

1) Find the set of zones that have indicated at least one event:

$$\mathcal{N}(i) = \Big\{ n \,|\, |\mathcal{M}_n(i)| > 0 \Big\} \tag{23}$$

2) If $|\mathcal{N}(i)| = 0$, then set $W_n(i+1) = W_n(i)$, for $1 \le n \le N_z$. Otherwise, go to step 3.

3) For every $m \in \mathcal{N}(i)$ set

$$W_m(i+1) = W_m(i) + |\mathcal{M}_m(i)| \Delta W, \tag{24}$$

and for every $m' \notin \mathcal{N}(i)$ set the parameters

$$W_{m'}(i+1) = W_{m'}(i) + \Big\lceil \sum_{m \in \mathcal{N}(i)} \sum_{j=1}^{i-1} \Gamma_{m,m'}(i-j) \Big\rceil \Delta W \tag{25}$$

4) If $W_n(i+1) \ne W_n(i)$, then send $W_n(i+1)$ to the DER system in zone z_n.

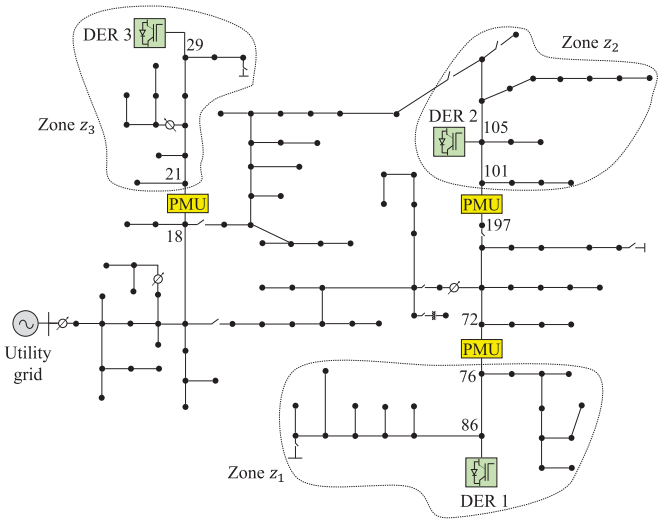


Fig. 4. IEEE 123-bus distribution network with three PV-based DER systems and PMUs.

IV. Simulation Results

Identification of DER events is investigated in the IEEE 123-bus distribution network with single-line diagram depicted in Fig. 4. As shown in Fig. 4, the network is divided into three zones, i.e., $N_z = 3$ and each zone includes a PV-based DER

system. Unless otherwise stated, the following simulation parameters are used: data reporting time $T_r = 1$ s, the length of the detection interval $\tau_D = 300$ s, the event threshold $\beta_{\text{th}} = 0.9$, the minimum separation time $\tau_s = 10$ s, and the significance level $\lambda = 2$. It is assumed that the PMUs have identical noise covariances as

$$\mathbf{\Sigma}_n = \begin{bmatrix} 0.1 & 0.02 \\ 0.02 & 0.1 \end{bmatrix}, \quad n = 1, 2, 3.$$

It should be noted that the variances of $\varepsilon_{n,P}(.)$ and $\varepsilon_{n,Q}(.)$ are given in kW^2 and kVAr^2, respectively.

A. Implementation of the Proposed Approach

The electrical components of the network including the lines, loads, PMUs, and DERs are implemented in EMTP-RV [31]. EMTP-RV simulates realistic behavior of the network with time-varying loads and solar power generation. The PDC and the network controller are implemented in MATLAB modules that carry out data gathering, event detection, and centralized control functionalities. The MATLAB modules are interfaced with EMTP-RV for bidirectional transfer of data and control commands in real-time.

B. PV-Based DER Systems

The inverters in DER systems operate under unity power factor at the full efficiency and are locally controlled based on realistic solar irradiance profiles, which are measured during a day in Varennes, Quebec [32]. The subsecond resolution of the solar irradiance allows for realistic simulation of ramp-rate DER events due to moving clouds and shadowing of objects. This methodology ensures that the spatial and temporal characteristics of energy resources are taken into account and thus event identification is assessed under realistic conditions.

Two solar irradiance profiles are adopted in the simulations [32].

