This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Distributed Grid State Estimation Under Cyber Attacks Using Optimal Filter and Bayesian Approach

Md Masud Rana, Rui Bo, and Ahmed Abdelhadi

Abstract—**Smart grid is built by combination of electric and information technologies and achieves the two-way interaction between power utilization and power generation. Unfortunately, new security threats appears together with cyber-physical communication systems. In order to properly monitor power network, an effective cyber attack detection and state estimation method are required to know attack and system states. This article considers the problem of robust grid state estimation and suggests a technique for distributed state estimation in power networks. First, the distribution power system incorporating multiple synchronous generators is modeled as a state-space framework, where attack occurs in measurements. Basically, the false data injection attacks can interfere with state estimation process by tampering with sensor measurements. Using mean squared error principle, the distributed dynamic state estimation algorithm is designed where local and neighboring gains are obtained using optimal filter and graph theory. For local gain computation, the attack parameter is obtained using the Bayesian learning process. The convergence condition of the proposed approach is derived. Extensive simulation results show that the proposed approach is able to estimate the system state within a short period of time. Hopefully, the proposed methodology can be used to tolerate the cyber attacks for improving the confidence of the grid state estimation process.**

Index Terms—**Bayesian approach, cyber attacks, distributed dynamic state estimation, false data injection attack (FDIA), graph theory, optimal filter.**

I. Introduction

THE conventional electric grid is undergoing a significant transformation in its power generation, transmission, and distribution units [1]. Interestingly, the use of advanced information and communication technology, sensors, and actuators is able to achieve these imperative milestones [2], [3]. Basically, the smart grid enables two-way communications between the utility operator and consumer, so it is more vulnerable to cyber attacks. Therefore, significant technical challenges arise for wide area monitoring, planning, and controlling the smart grid network [4]. To fulfill these challenges and meet customer satisfaction, the utility operator is monitored operational characteristics of power networks through a process called state estimation, which performs the task by filtering and fusing various sensory measurements. The attacker is manipulated sensory measurements that can mislead to the energy management system as shown Fig. 1. Generally speaking, the transmission of massive measurement to the centralized control center is expensive and infeasible, so the distributed estimation is gaining more popular. In distributed estimation, each agent in the power network is locally processed and exchanged information to recover system states [5]. Therefore, the distributed state estimation considering cyber attack is an important area of research, and this article deals with this emerging security issue.

Manuscript received December 9, 2019; revised June 2, 2020; accepted July 10, 2020. This work was supported by the National Science Foundation under Grant 1837472. *(Corresponding authors: Md Masud Rana; Rui Bo.)*

Md Masud Rana and Ahmed Abdelhadi are with the Engineering Technology Department, University of Houston,, TX77004 USA (e-mail: mrana928@yahoo.com).

Rui Bo is with the Electrical and Computer Engineering Department, Missouri University of Science and Technology, RollaMO 65409 USA.

Digital Object Identifier 10.1109/JSYST.2020.3010848

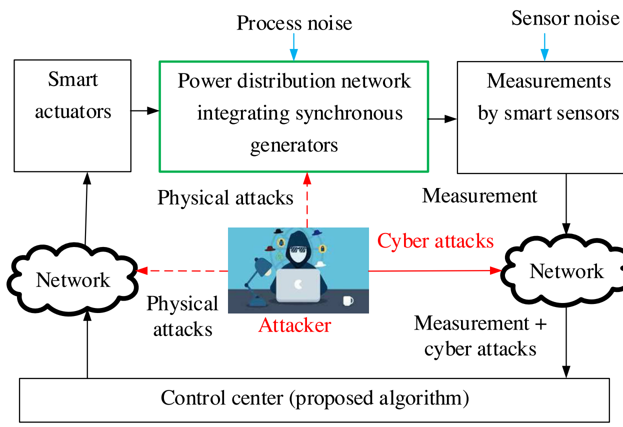

Fig. 1. Cyber-physical attacks for SCADA smart control center.

