IEEE TRANSACTIONS ON
SMART GRID

Quickest Fault Detection in Photovoltaic Systems

Leian Chen, Shang Li, and Xiaodong Wang, *Fellow, IEEE*

***Abstract*— Photovoltaic (PV) systems play an important role in contemporary electricity production as a ubiquitous renewable energy source. However, the performance of a PV system is susceptible to unexpected faults that occur inside its various components. In this paper, we propose a quickest fault detection algorithm for PV systems under the sequential change detection framework. In particular, multiple meters are employed to measure different output signals of the PV system. The time correlation of the faulty signal and the signal correlation among different meters are exploited by a vector AR model in modeling the post-change signal. In order to tackle the difficulty that no prior knowledge about the fault is available, we develop a change detection algorithm based on the generalized local likelihood ratio test. Extensive simulation results demonstrate that the proposed method achieves high adaptivity and fast detection in dealing with various types of faults in PV systems.**

***Index Terms*— Photovoltaic (PV) array, fault detection, autoregressive model, generalized local likelihood, quickest detection.**

I. Introduction

DUE Due to the rapid deployment of renewable energy plants, such as the photovoltaic(PV) systems, especially the large-scale grid-connected PV systems for the generation of electricity, the development of effective and economical techniques to ensure their long-lasting reliable performance becomes of paramount importance. Typical faults in a PV system include the irradiance change, ground faults, line-line faults, arc faults and so on. The occurrence of faults can result in energy loss, system shutdown or even safety concerns. Therefore, efficient online monitoring and quick fault detection/diagonosis become an essential component of the PV system performance control.

A. Fault Detection in PV Systems

Most fault detection methods in the literature rely on the prediction reported by a system model which acts as reference data for comparison during the online monitoring process [1] [2] [3]. This method typically generates an expected I-V curve for a specific PV system by a well-defined model, and then it is compared with the measured operating current and voltage in real time. The model usually requires a frequently updated input (local irradiance and temperature) which is supposed to be synchronized with the instant local climate. The concept of remote monitoring is introduced which replaces frequent local environmental measurement by climate data from satellites [4] [5]. Though it reduces the local labor cost, the detection performance can degrade. Furthermore, the model-based method needs to constantly adjust the model due to the aging of the PV arrays. Overall, this approach is expensive and labor-intensive which makes it less attractive for wide adoption by the ever-growing deployment of PV systems.

An alternative approach to the model-based fault detection is the statistical methods where statistical properties of the various signals of PV arrays are exploited to perform detections. An ad-hoc method is proposed in [6] where given a series of online samples, the PV string's current deviation ratio is calculated and tracked over time to identify the occurrence of faults. And in [7], analysis of variance (ANOVA) and the Kruskal-Wallis test are applied based on the measurements at the inverter. However, the method is specific to the fault detection among PV strings or sub-arrays.

Recently, machine learning techniques have been applied to the fault detection in PV systems. Typical algorithms such as feed-forward neural networks and support vector machines are employed in [8] and [9]. In [10], a fault detection method based on extension theory with neural networks is proposed. However, the main problem with the machine learning approach is the difficulty of obtaining a comprehensive set of training data under both faulty and non-faulty conditions, since in real systems the type and feature of the fault are unknown a priori. A graph-based semi-supervised learning algorithm is proposed in [11] to detect faults with small amount of labeled data and sufficiently large number of unlabeled data to train the model.

B. Contributions

In this paper, we cast the PV fault detection problem as a sequential change detection problem with unknown post-change distributions, and develop an efficient quickest detection algorithm to detect the fault. In contrast to other existing detection methods for PV systems which require the prior knowledge to train the system model or complicated configurations with multiple sensor readings from different locations, our proposed approach does not need extensive field works or additional equipments for model validation, but only needs to monitor the commonly measured signals (power, voltage, etc.) of the PV array. We exploit both the time correlation of fault signals and the correlation among multiple simultaneously measured signals (e.g., voltage, current, power), by applying a vector autoregressive (AR) model to describe the faulty signal. Such a model manifests the fact that the burst change due to faults can exert similar or even the same impact on various components of the system at the same time.

This work was supported in part by the U.S. National Science Foundation (NSF) under grant CIF1064575, and in part by the U.S. Office of Naval Research (ONR) under grant N000141410667.

L. Chen, S. Li, and X. Wang are with the Department of Electrical Engineering, Columbia University, New York, NY, 10027 USA (e-mail: chen.leian@columbia.edu, {shang,wangx}@ee.columbia.edu).

Digital Object Identifier: 10.1109/TSG.2016.2601082

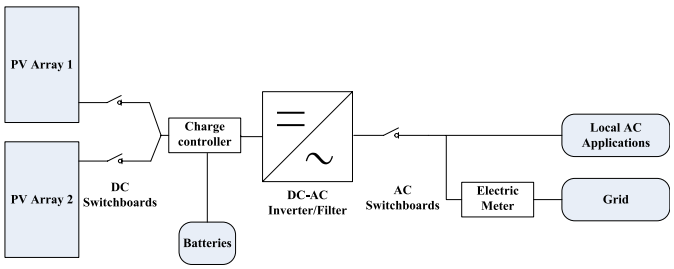


Fig. 1. A grid-connected PV system.

A major challenge is that no prediction or prior knowledge is available to help model the faulty signal. To that end, we propose a sequential fault detection scheme based on the generalized local likelihood ratio (GLLR) test, which does not need any prior knowledge on the change direction or amplitude of the signals, in contrast to classical detection algorithms such as the cumulative sum control chart (CUSUM) or the sequential probability ratio test (SPRT) used in [12].

Extensive simulation results demonstrate that our proposed algorithm outperforms the existing methods in detecting faults in PV systems with less delay and higher adaptivity.

The remainder of this paper is organized as follows. Section II gives an overview of the PV system and typical faults of interest. In Section III, we develop our proposed sequential fault detection algorithm based on the vector AR model and the GLLR test. In Section IV, we apply the proposed method in a simulated PV system and compare its performance with some existing methods. Finally, Section V concludes the paper. The Appendix contains some detailed mathematical derivations.

II. Overview of PV Systems

PV systems transform the energy from sunlight to generate electricity via an electronic process in semiconductors. As illustrated in Fig. 1, a typical PV system for residential or industrial electricity supply usually consists of the PV array which generates electricity directly from sun irradiance, and other follow-up components which are often referred to as "balance of components" (BOC).

A. Components in PV Systems

The PV array is composed of numerous PV cells which are encapsulated into PV modules. The modules are first serially connected to accumulate the desired output voltage and then connected in parallel to obtain more current. An $m \times n$ PV array is formed by m parallel strings of n serially-connected PV modules.

To utilize the electricity collected from PV arrays, the BOC transform and store the energy into the form which can be directly delivered for daily applications. The BOC basically include the mounting structures to fix and direct PV modules towards the sun, the DC-AC converters (also known as inverters) for applications requiring AC, the batteries for energy storage, and a charger regulator for smooth operation of the PV system.

B. Typical Faults in PV Systems

In an ideal PV array, the total output power is the sum of individual module powers. However, in a real system, the performance is vulnerable to environmental changes or hardware failures and it always deviates from the predicted level. As shown in Fig. 2, we next introduce four typical faults in a PV arrary.

1) Shading: The PV system is extremely sensitive to the amount of sunlight it effectively receives and a serious fault can occur when the surface of PV modules are blocked from the sunlight. Even a single blocked cell or module will result in an obvious fluctuation in the output power and voltage. If the neighboring unaffected cells in the same string can still provide enough voltage, the current going through the shaded portion will break down the junctions inside and the power will be absorbed and turned into heat. The conventional method to detect the failure by shading is based on tracking the performance of the maximum power point (MPP) in a running PV system [13] [14] [15].

2) Ground Fault: Ground fault is an unexpected connection between the circuits and the ground. The current will leak through the shortcut to the ground and results in lowered output power and voltage. In the PV system, it is usually caused by the damage to the protective insulation of normally current-carrying conductors [16]. The US National Electrical Code (NEC) requires the installation of ground-fault detection and interruption (GFDI) devices to protect the PV systems with voltage more than 50V. Even with GFDI devices, the ground fault can remain undetected when the input irradiance is not strong enough or the ground fault current is below the triggering threshold of GFDI devices [17]. In [18], the spread spectrum time domain reflectometry (SSTDR) method is applied to detect ground faults, which no longer relies on the amplitude of fault current even with a noisy signal.

3) Line-Line Fault: A line-line fault is an unexpected short circuiting between two points of different potentials among PV modules or arrays. Therefore, the electrical imbalance among PV arrays will result in overcurrent into the faulty string. Fig. 2 shows both the intra-string and cross-string line-line faults. As a conventional scheme to protect the system from line-line faults, the overcurrent protection devices in the code of NEC can only isolate the faulty circuits with large current under high illumination [19].

4) DC Arc Fault: The DC arc fault is triggered by a high power discharge among different conductors. In the PV system, it can act in either a series or parallel way. Series arcs can often take place within a module, at the wired connection between modules or at the junction box. Parallel arcs occur as two parallel conductors with different voltages are placed close to each other. In addition to its negative effects on the system efficiency, the arc faults lead to safety concerns by starting a fire. Since the DC arc faults behave in a transient way, it cannot be fully modeled as a change in the array. To detect such faults, existing methods are based on the spectrum analysis of the voltage or current waveforms [20] [21]. On the hardware side, arc fault circuit interrupters (AFCIs) which are installed on individual strings provide protections from arc

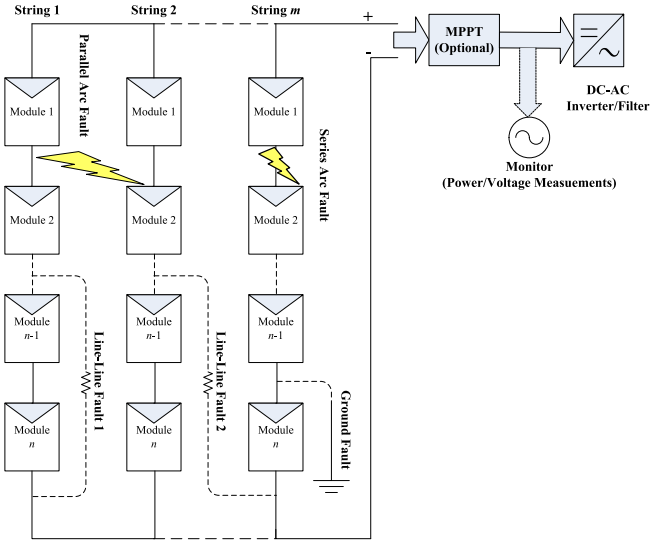


Fig. 2. Typical faults in a PV array.

faults and advanced methods to improve its performance are studied in [22] [23].

In this paper, we focus on the detection of faults due to irradiance change (shading), line-line faults and ground faults in a grid-connected PV system by monitoring the mean power and mean voltage at the output of PV arrays (i.e., DC side). The basic idea is to monitor the statistical properties of the observed signals over time and detect the change due to the occurrence of faults in a quickest way.

III. Sequential Fault Detection

In this section, we first formulate the problem of sequential fault detection in the PV array based on the vector autoregressive (AR) model. Then a quickest fault detection method based on the generalized local likelihood ratio (GLLR) test is presented. We assume that the statistical properties of the measurement signal (voltage, current or power) under normal conditions, i.e., before the fault occurs, are known (or available from empirical data) while properties of the signal after the change are totally unknown.

A. Vector AR Modelling of Faults

Before the fault occurs at some unknown time $t = t_0$, the signal observed at the output of PV arrays consists of the nominal (desired) DC output as shown in Fig. 2, which is supposed to be a constant, and the measurement noise. After t_0, the DC output is corrupted by both the faulty signal and noise.