1) *Variable profile:* The solar irradiance is measured on a winter day in February. In this case, the energy resource incurs remarkable fluctuations due to shadowing of intermittent clouds in the sky.
2) *Highly variable profile:* The solar irradiance is measured on a summer day in July and incurs significant and rapid fluctuations over the time. Fig. 5(a) and (b) shows the time-series of the generated power by solar PV resources in each daily scenario.

C. Simulation of Loads

The simulation parameters that describe the loads are given in Table I. The loads are modeled by determining the mean value and the range of variations of the consumed power within an interval of length τ_D. In order to simulate loads in a realistic manner, the simulator varies real/reactive powers of the loads using random walk processes [26]. The mean values of real/reactive powers are determined based on static power flow simulation of IEEE 123-bus network. The range of powers are chosen such that the loads exhibit different temporal properties

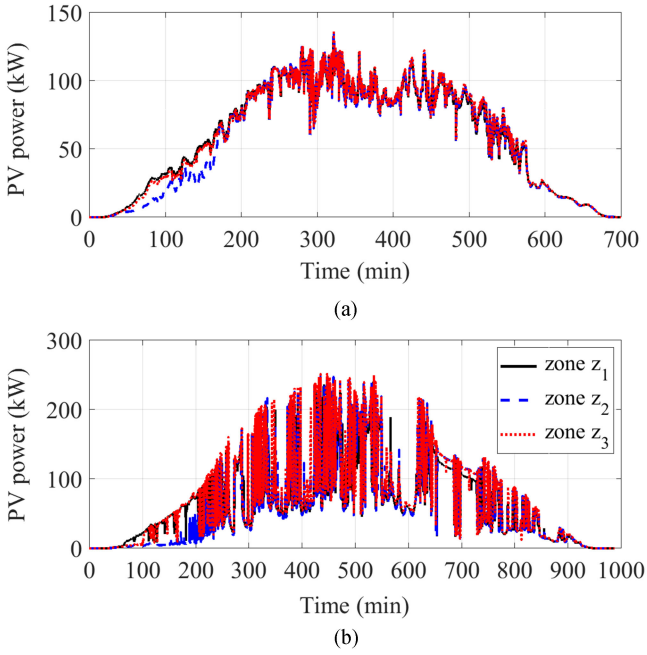


Fig. 5. Daily solar power profiles. (a) Variable solar power. (b) Highly variable solar power.

TABLE I
PARAMETERS OF THE LOADS IN THE TEST NETWORK

	Zone z_1	Zone z_2	Zone z_3
Number of loads within the zone	18	10	8
Mean of the real power (kW)	760	323	282
Mean of the reactive power (kVAr)	-365	162	141
Range of the real power (kW)	20	10	5
Range of the reactive power (kVAr)	10	4	1

within the detection interval. The largest load change is assigned to the zone z_1 as this zone has the highest range of consumed power.

D. Numerical Results and Discussions

The results shown in Fig. 6 demonstrate the impact of the reporting time after downsampling of the synchrophasors by the PDC. Fig. 6(a) shows the output power of the DER system of the zone z_2 with the variable PV profile. The irregularity begins at the time $t = 483$ min and ends at $t = 483.3$ min. Fig. 6(b) and (c) shows the event signatures with $T_r = 1$ s and $T_r = 0.5$ s, respectively. It is evident that, $T_r = 1$ s gives a very good estimate of the event duration, whereas $T_r = 0.5$ s misrepresents the event and the estimate is 13 s less than the actual event duration. In this test, the magnitude of the power does not change significantly during 0.5 s intervals. With $T_r = 0.5$ s, the DER event is partially masked by the severe noise and consequently the estimated duration becomes less than the actual event duration.