A. Related Work

In order to protect intrusion, the joint-transformation based false data injection attack (FDIA) scheme for smart grid is proposed in [6]. The idea is extended in [7], where H-infinity based attack-resilient algorithm is designed and verified. It can jointly estimate the system state and control input under the condition of cyber attacks [8]. Using same automatic generation control model, a dynamic watermarking framework is proposed to detect cyber attacks [9]. Furthermore, the semidefinite programming based AC state estimation scheme under cyber attack is presented in [10] and [11]. The computational complexity of this approach is very high. Moreover, the forecast-aided optimal state estimation algorithm is proposed in [12], where weighting sequence is obtained through the convex optimization process. When the system is highly dynamic, the algorithm cannot properly track the system behaviors. Moreover, the adaptive then combine distributed grid state estimation method under packet loss condition is proposed in [13]. In fusion center, the local estimation results are combined by designed weighting parameters, which are obtained through optimization process. For real-time implementation point of view, it is an unrealistic process as it involves two-stages (adaptive and combine) and requires more time to properly estimate grid states. Finally, computing the optimal weight parameters and convergence analysis are challenging tasks.

From filtering point of view, the Kalman filter (KF), extended KF (EKF), H-infinity EKF, unscented KF and cubature KF algorithms are used for power system state estimations [14], [15]. Moreover, the forecasted-aided KF algorithm considering cyber attack is explored in [16], where Euclidean distance metric is used to detect cyber attack. The observer-based anomaly detection scheme is presented in [17]. In addition, the wavelet transform-based mixed Kalman particle filter algorithm under FDIA is presented in [18]. The scenario based unsupervised learning algorithm for cyber physical power system is developed in [19]. All the aforementioned algorithms are designed for centralized state estimation, which requires all measurements and prone to vulnerable and single point failure. Due to deregulation of power systems, the distributed state estimation is gaining more attention in industrial and research communities.

In order to estimate the discrete time-varying cyber-physical system states, an iterative finite impulse response filter is designed [20]. It can effectively estimate the hidden system states without using any specific initialization scheme. For improvement of estimation accuracy, the robust type chandrasekhar-based maximum correntropy KF algorithm for cyber-physical system is proposed in [21]. The idea is extended in [22], where attack-resilient remote state estimation scheme is proposed and verified. The attackers are manipulated sensory measurements and the fusion center combines them for state estimations. Using residual prewhitening method, the cyber attack detection method is proposed in [23]. Technically, when the covariance matrix of the residual error is not full-rank, this method is used to solve the cyber attack detection and estimation problem.

Moreover, the distributed H-∞ algorithm for a virtual system is presented in [24]. The designed gains of the distributed filter are obtained by offline, and they are computed after solving several linear matrix inequalities (LMIs). The computational complexity of this approach is very high and cannot apply in real-time power systems. Moreover, the power system state estimation is jointly developed with an innovation-based cyber attack detection method to limit communication overhead in [25]. The alternating direction method of multipliers-based distributed state estimation algorithm for power system is presented in [26]. Moreover, the distributed optimal estimation algorithm for sensor networks are developed in [27]. The joint cyber attack detection and state estimation are presented in [28]. Furthermore, the distributed cyber attack detection and state estimation for a cyber-physical system under physical and cyber attacks are presented in [29]. After solving several LMIs and convex optimization, the joint estimation is performed, which incurs significant computational complexity.

Furthermore, the resource constraint based optimal state estimation algorithm for cyber-physical system is presented in [30]. Besides, the mixed integer linear programming based cyber attack protection scheme for power system is developed in [31]. The computational complexity is very high, and it requires significant amount of time as it is a bilevel optimization problem. In order to guarantee the cyber and operational security, a command authentication approach is proposed to detect intrusion [32], [33]. In addition, the mean squared error (MSE) based smart grid state estimation algorithm is presented in [34]. Due to the suboptimal nature of the computed gain, it cannot accurately estimate the grid states under cyber attacks [34].

Distributed state estimations face real environments, where cyber attacks and noisy measurements are present. Differentiated from prior literature, this article is the first of its kind to solve distributed state estimation problem for smart grid under cyber attacks using the optimal filter theory and Bayesian learning process.

B. Main Contributions

The main contributions of this article are as follows.