Under the normal operating conditions, the DC outputs of a PV system, such as voltage, current and power, are all constants that are known by design. Then we can subtract these constant signals from the corresponding measurements to obtain the post-processing signals that only comprise the faulty signal and noise. Suppose that K DC output parameters are measured for the PV system monitoring, then we can obtain a post-processing signal $\boldsymbol{y}_t \in \mathbb{R}^K$ that consists of the noise and possibly the disturbance signal. Specifically, before the fault occurs at $t = t_0$, the post-processing signal consists of the measurement noise $\boldsymbol{v}_t$ only, which is modeled as a white Gaussian process, i.e.,

$$\boldsymbol{y}_t = \boldsymbol{v}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{v}}), \quad t < t_0, \tag{1}$$

where $\boldsymbol{R}_{\boldsymbol{v}}$ is the $K \times K$ covariance matrix of the noise $\boldsymbol{v}_t$, which is positive-semidefinite and therefore can be written as $\boldsymbol{R}_{\boldsymbol{v}} = \boldsymbol{V}\boldsymbol{V}^T$ with $\boldsymbol{V} \in \mathbb{R}^{K\times K}$.

The post-processing signal after the fault occurs consists of the disturbance signal and the measurement noise, i.e.,

$$\boldsymbol{y}_t = \boldsymbol{x}_t + \boldsymbol{v}_t, \quad t \geq t_0, \tag{2}$$

where $\boldsymbol{x}_t \in \mathbb{R}^K$ is the disturbance signal caused by the faults. Different from the measurement noise $\boldsymbol{v}_t$ that is uncorrelated in time, the faulty signal $\boldsymbol{x}_t$ is correlated in time [24], which is the basis for the proposed fault detection scheme in this paper. In particular, we use an autoregressive (AR) model to characterize the statistical property of the faulty signal $\boldsymbol{x}_t$. The AR model has been employed to characterize the disturbance [25] [26] and inter-area oscillations [27] in power grid systems, the gear tooth fault signals in mechanical systems [28], speech signals [29], etc. Specifically, the faulty signal in the PV system is modeled as

$$\boldsymbol{x}_t = \tilde{\boldsymbol{\mu}} + \sum_{j=1}^{p} \boldsymbol{A}_j \left(\boldsymbol{x}_{t-j} - \tilde{\boldsymbol{\mu}}\right) + \boldsymbol{\omega}_t \tag{3}$$

where p denotes the order of the AR model, $\tilde{\boldsymbol{\mu}} \in \mathbb{R}^K$ is the mean vector, $\boldsymbol{\omega}_t \in \mathbb{R}^K$ is the innovation noise vector, and $\boldsymbol{A}_j \in \mathbb{R}^{K\times K}$, $j = 1, \ldots p$, are the matrix coefficients. Substituting (3) into (2), we have

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \sum_{j=1}^{p} \boldsymbol{A}_j \boldsymbol{y}_{t-j} + \boldsymbol{u}_t, \quad t \geq t_0 \tag{4}$$

where $\boldsymbol{\mu} \triangleq \tilde{\boldsymbol{\mu}} - \sum_{j=1}^{p} \boldsymbol{A}_j \tilde{\boldsymbol{\mu}}$ and $\boldsymbol{u}_t \triangleq \left(\boldsymbol{v}_t - \sum_{j=1}^{p} \boldsymbol{A}_j \boldsymbol{v}_{t-j} + \boldsymbol{\omega}_t\right) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{u}})$ with $\boldsymbol{R}_{\boldsymbol{u}} = \boldsymbol{R}_{\boldsymbol{v}} + \sum_{j=1}^{p} \boldsymbol{A}_j \boldsymbol{R}_{\boldsymbol{v}} \boldsymbol{A}_j^T + \boldsymbol{R}_{\boldsymbol{\omega}}$ reflecting the impacts of both the disturbance signal and the measurement noise. Since $\boldsymbol{R}_{\boldsymbol{u}}$ is positive-semidefinite, we can write $\boldsymbol{R}_{\boldsymbol{u}} = \boldsymbol{U}\boldsymbol{U}^T$ with $\boldsymbol{U} \in \mathbb{R}^{K\times K}$.

The cross-validation method [30] is employed to justify the validity of the above AR modeling of the disturbance signal. The details are given in Section III where we observe that the normalized root-mean-squared error (NRMSE) varies from 2% to 12% for different orders p, which demonstrates the high effectiveness of the AR model in characterizing the faulty signals in PV systems.

B. Sequential Change Detection Formulation

As given in (1) and (4), the statistical models for signals before and after the fault occurrence are the basis to formulate a change detection problem which aims at the quickest reaction to the sudden change [31] [32]. The basic strategy is to utilize the sequential measurements to achieve high time resolution, and thus minimize the detection delay subject to the constraint on the false alarm period. At each time instant k, the detector takes the current measurement $\boldsymbol{y}_k \in \mathbb{R}^K$, and computes a decision statistic g_k, based on which it decides whether or not to declare a change. The stopping time T is the instant

that a change is first declared. The optimal decision rule is obtained by finding the minimum expected delay after the change occurs at $t = t_0$, given that the false alarm period does not exceed γ, i.e.,

$$\inf_T \sup_{t_0} \mathbb{E}_{t_0}\{T - t_0 \,|T \geq t_0\}, \text{ s.t. } \mathbb{E}_\infty\{T\} \geq \gamma, \quad (5)$$

where $(x)^+ \triangleq \max\{x, 0\}$, $\mathbb{E}_{t_0}\{\cdot\}$ denotes the expected delay given the change occurs at $t = t_0$, and $\mathbb{E}_\infty\{\cdot\}$ corresponds to the expectation when no change ever occurs. This statistical framework is particularly useful for promptly detecting anomaly occurrence in a wide spectrum of applications. For instance, in [33] and [34], it is applied to detect and isolate the power system transmission line outages by drawing on the generalized likelihood ratio test. In [35], this framework is applied to detect the false data injection attacks in smart grid.

For the signal models in (1) and (4), we have the following parameters before and after the change at $t = t_0$:

$$\begin{aligned}&\boldsymbol{\theta}_0 \triangleq \left[0, 0, ...0, \text{vec}(\boldsymbol{V})^T\right]^T \in \mathbb{R}^{K+K^2p+K^2}, t < t_0,\\ &\boldsymbol{\theta}_1 \triangleq \left[\boldsymbol{\mu}^T, \text{vec}\,(\boldsymbol{A})^T, \text{vec}(\boldsymbol{U})^T\right]^T \in \mathbb{R}^{K+K^2p+K^2}, t \geq t_0.\end{aligned} \quad (6)$$

where $\boldsymbol{A} = \left[\boldsymbol{A}_1, \boldsymbol{A}_2, ..., \boldsymbol{A}_p\right]$.

Denoting $\boldsymbol{Y}_j^k \triangleq \left[\boldsymbol{y}_j, \boldsymbol{y}_{j+1}, \ldots, \boldsymbol{y}_k\right]$, then we have the following conditional joint probability density functions:

$$\begin{aligned}&f_{\boldsymbol{\theta}}\left(\boldsymbol{y}_t \middle| \boldsymbol{Y}_1^{t-1}\right)\\ &= \begin{cases} \frac{1}{\sqrt{(2\pi)^K(\det \boldsymbol{R}_v)}} \exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}_{t,\boldsymbol{\theta}_0}^T \boldsymbol{R}_v^{-1}\boldsymbol{\varepsilon}_{t,\boldsymbol{\theta}_0}\right), \boldsymbol{\theta} = \boldsymbol{\theta}_0 \\ \frac{1}{\sqrt{(2\pi)^K(\det \boldsymbol{R}_u)}} \exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}_{t,\boldsymbol{\theta}_1}^T \boldsymbol{R}_u^{-1}\boldsymbol{\varepsilon}_{t,\boldsymbol{\theta}_1}\right), \boldsymbol{\theta} = \boldsymbol{\theta}_1 \end{cases}\end{aligned} \quad (7)$$

with $\boldsymbol{\varepsilon}_{t,\boldsymbol{\theta}_0} \triangleq \boldsymbol{y}_t$, and $\boldsymbol{\varepsilon}_{t,\boldsymbol{\theta}_1} \triangleq \boldsymbol{y}_t - \boldsymbol{\mu} - \sum_{j=1}^p \boldsymbol{A}_j \boldsymbol{y}_{t-j}$. (8)

The occurrence of the faults can be detected via the following sequential change detection procedure [31] which is the optimal solution to (5) when the post-change parameter $\boldsymbol{\theta}_1$ is available:

$$g_k = \max_{1 \leq j \leq k} S_j^k, \quad (9)$$

$$T = \inf\{k : g_k \geq h\}, \quad (10)$$

where h is a threshold, and

$$\begin{aligned}S_j^k &\triangleq \ln \frac{f_{\boldsymbol{\theta}_1}\left(\boldsymbol{Y}_j^k \middle| \boldsymbol{Y}_1^{j-1}\right)}{f_{\boldsymbol{\theta}_0}\left(\boldsymbol{Y}_j^k \middle| \boldsymbol{Y}_1^{j-1}\right)} = \sum_{i=j}^k \ln \frac{f_{\boldsymbol{\theta}_1}\left(\boldsymbol{y}_i \middle| \boldsymbol{Y}_{i-p}^{i-1}\right)}{f_{\boldsymbol{\theta}_0}\left(\boldsymbol{y}_i \middle| \boldsymbol{Y}_{i-p}^{i-1}\right)}\\ &= \sum_{i=j}^k \underbrace{\left[\frac{1}{2}\ln\frac{\det(\boldsymbol{R}_v)}{\det(\boldsymbol{R}_u)} + \frac{1}{2}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_0}^T \boldsymbol{R}_v^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_0} - \frac{1}{2}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T \boldsymbol{R}_u^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\right]}_{s_i}.\end{aligned} \quad (11)$$

Given a pre-determined false alarm period γ, the threshold is set by $h \approx \ln\gamma$ [31]. In our case of fault detection, the stopping time T is the first time that g_k exceeds the threshold, indicating the occurrence of the fault and terminating the detection process.

The detection procedure in (9)-(10) is equivalent to the CUSUM test where the decision statistics g_k can be computed recursively as [31]

$$g_k = (g_{k-1} + s_k)^+, \quad (12)$$

where s_k is defined in (11).

In contrast to (9) where all the measurements from the beginning are stored and needed in computing the statistic g_k at each instant k, (12) implies that the detector will discard all previous measurements and restart the calculation of g_k whenever $g_k \leq 0$. Thus, (12) can be rewritten as [31]

$$\begin{aligned}N_k &= N_{k-1}\mathbb{1}\{g_{k-1} > 0\} + 1,\\ g_k &= \left(S_{k-N_k+1}^k\right)^+,\end{aligned} \quad (13)$$

where N_k is the number of measurements at time $t = k$ since the last reset at time $t = k - N_k$, and $\mathbb{1}\{.\}$ is the indicator function.

C. Generalized Local Likelihood Ratio Test

Though powerful, the classic approach of CUSUM cannot be directly applied to detect faults in PV systems since the patterns of all possible faults in an integrated system cannot be characterized and thus the parameters of the post-change distribution are not available to implement the CUSUM test. An alternative method is the weighted CUSUM [31] which does not require the exact post-change parameters but requires a prior distribution on the unknown parameters as a weighting function to average the decision statistic. However, there is no well-justified prior distribution for PV faults, and moreover the computational complexity due to high-dimensional integration is usually too high for real-time applications.

Here we adopt an adaptive approach based on the generalized likelihood ratio test, where the post-change parameters $\boldsymbol{\theta}_1$ in (13) is replaced by its maximum likelihood estimate

$$g_k = \left(\sup_{\boldsymbol{\theta}_1} S_{k-N_k+1}^k\right)^+. \quad (14)$$

Thus the key to obtaining the detection rule is to obtain g_k based on the calculation of S_j^k. In what follows, we present an approximation to S_j^k based on the local assumption, $\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_0$, and the second-order expansion of S_j^k.