Fig. 7 shows the impact of the minimum separation time on the event identification with the highly variable PV profile. As shown in Fig. 7(a), DER 1 incurs two adjacent power fluctuations in the detection interval $[295, 300]$ min. In this test,

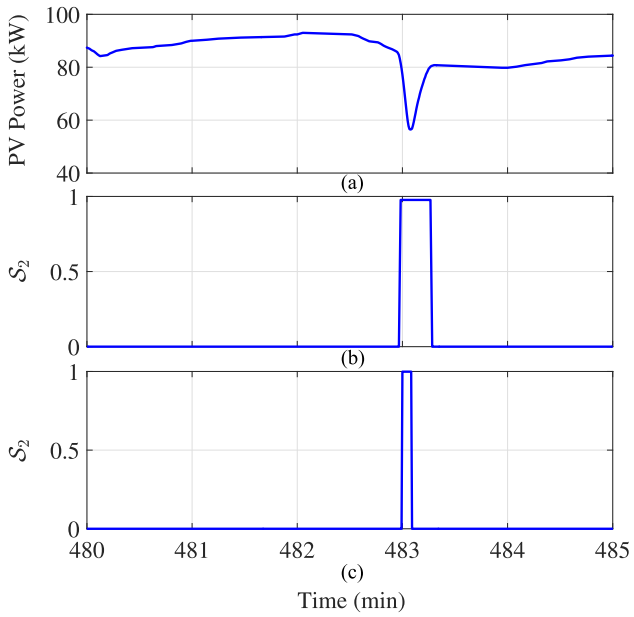


Fig. 6. Detection with different reporting times. (a) PV power in the zone z_2. (b) Signature sequence for z_2 with $T_r = 1$ s. (c) Signature sequence for z_2 with $T_r = 0.5$ s.

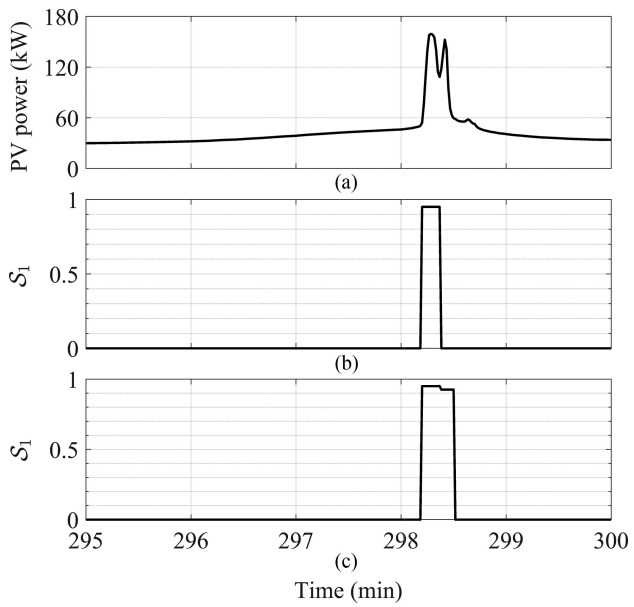


Fig. 7. Impact of minimum separation time. (a) PV power in the zone z_1. (b) Signature sequence for z_1 with $\tau_s = 10$ s. (c) Signature sequence for z_1 with $\tau_s = 5$ s.

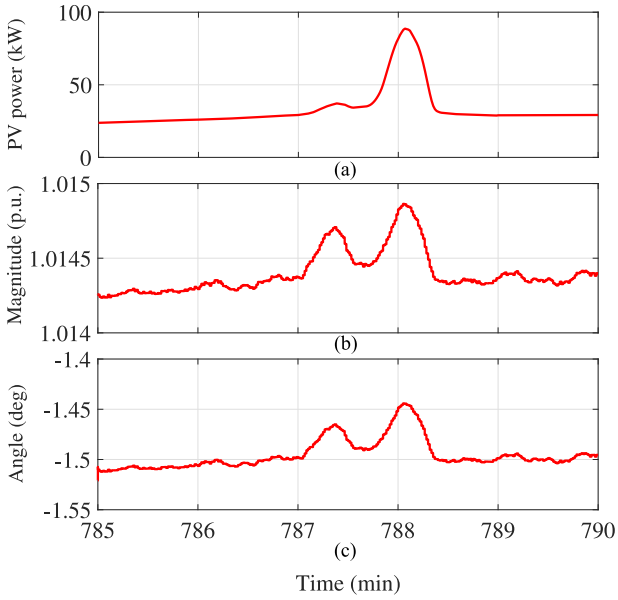


Fig. 8. Power irregularity in the zone z_3. (a) PV power. (b) Magnitude of the voltage synchrophasor. (c) Phase angle of the voltage synchrophasor.