1) The power distribution system incorporating multiple synchronous generators is modeled as a state-space framework. The sensors are used to obtain measurement, which is corrupted by noise and cyber attacks.
2) Based on the optimal filter and graph theory, the distributed smart grid state estimation algorithm is proposed. Specifically, the local gain of the distributed scheme is obtained using the optimal filter theory, whereas the neighboring gain is determined through convex optimization process and graph theory. For local gain computation, the attack parameter is obtained using the Bayesian learning process.
3) The convergence condition of the proposed approach is derived.
4) Numerical simulations show that the proposed method can properly tolerate the cyber attack and noises leads to an accurate estimation result. The research issue and potential solution identified in this article can open up several avenues for future research in this area.

Organizations: The organization of this article is as follows: In Section II, power system model is described, which follows measurement and cyber attack frameworks in Section III. The proposed algorithm is derived in Section IV. Analysis of the computer simulations is given in Section V and Section VI concludes this article.

Notations: This article employs some standard notations. Bold face upper and lower case letters are used to represent matrices and vectors, respectively. Superscripts $\mathbf{x}'$ denotes the transpose of $\mathbf{x}$, $E(\cdot)$ denotes the expectation operator, diag($\mathbf{x}$) denotes the diagonal matrix, $\lambda(\mathbf{X})$ is the spectral radius of $\mathbf{X}$, and $\mathbf{I}$ denotes the identity matrix with appreciate dimension.

II. State-Space Representation of Power Network Incorporating Synchronous Generators

The smart grid is the backbone of a nation economy and is crucial to the homeland security. Generally, there are many synchronous generators and loads are connected to the distribution power networks. To illustrate, Fig. 2 shows the typical synchronous generators and loads that are connected to the 8-bus distribution lines [35]–[37]. Basically, the nth-synchronous

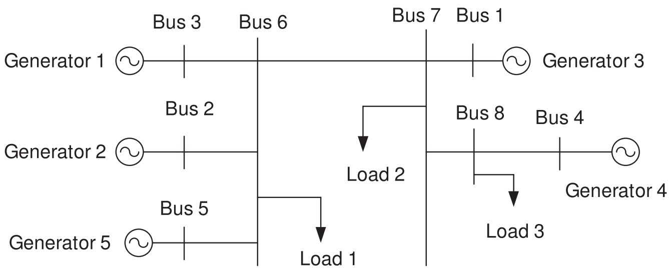

Fig. 2. Distributed power network incorporating synchronous generators.

generators can be represented by the following third-order differential equations as follows [36]– [38]:

$$\Delta\dot{\delta}_n = \Delta\omega_n \quad (1)$$

$$\Delta\dot{\omega}_n = -\frac{D_n}{H_n}\Delta\omega_n - \frac{\Delta P_{en}}{H_n} \quad (2)$$

$$\Delta\dot{E}'_{qn} = -\frac{\Delta E'_{qn}}{T'_{\text{don}}} + \frac{\Delta E_{fn}}{T'_{\text{don}}} + \frac{X_{dn}}{T'_{\text{don}}}\Delta I_{dn} - \frac{X'_{dn}}{T'_{\text{don}}}\Delta I_{dn} \quad (3)$$

here, δ_n is the rotor angle, ω_n is the rotor speed, H_n is the inertia constant, D_n is the damping constant, P_e is the active power delivered at the terminal, E'_{qn} is the quadrature-axis transient voltage, E_{fn} is the exciter output voltage, T'_{don} is the direct-axis open-circuit transient time constant, X_{dn} is the direct-axis synchronous reactance, X'_{dn} is the direct-axis transient reactance, and I_{dn} is the direct-axis current [35].

Generally, an automatic voltage regulator (AVR) is used to control the excitation current that leads to control the terminal voltage [36], [39]. A second-order transfer function is used to represent the AVR as follows [36]:

$$\Delta E_{fn} = b_{0n}z_{1n} + b_{1n}z_{2n} \quad (4)$$

$$\dot{z_{1n}} = z_{2n} \quad (5)$$

$$\dot{z_{2n}} = -c_{1n}z_{2n} - c_{0n}z_{1n} + \Delta v_n \quad (6)$$

here, z_{1n} and z_{2n} are the AVR internal states, b_{0n} and b_{1n} are transfer function coefficients of the AVR, c_{0n} and c_{1n} are the transfer function coefficients of the excitation system and Δv_n is the control input signal.