The basic idea of local detection is to approximate the decision statistic by linearly expanding around the pre-change parameter. Specifically, assuming the parameter change is small and denoting $\boldsymbol{\delta} \triangleq \boldsymbol{\theta}_1 - \boldsymbol{\theta}_0$, we use the second-order expansion of the conditional log-likelihood ratio in (11):

$$\begin{aligned}S_j^k &\approx \bar{S}_j^k\\ &= \boldsymbol{\delta}^T\left(\sum_{i=j}^k \frac{\partial s_i}{\partial \boldsymbol{\theta}_1}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right) + \frac{1}{2}\boldsymbol{\delta}^T\left(\sum_{i=j}^k \frac{\partial^2 s_i}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1^T}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right)\boldsymbol{\delta}\\ &= \boldsymbol{\delta}^T \sum_{i=j}^k \boldsymbol{z}_i - \frac{k-j+1}{2}\boldsymbol{\delta}^T\left(\sum_{i=j}^k \frac{\boldsymbol{W}_i}{k-j+1}\right)\boldsymbol{\delta},\end{aligned} \quad (15)$$

where $\boldsymbol{z}_i \triangleq \frac{\partial s_i}{\partial \boldsymbol{\theta}_1}\big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}$ and $\boldsymbol{W}_i \triangleq -\frac{\partial^2 s_i}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1{}^T}\big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}$. Note that $\mathbb{E}_{\boldsymbol{\theta}_0}\{\boldsymbol{W}_i\}$ is actually the Fisher information of the parameter $\boldsymbol{\theta}_0$, i.e., $\mathcal{I}(\boldsymbol{\theta}_0) = \mathbb{E}_{\boldsymbol{\theta}_0}\{\boldsymbol{W}_i\} = \mathbb{E}\{-\frac{\partial^2 \ln f_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\}$.

The explicit expressions of $\boldsymbol{z}_i$ and $\boldsymbol{W}_i$ are given in the Appendix A. Note that since (15) is a quadratic function of $(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_0)$, the optimization problem can be analytically solved by replacing $S_{k-N_k+1}^k$ in (14) with its second-order approximation $\bar{S}_{k-N_k+1}^k$, i.e., $\sup_{\boldsymbol{\theta}_1} S_{k-N_k+1}^k \approx \sup_{\boldsymbol{\theta}_1-\boldsymbol{\theta}_0} \bar{S}_{k-N_k+1}^k$. Based on the local assumption $\boldsymbol{\theta}_0 \approx \boldsymbol{\theta}_1$, as $k - j + 1 \to \infty$,

$$\sum_{i=j}^k \frac{\boldsymbol{W}_i}{k-j+1} \to \mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{W}_i\} \approx \mathbb{E}_{\boldsymbol{\theta}_0}\{\boldsymbol{W}_i\}. \quad (16)$$

$$\mathcal{I}(\boldsymbol{\theta}_0) = \mathbb{E}_{\boldsymbol{\theta}_0}\{\boldsymbol{W}_i\} = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{v}}^{-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_p \otimes \boldsymbol{R}_{\boldsymbol{v}} \otimes \boldsymbol{R}_{\boldsymbol{v}}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \begin{array}{l} -\boldsymbol{K}\left[\left(\boldsymbol{V}^{-1}\right)^T \otimes \left(\boldsymbol{V}^{-1}\right)\right] - \boldsymbol{I}_K \otimes \boldsymbol{R}_{\boldsymbol{v}}^{-1} \\ +4\left(\boldsymbol{V}^T \otimes \boldsymbol{I}_K\right)\left[\boldsymbol{N}_K\left(\boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{V} \otimes \boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\right] \end{array} \end{bmatrix}, \tag{17}$$

From the Appendix, we have $\mathbb{E}_{\boldsymbol{\theta}_0}\{\boldsymbol{W}_i\}$ in (17) with $\otimes$ denoting the Kronecker product. The commutation matrix $\boldsymbol{K} \in \mathbb{R}^{K^2 \times K^2}$ is defined such that $\boldsymbol{K}\,\mathrm{vec}(\boldsymbol{X}) = \mathrm{vec}\left(\boldsymbol{X}^T\right)$ for any matrix $\boldsymbol{X} \in \mathbb{R}^{K \times K}$, and $\boldsymbol{N}_K = \frac{1}{2}\left(\boldsymbol{I}_{K^2} + \boldsymbol{K}\right)$.

We now assume that the parameter $\boldsymbol{\theta}_0$ before change is known and that the parameter $\boldsymbol{\theta}_1$ after change is on the surface of an ellipsoid centered at $\boldsymbol{\theta}_0$, i.e.,

$$\boldsymbol{\theta}(t) = \begin{cases} \boldsymbol{\theta}_0, & t < t_0 \\ \boldsymbol{\theta} : (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \mathcal{I}(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) = b^2, & t \geq t_0 \end{cases} \tag{18}$$

where $b > 0$ is small. Based on the local assumption and large sample size, and decomposing $\mathcal{I}(\boldsymbol{\theta}_0) = \boldsymbol{B}^T\boldsymbol{B}$, (15) is approximated as

$$\begin{aligned} \tilde{S}_j^k &\triangleq \sup_{\boldsymbol{\delta}^T\mathcal{I}(\boldsymbol{\theta}_0)\boldsymbol{\delta} = b^2} \bar{S}_j^k \\ &\approx \sup_{\boldsymbol{\delta}^T\mathcal{I}(\boldsymbol{\theta}_0)\boldsymbol{\delta} = b^2} \boldsymbol{\delta}^T \textstyle\sum_{i=j}^k \boldsymbol{z}_i - \frac{k-j+1}{2}\boldsymbol{\delta}^T\mathcal{I}(\boldsymbol{\theta}_0)\boldsymbol{\delta} \\ &= \sup_{\tilde{\boldsymbol{\theta}}^T\tilde{\boldsymbol{\theta}} = b^2} \tilde{\boldsymbol{\theta}}^T \textstyle\sum_{i=j}^k \tilde{\boldsymbol{z}}_i - \frac{k-j+1}{2}b^2 \\ &= b\left\|\textstyle\sum_{i=j}^k \tilde{\boldsymbol{z}}_i\right\| - \frac{k-j+1}{2}b^2, \end{aligned} \tag{19}$$

where $\tilde{\boldsymbol{\theta}} = \boldsymbol{B}\boldsymbol{\delta}$, $\tilde{\boldsymbol{z}}_i = \left(\boldsymbol{B}^T\right)^{-1}\boldsymbol{z}_i$ and the last equality follows from $\left|\boldsymbol{x}^T\boldsymbol{y}\right| \leq \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$. Note that $\tilde{\boldsymbol{z}}_i$ involves only the observed signals and the covariance of the measurement noise $\boldsymbol{R}_{\boldsymbol{v}}$ which can be obtained under the normal operating condition.

Given the approximate decision statistics in (19), the detection rule of the proposed GLLR method is summarized by the pseudo code in Algorithm 1.

D. Discussions

1) Statistical Detectability: The detectability of the sequential change detection method introduced in [31] provides a guideline to properly choose the parameters for a specific algorithm which is b in our case. The two distributions before and after the occurrence of a fault is detectable if and only if $\mathbb{E}_{\boldsymbol{\theta}_0}\{\phi_k\} < 0 < \mathbb{E}_{\boldsymbol{\theta}_1}\{\phi_k\}$, where ϕ_k is the k^{th} sample statistic, which implies the statistic of N samples, $S_1^N \triangleq \sum_{k=1}^N \phi_k$, shifts in different directions before and after the fault occurrence. With the specific model in this paper, where S_1^N cannot be expressed by the sum of individual sample statistics, the corresponding quantity is defined as $\mathbb{E}_{\boldsymbol{\theta}}\{\phi_k\} \triangleq \lim_{N\to\infty}\frac{1}{N}\mathbb{E}_{\boldsymbol{\theta}}\left\{S_1^N\right\}$. With (19), we have

$$\begin{aligned} &\mathbb{E}_{\boldsymbol{\theta}}\left\{\tilde{\phi}_k\right\} \triangleq \lim_{N\to\infty}\tfrac{1}{N}\mathbb{E}_{\boldsymbol{\theta}}\left\{\tilde{S}_1^N\right\} \\ &= \lim_{N\to\infty}\tfrac{b}{N}\mathbb{E}_{\boldsymbol{\theta}}\left\{\left\|\textstyle\sum_{i=1}^N \tilde{\boldsymbol{z}}_i\right\|\right\} - \frac{b^2}{2}, \quad \boldsymbol{\theta} = \boldsymbol{\theta}_0, \boldsymbol{\theta}_1. \end{aligned} \tag{20}$$

We choose b such that $\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\tilde{\phi}_k\right\} < 0 < \mathbb{E}_{\boldsymbol{\theta}_1}\left\{\tilde{\phi}_k\right\}$. Note that the signal before the fault occurrence only contains the Gaussian white noise, and it can be shown that for any positive value of b we have

$$\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\tilde{\phi}_k\right\} = -b^2\Big/2 < 0. \tag{21}$$

After the fault occurrence, we need to have $0 < b < \lim_{N\to\infty}\frac{2}{N}\mathbb{E}_{\boldsymbol{\theta}_1}\left\{\left\|\sum_{i=1}^N \tilde{\boldsymbol{z}}_i\right\|\right\}$ such that $\mathbb{E}_{\boldsymbol{\theta}_1}\left\{\tilde{\phi}_k\right\} > 0$.

2) Delay Minimization: The bounds of b above ensure the detectability of the GLLR algorithm. Furthermore, we can find a proper value of b which is aimed to minimize the detection delay. In the asymptotic regime of large false alarm period γ, the expected delay $\mathbb{E}_{\boldsymbol{\theta}_1}\{T\}$ can be approximated as [31]:

$$\mathbb{E}_{\boldsymbol{\theta}_1}\{T\} \approx \frac{\ln\gamma}{\mathbb{E}_{\boldsymbol{\theta}_1}\left\{\tilde{\phi}_k\right\}} = \frac{\ln\gamma}{\lim_{N\to\infty}\frac{b}{N}\mathbb{E}_{\boldsymbol{\theta}_1}\left\{\left\|\sum_{i=1}^N \tilde{\boldsymbol{z}}_i\right\|\right\} - \frac{b^2}{2}}. \tag{22}$$

It implies when $\mathbb{E}_{\boldsymbol{\theta}_1}\left\{\tilde{\phi}_k\right\}$ attains its maximum, i.e., $b = \lim_{N\to\infty}\frac{1}{N}\mathbb{E}_{\boldsymbol{\theta}_1}\left\{\left\|\sum_{i=1}^N \tilde{\boldsymbol{z}}_i\right\|\right\}$, which is within the bounds discussed above, the expected detection delay is minimized. In practice, we can evaluate $\lim_{N\to\infty}\frac{1}{N}\mathbb{E}_{\boldsymbol{\theta}_1}\left\{\left\|\sum_{i=1}^N \tilde{\boldsymbol{z}}_i\right\|\right\}$ for a system of interest based on its specific signal properties to give a guideline for the choice of b.

IV. Simulation Results

In this section, we apply the proposed algorithm to a simulated PV system to detect typical faults introduced in Section II. The experiments are carried out for both the single-meter case ($K = 1$) and the multi-meter case ($K = 2, 3$). To demonstrate the effectiveness of the proposed algorithm, we also compare its delay performance with that of existing methods, including the weighted CUSUM in [36], the SPRT in [12], the deviation test in [6], the support vector machine method, and the semi-supervised learning method [11].