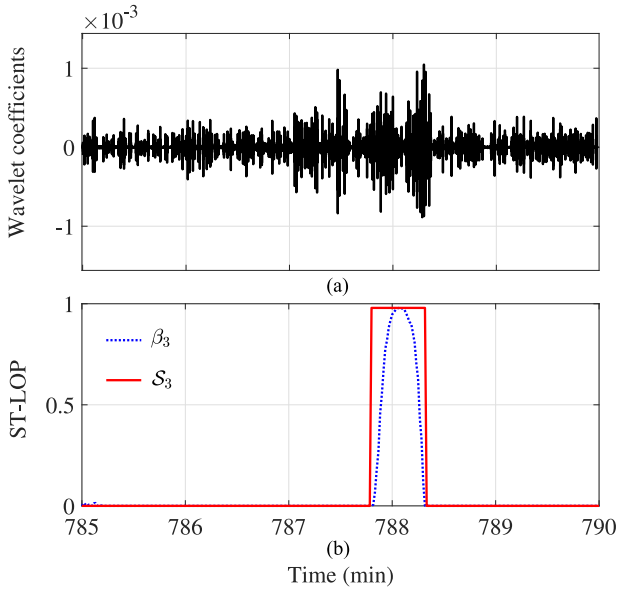


Fig. 9. Detection of the power irregularity in the zone z_3. (a) Sum of the wavelet coefficients. (b) ST-LOP sequence and the corresponding event signature give by Algorithm 3.

with $\tau_s = 10$ s, one event is detected with the signature shown in Fig. 7(b). By decreasing the minimum separation time to $\tau_s = 5$ s, two events are detected and the signature sequence becomes broader, as shown in Fig. 7(c). The average duration of events is estimated as 10 and 9 s for $\tau_s = 10$ and $\tau_s = 5$ s, respectively. The signature in Fig. 7(c) indicates that the events in that occur within the interval $[295, 300]$ min are very similar in terms of their features, i.e., magnitude and duration of the power fluctuation.

Fig. 8 demonstrates a DER event and the resulting synchrophasors of the voltage signal for the zone z_3. Fig. 9 compares the detection results with respect to the DER event demonstrated in Fig. 8. In this test, the PMU sends the synchrophasors to the PDC at the rate 60 frames/s. The small-scale phase angle variations stem from the noise, the load changes in the zone z_3, as well as power generation/consumption in the other zones. Moreover, the phase angle does not exhibit any transients during the detection interval. The performance of the wavelet-based methods (see [15] and [16]) is shown in Fig. 9(a). The vertical

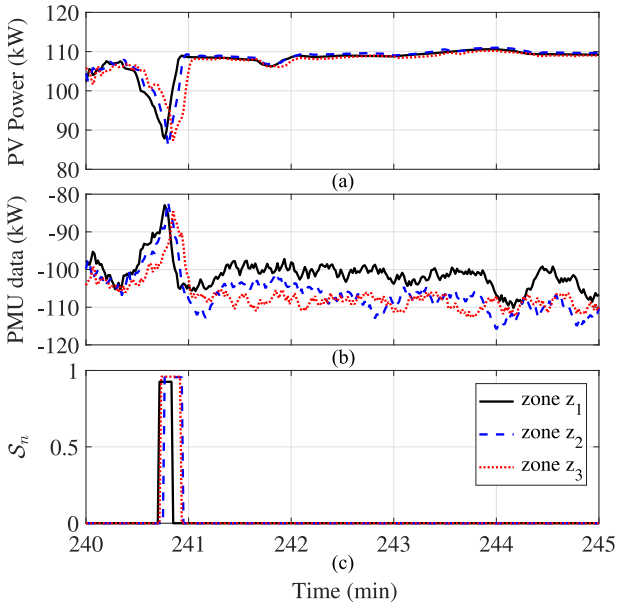


Fig. 10. Spatio-temporal correlation between DER events. (a) PV powers. (b) Real power data reported by PMUs after centering. (c) Signature sequences.