A. Data Acquisition and AR Model Validation

In our experiments, we simulated different types of fault occurrences in a grid-connected PV system constructed in Matlab by applying the toolbox SimPowerSystems. As shown in Fig. 3, the irradiance and temperature can be adjusted on the input side, and inside the two dashed boxes are the optional circuits for simulating ground faults and line-line faults respectively. The remaining part of Fig. 3 includes the common essential components to simulate a normal PV system. It consists of two 100-kW PV panel arrays (each with 5×66 PV modules) as the input, a DC-DC boost converter, a three-phase three-level voltage source converter (VSC), and a 25-kV grid as the output. The maximum power point tracking (MPPT) controller using the "Incremental Conductance and Integral Regulator" technique is implemented. Considering the application in a real large-scale PV system, we choose to utilize the key indicators for the performance of PV arrays which are commonly monitored at the output of the MPPT controller, and three tracked parameters, i.e., mean power,

Algorithm 1 The proposed sequential fault detection algorithm

1: Initialization: $k \leftarrow 0$, $N \leftarrow 1$, $\gamma \leftarrow 0$, $\tilde{S}^k_{k-N+1} \leftarrow 0$. Set the parameters h and b. Obtain the covariance of the measurement noise $\boldsymbol{R}_{\boldsymbol{v}} = \boldsymbol{V}\boldsymbol{V}^T$ under the normal operating condition. Calculate $\boldsymbol{B}$, such that $\boldsymbol{B}^T\boldsymbol{B} =$
$$\begin{bmatrix} \boldsymbol{R}_{\boldsymbol{v}}^{-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_p \otimes \boldsymbol{R}_{\boldsymbol{v}} \otimes \boldsymbol{R}_{\boldsymbol{v}}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{K}\left[\left(\boldsymbol{V}^{-1}\right)^T \otimes \left(\boldsymbol{V}^{-1}\right)\right] - \boldsymbol{I}_K \otimes \boldsymbol{R}_{\boldsymbol{v}}^{-1} \\ & & +4\left(\boldsymbol{V}^T \otimes \boldsymbol{I}_K\right)\left[\boldsymbol{N}_K\left(\boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{V} \otimes \boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\right] \end{bmatrix}.$$
2: **while** $\tilde{S}^k_{k-N+1} < h$ **do**
3: **if** $\tilde{S}^k_{k-N+1} \leq 0$ **then**
4: Discard previous decision statistics. Reset: $N \leftarrow 1$, $\gamma \leftarrow 0$.
5: **else**
6: $N \leftarrow N + 1$
7: **end if**
8: $k \leftarrow k + 1$. Take the next measurement $\boldsymbol{y}_k$. Obtain $\tilde{\boldsymbol{z}}_k = \left(\boldsymbol{B}^T\right)^{-1}$
$$\begin{bmatrix} \boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{y}_k \\ \left(\mathrm{vec}\left(\boldsymbol{Y}^{k-1}_{k-p}\right) \otimes \boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\boldsymbol{y}_k \\ \mathrm{vec}\left(-\left(\boldsymbol{V}^{-1}\right)^T + \boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{y}_k\boldsymbol{y}_k^T\boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{V}\right) \end{bmatrix}.$$
9: $\gamma \leftarrow \gamma + \tilde{\boldsymbol{z}}_k$
10: Compute $\tilde{S}^k_{k-N+1} \approx b\,\|\gamma\| - \frac{N}{2}b^2$.
11: **end while**
12: Trigger the fault alarm.

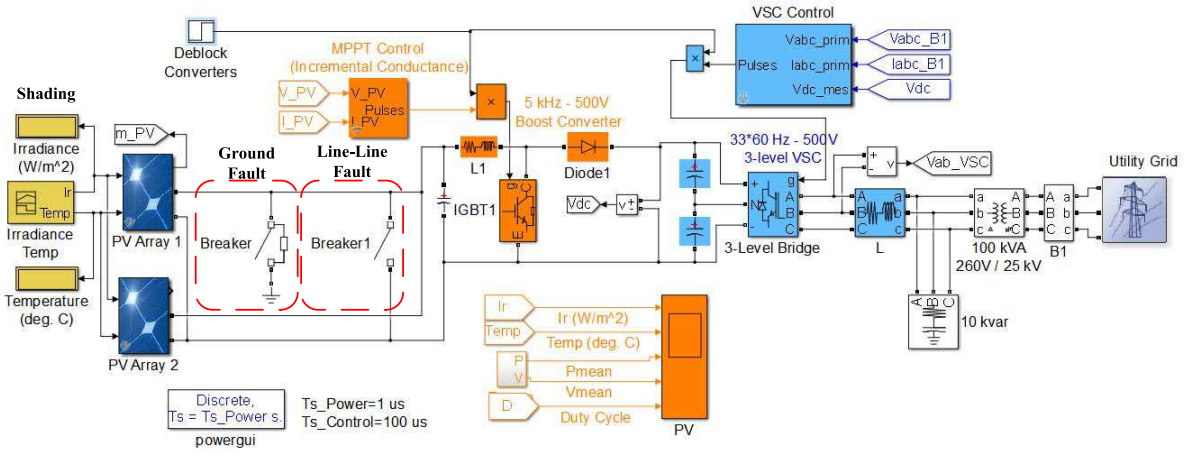


Fig. 3. The simulation system for faults induced by irradiance fluctuation, line-line faults, and ground faults.

mean voltage and duty cycle, are displayed and recorded over time.

As we observed from the readings when the simulation system was under normal condition, the outputs did not exactly stick to the theoretical value but fluctuated around it (e.g., readings of the mean power varied from 99.8kW to 100.8kW while it was supposed to be 100kW). Hence the measurement noise was considered to be included at the output and no artificial noise was added. The covariance of the measurement noise $\boldsymbol{R}_{\boldsymbol{v}}$ was calculated for each detection process based on the data under normal condition. The practical importance of this strategy lies in the fact that there are no justified ways to predict the measurement noise in various PV systems exposing to changeable conditions.

To evaluate the performance of the AR model in characterizing faulty signals, we employed the cross-validation method and compared the model predicted output with the actual output. The normalized root mean squared error (NRMSE) performance with different AR orders is presented in Table. I which demonstrates the high suitability of the AR model in our application. We also observed that the differences among detection performances with $p = 1, 2, 3, 4$ and 5 in terms of the average delays are very small, which implies that the order of the AR model in the proposed algorithm does not have a significant impact. This can be explained by the fact that in Algorithm 1, only the second element of $\tilde{\boldsymbol{z}}_k$ in Line 8 is affected by the order p.

Throughout the experiments, the sampling frequency $f_s = 10^4$ Hz and a fifth-order AR model is applied for the GLLR detector ($p = 5$) with the parameter $b = 0.5$. As for the output data fed to implement our detection algorithm, we only used the steady output after $t = 0.5$s which was considered

TABLE I
NRMSE VERSUS AR ORDER p

p	NRMSE		
	Voltage	Current	Duty cycle
1	12.1%	11.8%	11.2%
3	7.4%	7.6%	6.6%
5	4.6%	4.5%	4.1%
7	2.8%	2.5%	2.1%
9	2.0%	1.9%	1.9%

as normal operation after the simulation system completed its initialization. For the simulation in the multi-meter case, we utilized the mean power, the mean voltage, and the duty circle tracked at the output of a PV array.

In our experiments, the first three types of faults introduced in Section II were simulated to examine the performance of the proposed algorithm while arc fault was not included due to the difficulty in modeling it. For each type of fault, we first illustrate the effectiveness of our scheme by showing its decision statistics. Then, the average delays (measured by the number of samples) with varying false alarm periods are presented as a performance comparison with other existing methods. [1].

B. Existing Methods for Comparison

For the weighted CUSUM with a decision threshold g, the weighted log-likelihood ratio s_k, which involves integrating out the parameter $\boldsymbol{\theta}$, cannot be expressed in closed-form. To that end, we employ Monte Carlo method to numerically evaluate it. The Gaussian weighting function is adopted in our simulation as in [36].

The conventional SPRT method [12] is based on the sequential hypothesis test where the amplitudes of the signal before and after the fault are empirically predetermined, and the lower/upper thresholds (A/B) of the likelihood ratio are determined by the given false/miss alarm probabilities (α/β). In particular, the null hypothesis H_0 represents the normal state where the post-processing signal has zero mean and variance σ^2, and two alternative hypothesis, H_1 and H_2, denote two possible faulty outcomes where the mean is shifted by $+M$ or $-M$. Given a sequence of n measurements, $(y_1, y_2, \ldots y_n)$, the sequential log-likelihood ratio is calculated as

$$SPRT_k = \sum_{i=1}^{n} \ln \frac{p(y_i|H_k)}{p(y_i|H_0)} = \sum_{i=1}^{n} \ln \frac{f_k(y_i)}{f_0(y_i)}, k = 1, 2 \quad (23)$$

where the probability density function f_k is usually assumed to take the Gaussian form to model the electrical disturbance, yielding $SPRT_1 = \frac{M}{\sigma^2} \sum_{i=1}^{n} \left(y_i - \frac{M}{2}\right)$, and $SPRT_2 = \frac{M}{\sigma^2} \sum_{i=1}^{n} \left(-y_i - \frac{M}{2}\right)$. The upper and lower thresholds to decide whether the null hypothesis is accepted or not are given as $A = \ln\left(\frac{\beta}{1-\alpha}\right)$ and $B = \ln\left(\frac{1-\beta}{\alpha}\right)$.

In our simulations, the above SPRT method is carried out at each time n with the thresholds $A = 0$ and $B = m$ [31]:

1) H_0 is accepted if $SPRT_k \leq A$. No alarm is triggered and the algorithm is restarted by discarding the previous measurements.
2) H_1 or H_2 is accepted if $SPRT_k \geq B$. The alarm is triggered and the algorithm is restarted.
3) Further measurements are taken if neither threshold is reached until a decision can be made.

We are only interested in the delay associated with the decision of fault occurrence (H_1 or H_2) while the decision of normal operation (H_0) is not of interest. In the comparison results below, the amplitude shift of the signal is set as $M = 0.04$.

To apply the deviation test method [6] in a sequential way, at time t, we first obtain the reference amplitude V_{ref}^t (considered as under normal operation) by calculating the maximum amplitude from L previous samples, i.e., $V_{ref}^t = \max\{V_{t-L}, V_{t-L+1}, ..., V_{t-1}\}$,where L is a window size. Then the current sample's amplitude V_t is compared with V_{ref}^t, and the decision on fault occurrence is made if $\left|V_t - V_{ref}^t\right| \Big/ V_{ref}^t > d$, where d is the predetermined threshold which corresponds to the sensitivity of the deviation test. If the deviation does not exceed the threshold d, the window will move forward by L to get the new reference amplitude and the deviation test is repeated until the decision on fault occurrence can be made. In our simulation, the average delays are calculated for both $L = 50$ and $L = 200$.

Note that for comparisons of the delay performance among the methods above, the thresholds, i.e., h (GLLR), g (weighted CUSUM), m (SPRT) and d (deviation test) were adjusted to obtain the same false alarm period.

To compare with the machine leaning method, we implement the two-class support vector machine (SVM) with a Gaussian radial basis function kernel, and the semi-supervised learning approach in [11]. In specific, in the off-line phase, the machine learning model is trained by the data from the simulated normal output signals and the faulty signals which all include three attributes, i.e., the power, the voltage, and the duty cycle. In the on-line phase, the measurements with three attributes are sampled consecutively and fed into the model in sequence. The trained model will make a decision between normal and fault on each sample. The fault alarm will be triggered as soon as the classifier first identifies a sample as a faulty signal. The parameter c which controls the tolerance to the misclassification for SVM, and the parameter α which controls the solution rule in the kernel function for the semi-supervised learning are tuned to satisfy the target false alarm periods.

C. Faults due to the Irradiance Fluctuation

To simulate the irradiance fluctuation, we adjusted the input parameter of the PV array by reducing the solar irradiance

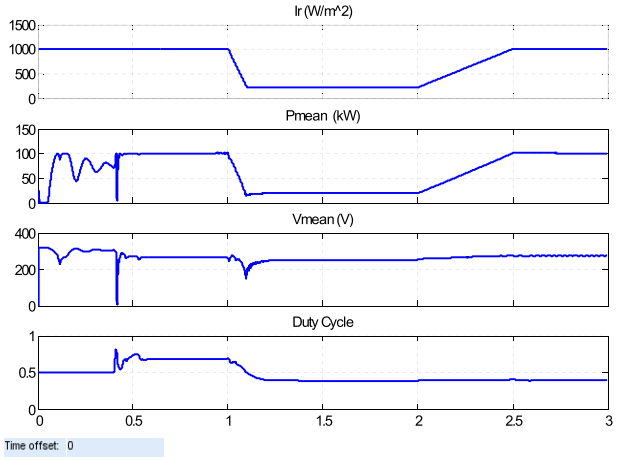
[1] In this work, the number of samples are used as the proxy metric for the decision latency on the account of the following assumptions. First, the sampling interval T_s (i.e., period of time between two sampling instants) is assumed to be uniform. Therefore, given the number of samples N, the real time decision latency can be readily computed as $N \times T_s$. In the numerical comparison, it is assumed that all the methods are implemented based on the same sampling interval, such that the number of samples can be used as a meaningful decision latency metric. Second, the algorithm running time at each sampling interval is neglected for all the mentioned methods. This is a reasonable assumption by noting that our proposed scheme admits closed-form implementation, and the data processing is almost immediate.

from 1000 W/m^2 to 250 W/m^2 at $t = 1$s and remaining steady for 0.9s.

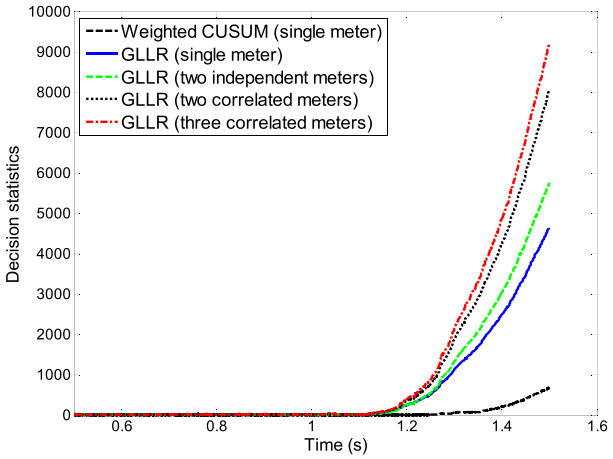
Fig. 4(a) shows the behaviors of the mean power and the mean voltage when the radiance begins to fluctuate at $t = 1$s with the environment temperature 25°C. As the input irradiance linearly decreases, the mean power drops down in a similar way while the mean voltage does not exhibit a dramatic decrease immediately until $t \approx 1.05$s and an abrupt change occurs at $t \approx 1.1$s. The duty cycle begins to decrease gently at $t = 1$s and remains steady after $t \approx 1.5$s.

In Fig. 4(b), the decision statistics of various methods are shown. For the single-meter case, the comparison between the proposed GLLR algorithm and the weighted CUSUM is presented when only the mean voltage is tracked. The GLLR algorithm exhibits a quick response to the linear change with short latency while the weighted CUSUM algorithm does not show an obvious reaction until $t = 1.4$s. For the multi-meter case, the method based on independent meters (where the signals observed by different meters are characterized by individual scalar AR models and the local statistics are aggregated to produce the total statistics) displays a better performance than the single-meter case with an increase before $t = 1.2$s. The proposed algorithm based on the vector AR model outperforms the others with a sharper response at $t = 1.1$s. The results above imply that the proposed algorithm demonstrates its prompt response to the output change.

Then we evaluate the average delay performance of different methods as a function of the false alarm period (in terms of the number of samples). The delay was the average of 1000 simulations. Fig. 5 shows the average delay performance for detecting the fault due to irradiance fluctuations. Given a specific false alarm period, the proposed correlated method has the shortest average delays in general and the differences are more obvious with larger false alarm periods. By employing multiple meters, the average delay is reduced. Overall, the deviation test is the least efficient method and its performance largely depends the window size.



(a) System output with fluctuating irradiance.



(b) Decision statistics for detecting faults due to irradiance fluctuation.

Fig. 4. Detection of faults due to irradiance fluctuation.

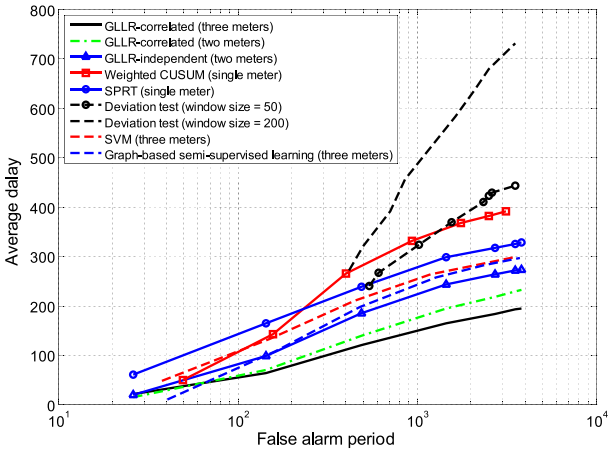


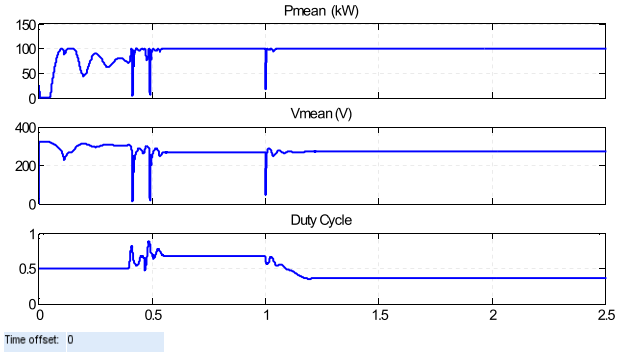
Fig. 5. Average detection delay for detecting faults due to irradiance fluctuation.

D. Line-line Faults

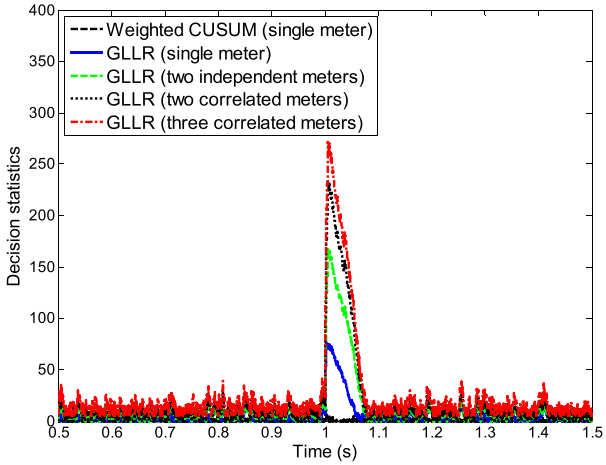
For the line-line fault and the ground fault below, we assumed standard test conditions (STC) of 1000 W/m^2 solar irradiance and 25°C PV module temperature. Since the line-line fault occurs when two points with different potentials are connected with low resistance, in the PV panel, it can be the connection between two PV modules in the same string or a cross connection among different strings/arrays. Here we consider the case that an accidental connection occurs between two PV arrays. In Fig. 3, we introduced a timed-breaker to simulate the line-line fault such that the positive output of the PV array above was accidentally connected to the negative output of the PV array below. The fault occurred when the breaker was closed at $t = 1$s with a breaker resistance $R_{on} = 1\Omega$ and a snubber resistance $R_s = \infty$.

Fig. 6(a) shows the system output. The fault results in a steep fall for both the mean power and the mean voltage which immediately return to the normal level after $t = 1$s with small fluctuations while the duty cycle gradually decreases and remains steady after $t \approx 1.2$s.

As shown in Fig. 6(b), the single-meter GLLR method captures the transient change in the mean power with a peak after $t = 1$s. However, the weighted CUSUM method does



(a) System output with line-line faults.



(b) Decision statistics for detecting line-line faults.

Fig. 6. Detection of line-line faults.

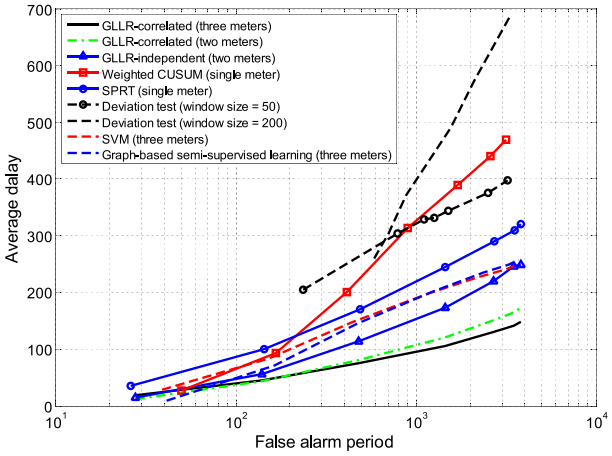
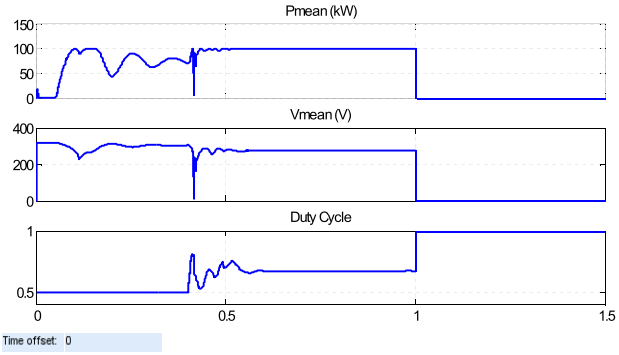
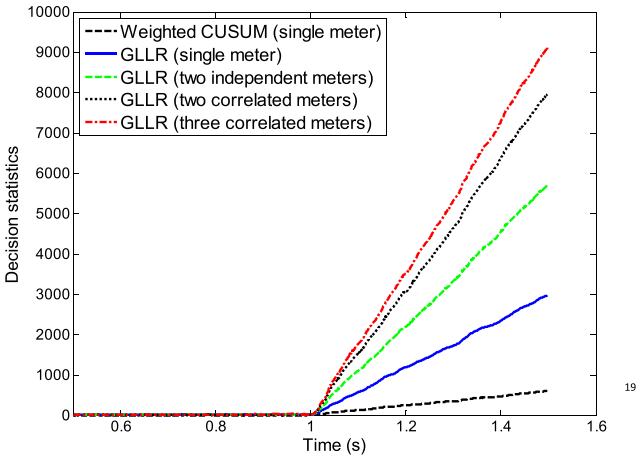


Fig. 7. Average detection delay for detecting line-line faults .



(a) System output with ground faults.



(b) Decision statistics for detecting ground faults.

Fig. 8. Detection of ground faults.

not show a large peak in the decision statistics, which implies it is not sensitive enough to capture the fault occurrence. For the multi-meter cases, the decision statistics of the correlated method exhibit an instant reaction to the abrupt disturbance. In contrast, the behaviors of other three methods show a smaller amount of increase in the decision statistics, which suggests that the correlated GLLR algorithm is more sensitive to the faults and makes prompt decisions.

The average detection delay performance for detecting the line-line fault is shown in Fig. 7. Similarly, the correlated GLLR method exhibits less detection delays. As the number of meters increases, the average delay is further reduced.

E. Ground Faults

The ground fault occurred as the breaker was closed and the output of PV panels was connected to the ground at $t = 1$s. The breaker resistance and the snubber resistance were set as $R_{on} = 10^{-4}\Omega$ and $R_s = 10^{12}\Omega$. The output behavior is shown in Fig. 8(a), where both the mean power and the mean voltage drop to zero at $t = 1$s and keep steady while the duty cycle increases to 1 at $t = 1$s. What differs from the previous two types of faults is that the outputs after the fault occurred do not change in a transient way or instantly return to normal, which allows for sufficient time for the change detection.