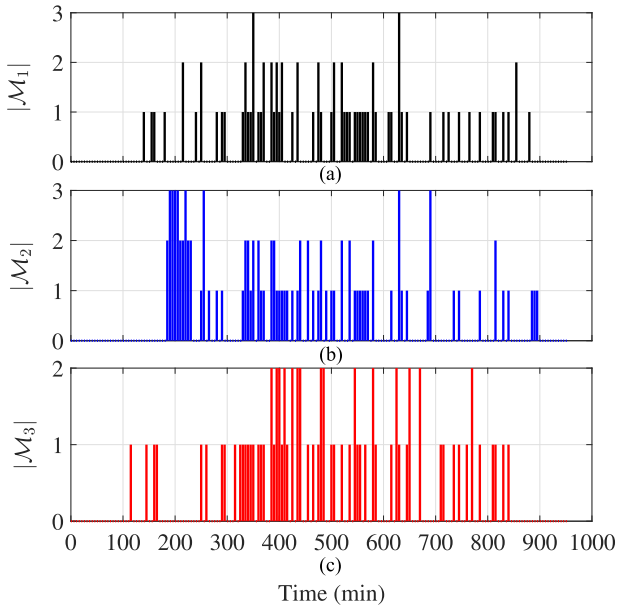


Fig. 11. Estimated number of DER events with the highly variable PV profile. (a) Zone z_1. (b) Zone z_2. (c) Zone z_3.

axis in Fig. 9(a) represents the sum of the wavelet coefficients for the phase angle data shown in Fig. 8(c). It is observed that the energy of the wavelet coefficients remains trivial when the event occurs. The ST-LOP sequence and the signature sequence are depicted in Fig. 9(b). It is seen that the ST-LOP sequence has a dominant peak, and thus, the power fluctuation is detected with estimated duration of 31 s. The wavelet-based methods fail to detect the event since abrupt transients do not exist in the synchrophasors. These results validate the effectiveness of the proposed approach and indicate its superiority over the previous methods discussed in Section I.

Estimation of event features with the variable PV profile is demonstrated in Fig. 10. Fig. 10(a) shows that the DER systems incur a steep power drop around the time instant $t = 240.8$ min. Fig. 10(b) shows the centered real power measurements reported by the PMUs. As shown in Fig. 10(c), the signature sequences for the three zones overlap and spatial correlation exists. The similarity degrees are found as $\Gamma_{1,2} = 0.52$, $\Gamma_{1,3} = 0.69$, and $\Gamma_{2,3} = 0.86$. The large values of the similarity degrees reveal that the DER systems incur power dips with remarkable temporal correlation. In this test, the average duration of events are estimated as: $\tau_1 = 7$ s, $\tau_2 = 10$ s, and $\tau_3 = 11$ s. The smallest magnitude of event signature belongs to the zone z_1, which is justified by the fact that the loads in the zone z_1 vary in a large range. This result confirms that the proposed method can properly identify DER events even when the resource fluctuation is moderate.

Fig. 11(a)–(c) demonstrate the estimated number of DER events over the time. In this simulation, the highly variable PV profile is used and no energy storage device is employed by the DER systems. As expected, multitude of ramp-rate events are detected within the detection intervals during this cloudy day. The highest number of DER events is observed in the zone z_2,

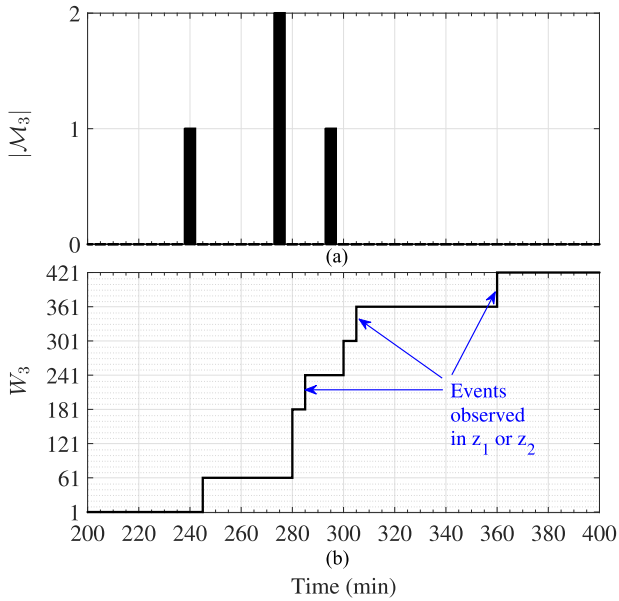


Fig. 12. Event-triggered control of the DER system in the zone z_3 with the variable PV profile. (a) Estimated number of events. (b) Size of the moving average window.

where 98 events are detected with total average duration of 445 s. The zones z_1 and z_3 indicate 76 and 75 events, respectively. However, the total average duration of events in the zones z_1 and z_3 are 423 and 452 s, respectively. It turns out that the DER system in the zone z_3 incurs longer events on average. Specifically, the average duration of events in the zone z_3 is around 6 s. The results point out that some of the ramp-rate solar power fluctuations may last up to several seconds.