As we can see in Fig. 8(b), in response to the abrupt output drop at instant $t = 1$s, both the single-meter GLLR algorithm and the weighted CUSUM reacts promptly at $t = 1$s while the GLLR method exhibits a larger slope, indicating its quicker

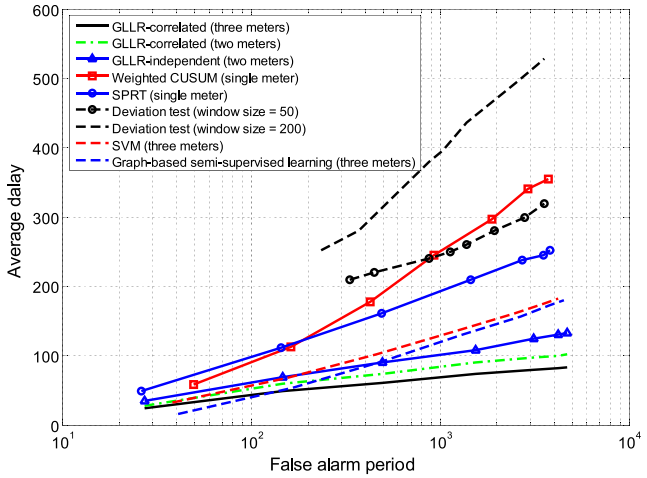


Fig. 9. Average detection delay for detecting ground faults.

response to the change. In the multi-meter case, the correlated GLLR method still outperforms the other methods in general.

Fig. 9 shows the average delay performance for detecting the ground fault. The proposed correlated GLLR method outperforms the other algorithms with less delays given the same false alarm period.

Based on the results above, the advantage of the proposed method lies in its high adaptivity, quick response and stable performance to detect different types of faults. In contrast, the SPRT and the deviation test require prior knowledge to determine the model parameters which is not always available in practice for various PV systems, and their performance can vary wildly with different parameters and various faults. Considering the unpredictable conditions in a real system, the proposed algorithm provides a solution with desired performance and wide applicability.

V. Conclusions

We have proposed a quickest detection approach to detecting unknown faults in PV systems that makes use of multiple online meters and exploits the correlation among them to enhance the detection performance. The framework is based on the vector AR modeling of the unknown faulty signals, and a generalized local likelihood ratio approach to handle the unknown parameters. The performance of the proposed detection method is simulated to detect typical faults in PV systems and compared with several existing methods. The results on decision statistics and average delays demonstrate its superior performance in making quick detection decisions upon different faults. Furthermore, its high adaptivity makes it suitable to be adopted in real PV systems.

Appendix

This section contains the derivations of z_i and $\boldsymbol{W}_i$ in (15).

A. Derivation of the First-Order Derivatives

Recall that $\boldsymbol{A} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p]$ and $\boldsymbol{Y}_{i-p}^{i-1} = [\boldsymbol{y}_{i-p}, \boldsymbol{y}_{i-p+1}, \ldots, \boldsymbol{y}_{i-1}]$. Using (11) and (8), we write the part of s_i that is a function of $\boldsymbol{\theta}_1$ as

$$\begin{aligned} s_{i,\boldsymbol{\theta}_1} &\triangleq -\tfrac{1}{2}\ln\left[\det\left(\boldsymbol{R}_{\boldsymbol{u}}\right)\right] \underbrace{-\tfrac{1}{2}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T \boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}} \\ &= -\tfrac{1}{2}\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \sum_{j=1}^{p}\boldsymbol{A}_j\boldsymbol{y}_{i-j}\Big]^T \boldsymbol{R}_{\boldsymbol{u}}^{-1}\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \sum_{j=1}^{p}\boldsymbol{A}_j\boldsymbol{y}_{i-j}\Big] \\ &= -\tfrac{1}{2}\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \boldsymbol{A}\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)\Big]^T \boldsymbol{R}_{\boldsymbol{u}}^{-1}\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \boldsymbol{A}\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)\Big] \\ &= -\tfrac{1}{2}\ln[\det\left(\boldsymbol{R}_{\boldsymbol{u}}\right)] \\ &\quad -\tfrac{1}{2}\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\mathrm{vec}(\boldsymbol{A})\Big]^T \boldsymbol{R}_{\boldsymbol{u}}^{-1} \\ &\quad \cdot\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\mathrm{vec}(\boldsymbol{A})\Big], \end{aligned} \tag{24}$$

where (24) follows from the fact that for matrices $\boldsymbol{A} \in \mathbb{R}^{k\times l}$ and $\boldsymbol{B} \in \mathbb{R}^{l\times m}$,

$$\mathrm{vec}\left(\boldsymbol{A}\boldsymbol{B}\right) = \left(\boldsymbol{B}^T \otimes \boldsymbol{I}_k\right)\mathrm{vec}\left(\boldsymbol{A}\right). \tag{25}$$

Hence we have

$$\begin{aligned} \boldsymbol{z}_i &= \tfrac{\partial s_i}{\partial \boldsymbol{\theta}_1}\big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} = \tfrac{\partial s_{i,\boldsymbol{\theta}_1}}{\partial \boldsymbol{\theta}_1}\big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0} \\ &= \Big[\tfrac{\partial s_i}{\partial \boldsymbol{\mu}}, \tfrac{\partial s_i}{\partial \mathrm{vec}(\boldsymbol{A})}, \tfrac{\partial s_i}{\partial \mathrm{vec}(\boldsymbol{U})}\Big]^T\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}. \end{aligned} \tag{26}$$

The three partial derivatives in (26) are computed as follows.

$$\begin{aligned} \tfrac{\partial s_i}{\partial \boldsymbol{\mu}} &= -\tfrac{1}{2}\partial\Big\{\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\mathrm{vec}(\boldsymbol{A})\Big]^T \boldsymbol{R}_{\boldsymbol{u}}^{-1} \\ &\quad \cdot\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\mathrm{vec}(\boldsymbol{A})\Big]\Big\}/\partial\boldsymbol{\mu} \\ &\overset{(41)}{=} -\tfrac{1}{2}\left\{2\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T \boldsymbol{R}_{\boldsymbol{u}}^{-1} \frac{\partial\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\mathrm{vec}(\boldsymbol{A})\Big]}{\partial \boldsymbol{\mu}^T}\right\}^T \\ &= \boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}. \end{aligned} \tag{27}$$

$$\begin{aligned} \tfrac{\partial s_i}{\partial \mathrm{vec}(\boldsymbol{A})} &= -\tfrac{1}{2}\partial\Big\{\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\mathrm{vec}(\boldsymbol{A})\Big]^T \boldsymbol{R}_{\boldsymbol{u}}^{-1} \\ &\quad \cdot\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\mathrm{vec}(\boldsymbol{A})\Big]\Big\}/\partial\mathrm{vec}\left(\boldsymbol{A}\right) \\ &\overset{(41)}{=} -\tfrac{1}{2}\left\{2 \times \boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T \boldsymbol{R}_{\boldsymbol{u}}^{-1} \frac{\partial\Big[\boldsymbol{y}_i - \boldsymbol{\mu} - \Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\mathrm{vec}(\boldsymbol{A})\Big]}{\partial \mathrm{vec}(\boldsymbol{A})^T}\right\}^T \\ &= \Big[\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T \boldsymbol{R}_{\boldsymbol{u}}^{-1}\Big(\mathrm{vec}\Big(\boldsymbol{Y}_{i-p}^{i-1}\Big)^T \otimes \boldsymbol{I}_K\Big)\Big]^T \\ &= \Big(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right) \otimes \boldsymbol{I}_K\Big)\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1} \\ &= \Big(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right) \otimes \boldsymbol{R}_{\boldsymbol{u}}^{-1}\Big)\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}. \end{aligned} \tag{28}$$

$$\begin{aligned} \frac{\partial s_i}{\partial \mathrm{vec}\left(\boldsymbol{U}\right)} &= \underbrace{\frac{\partial\left\{-\frac{1}{2}\log\det\left(\boldsymbol{U}\boldsymbol{U}^T\right)\right\}}{\partial \mathrm{vec}(\boldsymbol{U})}}_{\overset{(45)}{=} -\frac{1}{2}\times 2\mathrm{vec}\left(\boldsymbol{U}^{-1}\right)^T} \\ &\quad \underbrace{+\frac{\partial\left\{-\frac{1}{2}\left[\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\right]\right\}}{\partial \mathrm{vec}(\boldsymbol{U})}} \\ &\overset{(40)}{=} -\tfrac{1}{2}\times\frac{\partial \mathrm{vec}\left(\boldsymbol{U}\boldsymbol{U}^T\right)}{\partial \mathrm{vec}(\boldsymbol{U})}\cdot\frac{\partial\left[\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\right]}{\partial \mathrm{vec}\left(\boldsymbol{U}\boldsymbol{U}^T\right)} \\ &\overset{(43)}{=} -\tfrac{1}{2}\left[2\left(\boldsymbol{U}^T \otimes \boldsymbol{I}_K\right)\boldsymbol{N}_K\right]\Big[\mathrm{vec}\Big(-\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\Big)\Big] \\ &\overset{(44)}{=} -\tfrac{1}{2}\left[2\left(\boldsymbol{U}^T \otimes \boldsymbol{I}_K\right)\right]\Big[\mathrm{vec}\Big(-2\cdot\tfrac{1}{2}\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\Big)\Big] \\ &\overset{(25)}{=} \mathrm{vec}\Big(\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{U}\Big) \end{aligned} \tag{29}$$

Given (27)-(29), $\boldsymbol{z}_i$ is locally approximated as

$$\boldsymbol{z}_i = \left[\begin{array}{c} \boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1} \\ \left(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\right)\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1} \\ \mathrm{vec}\left(-\left(\boldsymbol{U}^{-1}\right)^T+\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{U}\right) \end{array}\right]\Bigg|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}$$
$$= \left[\begin{array}{c} \boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{y}_i \\ \left(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\boldsymbol{y}_i \\ \mathrm{vec}\left(-\left(\boldsymbol{V}^{-1}\right)^T+\boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{y}_i\boldsymbol{y}_i^T\boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{V}\right) \end{array}\right]. \quad (30)$$

B. Derivation of the Second-Order Derivatives

Based on (30), we calculate the second-order derivative in (17). First, the diagonal blocks are derived as follows.

$$\frac{\partial^2 s_i}{\partial\boldsymbol{\mu}\partial\boldsymbol{\mu}^T}=\frac{\partial\boldsymbol{R}_{\boldsymbol{u}}^{-1}\left[\boldsymbol{y}_i-\boldsymbol{\mu}-\left(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\otimes\boldsymbol{I}_K\right)\mathrm{vec}(\boldsymbol{A})\right]}{\partial\boldsymbol{\mu}^T}=-\boldsymbol{R}_{\boldsymbol{u}}^{-1}. \quad (31)$$

$$\begin{aligned}
&\frac{\partial s_i^2}{\partial\mathrm{vec}(\boldsymbol{A})\partial\mathrm{vec}(\boldsymbol{A})^T}=\frac{\partial\left(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\right)\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}}{\partial\mathrm{vec}(\boldsymbol{A})^T}\\
&=\partial\Big\{\left(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\right)\\
&\quad\cdot\left[\boldsymbol{y}_i-\boldsymbol{\mu}-\left(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\otimes\boldsymbol{I}_K\right)\mathrm{vec}(\boldsymbol{A})\right]\Big\}/\partial\mathrm{vec}(\boldsymbol{A})^T\\
&=-\left(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\right)\left(\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\otimes\boldsymbol{I}_K\right)\\
&=-\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}. \quad (32)
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial s_i^2}{\partial\mathrm{vec}(\boldsymbol{U})\partial\mathrm{vec}(\boldsymbol{U})^T}=\frac{\partial\mathrm{vec}\left(-\left(\boldsymbol{U}^{-1}\right)^T+\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{U}\right)}{\partial\mathrm{vec}(\boldsymbol{U})^T}\\
&\overset{(42)}{=}\underbrace{-\frac{\partial\boldsymbol{K}\mathrm{vec}\left(\boldsymbol{U}^{-1}\right)}{\partial\mathrm{vec}(\boldsymbol{U})^T}}\\
&\quad\overset{(46)}{=}-\boldsymbol{K}\left[-\left(\boldsymbol{U}^{-1}\right)^T\otimes\left(\boldsymbol{U}^{-1}\right)\right]\\
&\quad+\left[\boldsymbol{I}_K\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\right]\underbrace{\frac{\partial\mathrm{vec}\left(\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{U}\right)}{\partial\mathrm{vec}(\boldsymbol{U})^T}}\\
&\quad\overset{(42)}{=}\left[\boldsymbol{I}_K\otimes\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\right]\frac{\partial\mathrm{vec}(\boldsymbol{U})}{\partial\mathrm{vec}(\boldsymbol{U})^T}+\left[\boldsymbol{U}^T\otimes\boldsymbol{I}_K\right]\frac{\partial\mathrm{vec}\left(\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\right)}{\partial\mathrm{vec}(\boldsymbol{U})^T}\\
&\quad\overset{(47)}{=}\left[\boldsymbol{I}_K\otimes\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\right]+\left[\boldsymbol{U}^T\otimes\boldsymbol{I}_K\right]\left\{-2\boldsymbol{N}_K\left[\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{U}\otimes\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\right]\right\}\\
&\quad+\left[\boldsymbol{U}^T\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\otimes\boldsymbol{I}_K\right]\underbrace{\frac{\partial\mathrm{vec}\left(\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\right)}{\partial\mathrm{vec}(\boldsymbol{U})^T}}. \quad (33)\\
&\qquad\qquad\overset{(47)}{=}-2\boldsymbol{N}_K\left[\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{U}\otimes\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\right]
\end{aligned}$$

Hence we have

$$\mathbb{E}_{\boldsymbol{\theta}_0}\left\{-\frac{\partial s_i^2}{\partial\boldsymbol{\mu}\partial\boldsymbol{\mu}^T}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right\}=\mathbb{E}\left\{\boldsymbol{R}_{\boldsymbol{u}}^{-1}\big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right\}=\boldsymbol{R}_{\boldsymbol{v}}^{-1}, \quad (34)$$

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_0}\left\{-\frac{\partial s_i^2}{\partial\mathrm{vec}(\boldsymbol{A})\partial\mathrm{vec}(\boldsymbol{A})^T}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right\}\\
&=\mathbb{E}\left\{\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right\}\\
&=\mathbb{E}\left\{\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right\}\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\\
&=\boldsymbol{I}_p\otimes\boldsymbol{R}_{\boldsymbol{v}}\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}, \quad (35)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_0}\left\{-\frac{\partial s_i^2}{\partial\mathrm{vec}(\boldsymbol{U})\partial\mathrm{vec}(\boldsymbol{U})^T}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right\}\\
&=\mathbb{E}_{\boldsymbol{\theta}_0}\{-\boldsymbol{K}[\left(\boldsymbol{U}^{-1}\right)^T\otimes\left(\boldsymbol{U}^{-1}\right)]\}\\
&\quad-\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\boldsymbol{I}_K\otimes\left(\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\boldsymbol{R}_{\boldsymbol{u}}^{-1}\right)\right\}\\
&\quad+\mathbb{E}_{\boldsymbol{\theta}_0}\left\{2\left(\boldsymbol{U}^T\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\right)\left[\boldsymbol{N}_K\left(\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{U}\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\right)\right]\right\}\\
&\quad+\mathbb{E}_{\boldsymbol{\theta}_0}\left\{2\left[\boldsymbol{U}^T\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\otimes\boldsymbol{I}_K\right]\left[\boldsymbol{N}_K\left(\boldsymbol{R}_{\boldsymbol{u}}^{-1}\boldsymbol{U}\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\right)\right]\right\}\\
&=-\boldsymbol{K}[\left(\boldsymbol{V}^{-1}\right)^T\otimes\left(\boldsymbol{V}^{-1}\right)]-\boldsymbol{I}_K\otimes\left(\boldsymbol{R}_{\boldsymbol{v}}^{-1}\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\boldsymbol{y}_i\boldsymbol{y}_i^T\right\}\boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\\
&\quad+2\left(\boldsymbol{V}^T\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\boldsymbol{y}_i\boldsymbol{y}_i^T\right\}\right)\left[\boldsymbol{N}_K\left(\boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{V}\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\right]\\
&\quad+2\left[\boldsymbol{V}^T\boldsymbol{R}_{\boldsymbol{v}}^{-1}\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\boldsymbol{y}_i\boldsymbol{y}_i^T\right\}\otimes\boldsymbol{I}_K\right]\left[\boldsymbol{N}_K\left(\boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{V}\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\right]\\
&=-\boldsymbol{K}[\left(\boldsymbol{V}^{-1}\right)^T\otimes\left(\boldsymbol{V}^{-1}\right)]-\boldsymbol{I}_K\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\\
&\quad+4\left(\boldsymbol{V}^T\otimes\boldsymbol{I}_K\right)\left[\boldsymbol{N}_K\left(\boldsymbol{R}_{\boldsymbol{v}}^{-1}\boldsymbol{V}\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\right]. \quad (36)
\end{aligned}$$

Next, following the product rule in (42), we obtain the cross partial derivatives and their expectations as:

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\frac{\partial^2 s_i}{\partial\boldsymbol{\mu}\partial\mathrm{vec}(\boldsymbol{A})^T}\right\}=\mathbb{E}\left\{-\boldsymbol{R}_{\boldsymbol{u}}^{-1}\left[\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\otimes\boldsymbol{I}_K\right]\right\}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\\
&=-\boldsymbol{R}_{\boldsymbol{v}}^{-1}\underbrace{\left[\mathbb{E}\left\{\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\right\}\otimes\boldsymbol{I}_K\right]}_{\mathbf{0}}=\mathbf{0}, \quad (37)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\frac{\partial^2 s_i}{\partial\mathrm{vec}(\boldsymbol{U})\partial\boldsymbol{\mu}^T}\right\}\\
&=\mathbb{E}\Big\{\Big[\left(\boldsymbol{U}^T\boldsymbol{R}_{\boldsymbol{u}}^{-1}\right)\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\Big]\Big[\left(\boldsymbol{I}_K\otimes\boldsymbol{\varepsilon}_{i.\boldsymbol{\theta}_1}\right)\frac{\partial\mathrm{vec}\left(\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\right)}{\partial\boldsymbol{\mu}^T}\\
&\quad+\left(\boldsymbol{\varepsilon}_{i.\boldsymbol{\theta}_1}\otimes\boldsymbol{I}_K\right)\frac{\partial\mathrm{vec}\left(\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\right)}{\partial\boldsymbol{\mu}^T}\Big]\Big\}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\\
&=\Big[\left(\boldsymbol{V}^T\boldsymbol{R}_{\boldsymbol{v}}^{-1}\right)\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\Big]\Big[\Big(\boldsymbol{I}_K\otimes\underbrace{\mathbb{E}\left\{\boldsymbol{\varepsilon}_{i.\boldsymbol{\theta}_1}\right\}}_{\mathbf{0}}\Big)\underbrace{\frac{\partial\mathrm{vec}\left(\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}^T\right)}{\partial\boldsymbol{\mu}^T}}_{\text{constant}}\\
&\quad+\Big(\underbrace{\mathbb{E}\left\{\boldsymbol{\varepsilon}_{i.\boldsymbol{\theta}_1}\right\}}_{\mathbf{0}}\otimes\boldsymbol{I}_K\Big)\underbrace{\frac{\partial\mathrm{vec}\left(\boldsymbol{\varepsilon}_{i,\boldsymbol{\theta}_1}\right)}{\partial\boldsymbol{\mu}^T}}_{\text{constant}}\Big]\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}=\mathbf{0}, \quad (38)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_0}\left\{\frac{\partial^2 s_i}{\partial\mathrm{vec}(\boldsymbol{U})\partial\mathrm{vec}(\boldsymbol{A})^T}\right\}\\
&=\mathbb{E}\{-\Big[\Big(\boldsymbol{U}^T\boldsymbol{R}_{\boldsymbol{u}}^{-1}\Big)\otimes\boldsymbol{R}_{\boldsymbol{u}}^{-1}\Big]\Big[\left(\boldsymbol{I}_K\otimes\boldsymbol{\varepsilon}_{i.\boldsymbol{\theta}_1}\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\right)\frac{\partial\mathrm{vec}\left(\boldsymbol{A}^T\right)}{\partial\mathrm{vec}(\boldsymbol{A})^T}\\
&\quad+\left(\boldsymbol{\varepsilon}_{i.\boldsymbol{\theta}_1}\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\otimes\boldsymbol{I}_K\right)\Big]\Big\}\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}\\
&=-\Big[\Big(\boldsymbol{V}^T\boldsymbol{R}_{\boldsymbol{v}}^{-1}\Big)\otimes\boldsymbol{R}_{\boldsymbol{v}}^{-1}\Big]\Big[\Big(\boldsymbol{I}_K\otimes\underbrace{\mathbb{E}\left\{\boldsymbol{\varepsilon}_{i.\boldsymbol{\theta}_1}\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\right\}}_{\mathbf{0}}\Big)\frac{\partial\mathrm{vec}\left(\boldsymbol{A}^T\right)}{\partial\mathrm{vec}(\boldsymbol{A})^T}\\
&\quad+\Big(\underbrace{\mathbb{E}\left\{\boldsymbol{\varepsilon}_{i.\boldsymbol{\theta}_1}\mathrm{vec}\left(\boldsymbol{Y}_{i-p}^{i-1}\right)^T\right\}}_{\mathbf{0}}\otimes\boldsymbol{I}_K\Big)\Big]\Big|_{\boldsymbol{\theta}_1=\boldsymbol{\theta}_0}=\mathbf{0}. \quad (39)
\end{aligned}$$

Thus, using (34)-(39), we obtain $\mathbb{E}_{\boldsymbol{\theta}}\{\boldsymbol{W}_i\}$ in (19).

C. Formulas for Matrix Differentiation

1) Chain Rule for Vector Differentiation: Let $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ be $n\times 1$, $r\times 1$, and $m\times 1$ vectors, respectively. Suppose $\boldsymbol{z}$ is

a vector function of $\boldsymbol{y}$ and $\boldsymbol{y}$ is a vector function of $\boldsymbol{x}$. Then

$$\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{z}}{\partial \boldsymbol{y}}. \tag{40}$$

2) Product Rules for Vector Differentiation:

Suppose $\boldsymbol{\beta} \in \mathbb{R}^m$, $\boldsymbol{a}(\boldsymbol{\beta}) \in \mathbb{R}^n$, $\boldsymbol{c}(\boldsymbol{\beta}) \in \mathbb{R}^p$ and $\boldsymbol{A} \in \mathbb{R}^{n \times p}$ does not depend on $\boldsymbol{\beta}$. Then [37]

$$\frac{\partial \left[\boldsymbol{a}(\boldsymbol{\beta})^T \boldsymbol{A} \boldsymbol{c}(\boldsymbol{\beta})\right]}{\partial \boldsymbol{\beta}^T} = \boldsymbol{c}(\boldsymbol{\beta})^T \boldsymbol{A}^T \frac{\partial \boldsymbol{a}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T} + \boldsymbol{a}(\boldsymbol{\beta})^T \boldsymbol{A} \frac{\partial \boldsymbol{c}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T}. \tag{41}$$

Suppose $\boldsymbol{\beta} \in \mathbb{R}^m$, $\boldsymbol{A}(\boldsymbol{\beta}) \in \mathbb{R}^{n \times p}$, $\boldsymbol{D}(\boldsymbol{\beta}) \in \mathbb{R}^{q \times r}$ and $\boldsymbol{C} \in \mathbb{R}^{p \times q}$ does not depend on $\boldsymbol{\beta}$. Then [37],

$$\begin{aligned} \frac{\partial \mathrm{vec}(\boldsymbol{A}(\boldsymbol{\beta}) \boldsymbol{C} \boldsymbol{D}(\boldsymbol{\beta}))}{\partial \boldsymbol{\beta}^T} &= (\boldsymbol{I}_r \otimes \boldsymbol{A}(\boldsymbol{\beta}) \boldsymbol{C}) \frac{\partial \mathrm{vec}(\boldsymbol{D}(\boldsymbol{\beta}))}{\partial \boldsymbol{\beta}^T} \\ &\quad + \left(\boldsymbol{D}(\boldsymbol{\beta})^T \boldsymbol{C}^T \otimes \boldsymbol{I}_n\right) \frac{\partial \mathrm{vec}(\boldsymbol{A}(\boldsymbol{\beta}))}{\partial \boldsymbol{\beta}^T}. \end{aligned} \tag{42}$$