The event-triggered control of the DER system in the zone z_3 is investigated in Fig. 12. In this test, Algorithm 4 uses an

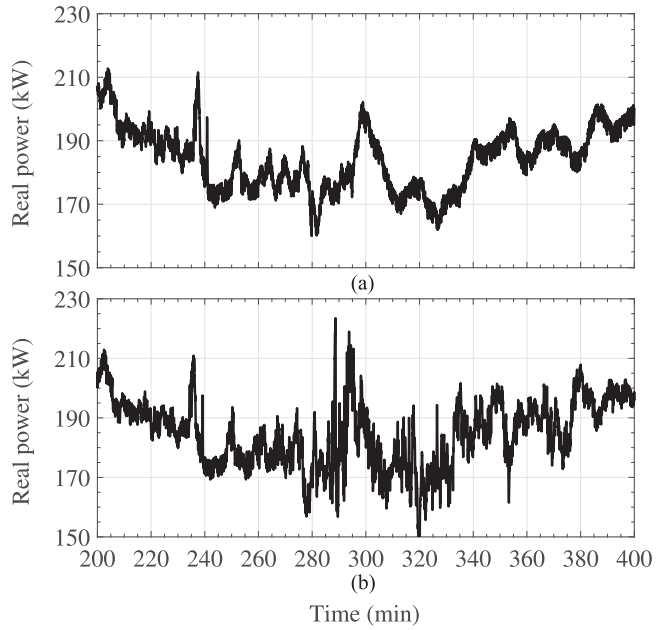


Fig. 13. Real power data reported by the PMU of the zone z_3 with the variable PV profile. (a) With controlled window size. (b) With no windowing.

initial window size of $W_3(1) = 1$ and the step size $\Delta W = 60$. Fig. 12(a) and (b) shows that the network controller updates the window size upon detection of solar power fluctuations. Two events are detected in the zone z_3 within the detection interval $[275, 280]$ min. Hence, the window size jumps from 61 to 181 and the smoothing effect doubles. It is worth noting that the window size increases at time instants 285, 305, and 360 min when the zone z_3 does not indicate any event. The rise in the window size in such cases stems from the nonzero similarity degrees, which are estimated in the interval $[275, 280]$ min. This observation reveals that the network controller can take advantage of the spatio-temporal correlations between DER events in order to enhance centralized control applications.

Fig. 13 demonstrates the exchange of the real power between the zone z_3 and the rest of the network within the time interval $[200, 400]$ min. Fig. 13(a) shows the real power measurements when the window size is controlled by Algorithm 4, i.e., when W_3 is updated according to Fig. 12(b). As a result of controlling the window size, the DER system in the zone z_3 counteracts ramp-rate events and thus the power exchange incurs less fluctuations. The real power measurements when the DER system lacks the compensation mechanism are shown in Fig. 13(b). The comparison between Fig. 13(a) and (b) reveals that the network controller can diminish the power fluctuations in terms of number and magnitude. This result corroborates the effectiveness of the developed analytics for centralized control applications in smart grids that incorporate renewable energy resources.

V. Conclusion

This paper presents a data analytics approach to centralized detection of irregular resource fluctuations in smart distribution networks and microgrids. The network is divided into different detection zones, and each zone is associated with a PMU that provides real and reactive power measurements. The power fluctuations are detected and analyzed by processing the ST-LOP sequences. Multiple DER events may occur in the network and they can overlap each other. It is shown that the event signatures can be used to estimate the average duration of events and measure the spatio-temporal correlations between the events. The proposed detection method can be combined with a control strategy to smooth out ramp-rate power fluctuations and improve power quality in the network. Simulations are carried out for the IEEE 123-bus network that validate the effectiveness of the proposed approach in event-triggered control of PV systems with energy storage devices.

References

[1] T. Som and N. Chakraborty, "Studies on economic feasibility of an autonomous power delivery system utilizing alternative hybrid distributed energy resources," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 172–181, Jan. 2014.

[2] Y. Yang, P. Enjeti, F. Blaabjerg, and H. Wang, "Wide-scale adoption of photovoltaic energy: Grid code modifications are explored in the distribution grid," *IEEE Ind. Appl. Mag.*, vol. 21, no. 5, pp. 21–31, Sep./Oct. 2015.

[3] Y. Seyedi, H. Karimi, and S. Grijalva, "Distributed generation monitoring for hierarchical control applications in smart microgrids," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2305–2314, May 2017.

[4] Q. Sun *et al.*, "A comprehensive review of smart energy meters in intelligent energy networks," *IEEE Internet Things J.*, vol. 3, no. 4, pp. 464–479, Aug. 2016.

[5] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2784–2794, Nov. 2014.

[6] Y. Wang, P, Yemula, and A. Bose, "Decentralized communication and control systems for power system operation," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 885–893, Mar. 2015.

[7] A. von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: A summary of applications," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2926–2936, Nov. 2017.

[8] Y. Weng, Y. Liao, and R. Rajagopal, "Distributed energy resources topology identification via graphical modeling," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2682–2694, Jul. 2017.

[9] H. Sugihara, K. Yokoyama, O. Saeki, K. Tsuji, and T. Funaki, "Economic and efficient voltage management using customer-owned energy storage systems in a distribution network with high penetration of photovoltaic systems," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 102–111, Feb. 2013.

[10] V. R. Pandi, H. H. Zeineldin, and W. Xiao, "Determining optimal location and size of distributed generation resources considering harmonic and protection coordination limits," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1245–1254, May 2013.

[11] V. C. Nikolaidis, E. Papanikolaou, and A. S. Safigianni, "A communication-assisted overcurrent protection scheme for radial distribution systems with distributed generation," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 114–123, Jan. 2016.

[12] X. Zhang and S. Grijalva, "A data-driven approach for detection and estimation of residential PV installations," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2477–2485, Sep. 2016.

[13] Y. Ge, A. J. Flueck, D.-K. Kim, J.-B. Ahn, J.-D. Lee, and D.-Y. Kwon, "Power system real-time event detection and associated data archival reduction based on synchrophasors," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2088–2097, Jul. 2015.

[14] M. Rafferty, X. Liu, D. Laverty, and S. McLoone, "Real-time multiple event detection and classification using moving window PCA," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2537–2548, Sep. 2016.

[15] D. Kim, T. Y. Chun, S.-H. Yoon, G. Lee, and Y.-J. Shin, "Wavelet-based event detection method using PMU data," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1154–1162, May 2017.