Applying the product rule, for an $(n \times r)$ matrix $\boldsymbol{X}$, we have [38]

$$\frac{\partial \mathrm{vec}(\boldsymbol{X}\boldsymbol{X}^T)}{\partial \mathrm{vec}(\boldsymbol{X})} = 2\left(\boldsymbol{X}^T \otimes \boldsymbol{I}_n\right) \boldsymbol{N}_n, \tag{43}$$

where $\boldsymbol{N}_n \in \mathbb{R}^{n^2 \times n^2}$ is associated with the commutation matrix $\boldsymbol{K}_n \in \mathbb{R}^{n^2 \times n^2}$ as $\boldsymbol{N}_n = \frac{1}{2}\left(\boldsymbol{I}_{n^2} + \boldsymbol{K}_n\right)$. For a square matrix $\boldsymbol{X} \in \mathbb{R}^{n \times n}$ [38],

$$\boldsymbol{N}_n \mathrm{vec}(\boldsymbol{X}) = \mathrm{vec}\left(\tfrac{1}{2}\left(\boldsymbol{X} + \boldsymbol{X}^T\right)\right). \tag{44}$$

3) Derivatives of a Logarithmic Determinant:

If $\det\left(\boldsymbol{X}\boldsymbol{X}^T\right)$ is positive so that $\ln\det\left(\boldsymbol{X}\boldsymbol{X}^T\right)$ exists, then [38]

$$\frac{\partial \ln\det(\boldsymbol{X}\boldsymbol{X}^T)}{\partial \mathrm{vec}(\boldsymbol{X})} = 2\mathrm{vec}\left(\left(\boldsymbol{X}\boldsymbol{X}^T\right)^{-1}\boldsymbol{X}\right). \tag{45}$$

4) Derivatives of an inverse matrix: For a nonsingular matrix $\boldsymbol{X} \in \mathbb{R}^{K \times K}$ [37],

$$\frac{\partial \mathrm{vec}(\boldsymbol{X}^{-1})}{\partial \mathrm{vec}(\boldsymbol{X})} = -\boldsymbol{X}^{-1} \otimes \left(\boldsymbol{X}^{-1}\right)^T. \tag{46}$$

Based on the rules (40), (43) and (46), for a $K \times K$ matrix $\boldsymbol{U}$, we have

$$\frac{\partial \mathrm{vec}\left(\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\right)}{\partial \mathrm{vec}(\boldsymbol{U})^T} = -2\boldsymbol{N}_K\left[\left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\boldsymbol{U} \otimes \left(\boldsymbol{U}\boldsymbol{U}^T\right)^{-1}\right]. \tag{47}$$

REFERENCES

[1] A. Chouder and S. Silvestre, "Automatic supervision and fault detection of PV systems based on power losses analysis," *Energy Conversion and Management*, vol. 51, no. 10, pp. 1929–1937, 2010.

[2] S. Firth, K. Lomas, and S. Rees, "A simple model of PV system performance and its use in fault detection," *Solar Energy*, vol. 84, no. 4, pp. 624–635, 2010.

[3] S. Silvestre, A. Chouder, and E. Karatepe, "Automatic fault detection in grid connected PV systems," *Solar Energy*, vol. 94, pp. 119–127, 2013.

[4] A. Drews, A. de Keizer, H. Beyer, E. Lorenz, J. Betcke, W. van Sark, W. Heydenreich, E. Wiemken, S. Stettler, P. Toggweiler, S. Bofinger, M. Schneider, G. Heilscher, and D. Heinemann, "Monitoring and remote failure detection of grid-connected PV systems based on satellite observations," *Solar Energy*, vol. 81, no. 4, pp. 548–564, 2007.

[5] M. Tadj, K. Benmouiza, A. Cheknane, and S. Silvestre, "Improving the performance of PV systems by faults detection using GISTEL approach," *Energy Conversion and Management*, vol. 80, pp. 298–304, 2014.

[6] H. Zhiqiang and G. Li, "Research and implementation of microcomputer online fault detection of solar array," in *2009 ICCSE 4th International Conference on Computer Science Education*, July 2009, pp. 1052–1055.

[7] S. Vergura, G. Acciani, V. Amoruso, and G. Patrono, "Inferential statistics for monitoring and fault forecasting of PV plants," in *2008 IEEE International Symposium on Industrial Electronics (ISIE)*, June 2008, pp. 2414–2419.

[8] S. Syafaruddin, E. Karatepe, and T. Hiyama, "Controlling of artificial neural network for fault diagnosis of photovoltaic array," in *2011 16th International Conference on Intelligent System Application to Power Systems (ISAP)*, Sept. 2011, pp. 1–6.

[9] C. C. K. Hong Chan Chang, Shang Chih Lin and H. P. Yu, "Cloud monitoring for solar plants with support vector machine based fault detection system," *Mathematical Problems in Engineering,*, vol. 2014, no. 564517, 2014.

[10] M. H. W. Kuei Hsiang Chao, Pi Yun Chen and C. T. Chen, "An intelligent fault detection method of a photovoltaic module array using wireless sensor networks," *International Journal of Distributed Sensor Networks,*, vol. 2014, no. 540147, 2014.

[11] Y. Zhao, R. Ball, J. Mosesian, J.-F. de Palma, and B. Lehman, "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2848–2858, May 2015.

[12] M. Davarifar, A. Rabhi, A. Hajjaji, and Z. Daneshifar, "Real-time diagnosis of PV system by using the sequential probability ratio test (SPRT)," in *2014 16th International Power Electronics and Motion Control Conference and Exposition (PEMC)*, Sept. 2014, pp. 508–513.

[13] E. Koutroulis and F. Blaabjerg, "A new technique for tracking the global maximum power point of PV arrays operating under partial-shading conditions," *IEEE Journal of Photovoltaics*, vol. 2, no. 2, pp. 184–190, April 2012.

[14] M. Seyedmahmoudian, R. Rahmani, S. Mekhilef, A. Maung Than Oo, A. Stojcevski, T. K. Soon, and A. Ghandhari, "Simulation and hardware implementation of new maximum power point tracking technique for partially shaded PV system using hybrid DEPSO method," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 850–862, July 2015.

[15] N. A. Ahmed and M. Miyatake, "A novel maximum power point tracking for photovoltaic applications under partially shaded insolation conditions," *Electric Power Systems Research*, vol. 78, no. 5, pp. 777–784, 2008.

[16] J. P. Dunlop, *Photovoltaic Systems*, 2nd ed. Amer Technical Pub., 2010.

[17] B. Brooks, *The Ground-Fault Protection BLIND SPOT: A Safety Concern for Larger Photovoltaic Systems in the United States*. A Solar ABCs White Paper, January 2012.

[18] M. Alam, F. Khan, J. Johnson, and J. Flicker, "PV ground-fault detection using spread spectrum time domain reflectometry (SSTDR)," in *2013 IEEE Energy Conversion Congress and Exposition (ECCE)*, Sept. 2013, pp. 1015–1020.

[19] Y. Zhao, J. de Palma, J. Mosesian, R. Lyons, and B. Lehman, "Line-line fault analysis and protection challenges in solar photovoltaic arrays," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3784–3795, Sept. 2013.

[20] X. Yao, L. Herrera, S. Ji, K. Zou, and J. Wang, "Characteristic study and time-domain discrete- wavelet-transform based hybrid detection of series DC arc faults," *IEEE Transactions on Power Electronics*, vol. 29, no. 6, pp. 3103–3115, June 2014.

[21] C. Strobl and P. Meckler, "Arc faults in photovoltaic systems," in *2010 Proceedings of the 56th IEEE Holm Conference on Electrical Contacts (HOLM)*, Oct. 2010, pp. 1–7.

[22] K. Lippert and T. Domitrovich, "AFCIs - from a standards perspective," in *2013 IEEE IAS Electrical Safety Workshop (ESW)*, March 2013, pp. 57–61.

[23] J. Johnson, B. Pahl, C. Luebke, T. Pier, T. Miller, J. Strauch, S. Kuszmaul, and W. Bower, "Photovoltaic DC arc fault detector testing at sandia national laboratories," in *2011 IEEE 37th Photovoltaic Specialists Conference (PVSC)*, June 2011, pp. 003 614–003 619.

[24] I. Gu, M. Bollen, and E. Styvaktakis, "The use of time-varying ar models for the characterization of voltage disturbances," in *2000 IEEE Power Engineering Society Winter Meeting,*, vol. 4, 2000, pp. 2943–2948.

[25] S. Li and X. Wang, "Cooperative change detection for voltage quality monitoring in smart grids," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 1, pp. 86–99, Jan. 2016.

[26] ——, "Monitoring disturbances in smart grids using distributed sequential change detection," in *2013 IEEE 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2013, pp. 432–435.

[27] D. Sidorov, D. Panasetsky, and V. SmadI, "Non-stationary autoregressive model for on-line detection of inter-area oscillations in power systems," in *2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, Oct. 2010, pp. 1–5.

[28] H. Endo and R. Randall, "Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 906–919, 2007.
[29] R. Andre-Obrecht, "A new statistical approach for the automatic segmentation of continuous speech signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, no. 1, pp. 29–40, Jan. 1988.
[30] T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction: with 200 full-color illustrations*. New York: Springer-Verlag, 2001.
[31] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.
[32] H. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge University Press, 2008.
[33] T. Banerjee, Y. C. Chen, A. D. Dominguez-Garcia, and V. V. Veeravalli, "Power system line outage detection and identification – a quickest change detection approach," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 3450–3454.
[34] Y. C. Chen, T. Banerjee, A. D. Dominguez-Garcia, and V. V. Veeravalli, "Quickest line outage detection and identification," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 749–758, Jan. 2016.
[35] S. Li, Y. Yilmaz, and X. Wang, "Quickest detection of false data injection attack in wide-area smart grids," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2725–2735, Nov. 2015.
[36] X. He, M.-O. Pun, C.-C. Kuo, and Y. Zhao, "A change-point detection approach to power quality monitoring in smart grids," in *2010 IEEE International Conference on Communications Workshops (ICC)*, May 2010, pp. 1–5.
[37] H. Ltkepohl, *New Introduction to Multiple Time Series Analysis*. Springer Publishing Company, Incorporated, 2007.
[38] D. A. Turkington, *Matrix Calculus and Zero-One Matrices*. Cambridge University Press, 2005.

Leian Chen received the B.E. degree in communication engineering from Fudan University, China, in 2013 and the M.S. degree in electrical engineering from Columbia University, New York, in 2015. She is currently working towards the Ph.D. degree in electrical engineering at Columbia University. Her current research interests broadly fall into statistical signal processing, and stochastic control with applications in sensor networks and smart grids.

Shang Li received the B.Sc. degree in electronics from Peking University, China, in 2010 and the M.Phil. degree in electronic and computer engineering (ECE) from the Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2012.

He is currently working towards the Ph.D. degree in electrical engineering at Columbia University, New York. His current research interests broadly fall into statistical signal processing, online inference over networks, and online decision-making with applications in social networks and smart grids.

Xiaodong Wang (S'98-M'98-SM'04-F'08) received the Ph.D. degree in electrical engineering from Princeton University. He is a Professor of electrical engineering at Columbia University in New York. Dr. Wang's research interests fall in the general areas of computing, signal processing, and communications, and he has published extensively in these areas. Among his publications is a recent book entitled Wireless Communication Systems: Advanced Techniques for Signal Reception, published by Prentice Hall in 2003.

His current research interests include wireless communications, statistical signal processing, and genomic signal processing. Dr. Wang received the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an Associate Editor for the *IEEE Transactions on Communications*, the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Signal Processing*, and the *IEEE Transactions on Information Theory*. He is a Fellow of the IEEE and listed as an ISI Highly-cited Author.