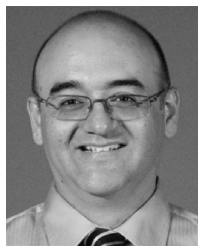
[16] S. S. Negi, N. Kishor, K. Uhlen, and R. Negi, "Event detection and its signal characterization in PMU data stream," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 3108–3118, Dec. 2017.
[17] Y. Zhou *et al.*, "Abnormal event detection with high resolution micro-PMU data," in *Proc. Power Syst. Comput. Conf.*, Jun. 2016, pp. 1–7.
[18] S. Das and T. S. Sidhu, "Application of compressive sampling in synchrophasor data communication in WAMS," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 450–460, Feb. 2014.
[19] Y. Seyedi, H. Karimi, and S. Grijalva, "Coordinated protection and control based on synchrophasor data processing in smart distribution networks," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 634–645, Jan. 2018.
[20] R. Pourramezan, Y. Seyedi, H. Karimi, G. Zhu, and M. Mont-Briant, "Design of an advanced phasor data concentrator for monitoring of distributed energy resources in smart microgrids," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 3027–3036, Dec. 2017.
[21] V. P. Singh, N. Kishor, and P. Samuel, "Load frequency control with communication topology changes in smart grid," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1943–1952, Oct. 2016.
[22] J. Zhao, G. Zhang, M. La Scala, and Z. Wang, "Enhanced robustness of state estimator to bad data processing through multi-innovation analysis," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1610–1619, Aug. 2017.
[23] M. J. E. Alam, K. M. Muttaqi, and D. Sutanto, "A novel approach for ramp-rate control of solar PV using energy storage to mitigate output fluctuations caused by cloud passing," *IEEE Trans. Energy Convers.*, vol. 29, no. 2, pp. 507–518, Jun. 2014.
[24] *Communication Networks and Systems for Power Utility Automation—Part 1: Introduction and Overview*, IEC Technical Report 61850-1, Mar. 14, 2013.
[25] H. P. Kriegel, P. Kroger, E. Schubert, and A. Zimek, "LoOP: Local outlier probabilities," in *Proc. 18th ACM Conf. Inf. Knowl. Manage.*, Nov. 2009, pp. 1649–1652.
[26] S. Barker, S. Kalra, D. Irwin, and P. Shenoy, "Empirical characterization, modeling, and analysis of smart meter data," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1312–1327, Jul. 2014.
[27] A. Ahmad, N. Javaid, M. Guizani, N. Alrajeh, and Z. A. Khan, "An accurate and fast converging short-term load forecasting model for industrial applications in a smart grid," *IEEE Trans. Ind. Informat.*, vol. 13, no. 5, pp. 2587–2596, Oct. 2017.
[28] G. Valverde, A. T. Saric, and V. Terzija, "Stochastic monitoring of distribution networks including correlated input variables," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 246–255, Feb. 2013.
[29] A. Ellis *et al.*, "PV output smoothing with energy storage," in *Proc. Photovolt. Spec. Conf.*, Jun. 2012, pp. 1523–1528.
[30] A. Hoke, R. Butler, J. Hambrick, and B. Kroposki, "Steady-state analysis of maximum photovoltaic penetration levels on typical distribution feeders," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 350–357, Apr. 2013.
[31] J. Mahseredjian, S. Dennetire, L. Dub, B. Khodabakhchian, and L. Grin-Lajoie, "On a new approach for the simulation of transients in power systems," *Electric Power Syst. Res.*, vol. 77, no. 11, pp. 1514–1520, Sep. 2007.
[32] NRCAN, "High Resolution Solar Radiation Datasets," 2016. [Online]. Available: http://www.nrcan.gc.ca/energy/renewable-electricity/solar-photovoltaic/18409

Younes Seyedi received the B.Sc. degree in electrical engineering from the University of Tehran, Tehran, Iran, and the Ph.D. degree in electrical engineering-power systems from Polytechnique Montreal, Montreal, QC, Canada, in 2010 and 2017, respectively.

He is currently a Research Fellow with the Department of Electrical Engineering, Polytechnique Montreal, Montreal, QC, Canada. His recent research is focused on smart grids with emphasis on data analytics and intelligent learning algorithms for grid monitoring, control, and protection applications.

Houshang Karimi (M'07–SM'12) received the Ph.D. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2007.

He was an Assistant Professor with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran, from 2009 to 2012. From June 2012 to January 2013, he was a Visiting Professor with the ePower Lab, Department of Electrical and Computer Engineering, Queens University, Kingston, ON, Canada. In 2013, he joined the Department of Electrical Engineering, Polytechnique Montreal, Montreal, QC, Canada, where he is currently an Associate Professor. His research interests include control systems, microgrid control, and smart grids.



Santiago Grijalva (M'02–SM'07) received the graduate degrees in electrical and computer engineering, the M.Sc. and Ph.D. degrees from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 1999 and 2002, respectively.

He is the Georgia Power Distinguished Professor of electrical and computer engineering and the Director of the Advanced Computational Electricity Systems (ACES) Laboratory, The Georgia Institute of Technology, Atlanta, GA, USA. His research interest include decentralized power system control, power system analytics and economics, and future sustainable energy systems. He has been a Principal Investigator for research under the Department of Energy, ARPA-E, EPRI, PSERC, NSF, and other industry and Government sponsors.

Dr. Grijalva, he was with PowerWorld Corporation as a Software Architect and consultant, from 2002 to 2009. From 2013 to 2014, he was on assignment to the National Renewable Energy Laboratory (NREL) as the Founding Director of the Power System Engineering Center (PSEC). He is a Member of the Federal Smart Grid Advisory Committee of the National Institute of Standards and Technology (NIST).