Considering N generators in the power network, the d-axis current I_{di} and electrical power P_{ei} are represented as [39]

$$I_{dn} = \sum_{m=1}^{N} \Delta E'_{qn}[B_{nm}\cos(\delta_n - \delta_m) - G_{nm}\sin(\delta_n - \delta_m)]. \quad (7)$$

$$P_{en} = \Delta E'_{qn}\sum_{m=1}^{N}[B_{nm}\sin(\delta_n - \delta_m) + G_{nm}\cos(\delta_n - \delta_m)]\Delta E'_{qm} \quad (8)$$

here, $n, m \in \{1, \ldots, N\}$, G_{nm} and B_{nm} are the real and imaginary part of the admittance $\mathbf{Y} \in \mathbb{R}^{N\times N}$, which is described in the Appendix A.

After linearizing (7) and (8), ΔP_{en} and ΔI_{dn} are written as follows [36], [40], [13]:

$$\Delta P_{en} = \left[\frac{\partial P_{en}}{\partial \delta}\ \frac{\partial P_{en}}{\partial E'_q}\right][\Delta\delta \Delta E'_q]' \quad (9)$$

$$\Delta I_{dn} = \left[\frac{\partial I_{dn}}{\partial \delta}\ \frac{\partial I_{dn}}{\partial E'_q}\right][\Delta\delta \Delta E'_q]' \quad (10)$$

here, $\Delta E'_q$ and $\Delta\delta$ are the transient voltage deviations and rotor angle deviations. By combining (1)–(6) and (9)–(10), it can be written as follows:

$$\dot{\mathbf{s}}_n = \mathbf{A}_n\mathbf{s}_n + \mathbf{B}_n u_n + \sum_{m\in N_n}\mathbf{A}_{nm}\mathbf{s}_m \quad (11)$$

here, the generator state $\mathbf{s}_n = [\Delta\delta_n \Delta\omega_n \Delta E'_{qn} z_{2n} z_{1n}]'$, the control input signal $u_n = \Delta v_n$, N_n indicates a set of connected generators, the system matrices $\mathbf{A}_n \in \mathbb{R}^{5\times5}$, $\mathbf{B}_n \in \mathbb{R}^{5\times1}$ and $\mathbf{A}_{nm} \in \mathbb{R}^{5\times5}$ are

$$\mathbf{A}_n = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -\frac{1}{H_n}\frac{\partial P_{en}}{\partial \delta n} & -\frac{D_n}{H_n} & -\frac{1}{H_n}\frac{\partial P_{en}}{\partial E'_{qn}} & 0 & 0 \\ X_n\frac{\partial I_{dn}}{\partial \delta_n} & 0 & -\frac{1}{T'_{don}} + X_n\frac{\partial I_{dn}}{\partial E'_{qn}} & \frac{b_{1n}}{T'_{don}} & \frac{b_{on}}{T'_{don}} \\ 0 & 0 & 0 & -c_{1n} & -c_{0n} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{A}_{nm} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{H_n}\frac{\partial P_{en}}{\partial \delta m} & 0 & -\frac{1}{H_n}\frac{\partial P_{en}}{\partial E'_{qm}} & 0 & 0 \\ X_n\frac{\partial I_{dn}}{\partial \delta_m} & 0 & -\frac{1}{T'_{don}} + X_n\frac{\partial I_{dn}}{\partial E'_{qm}} & \frac{b_{1n}}{T'_{dom}} & \frac{b_{on}}{T'_{dom}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\mathbf{B}_n = [0\ \ 0\ \ 0\ \ 1\ \ 0]'$[36] and $X_n = \dfrac{X_{dn} - X'_{dn}}{T'_{don}}$.

The aforementioned system can be written in continuous-time form

$$\dot{\mathbf{s}} = \mathbf{A}^c\mathbf{s} + \mathbf{B}^c\mathbf{u} + \mathbf{w} \quad (12)$$

here, $\mathbf{s} \in \mathbb{R}^{5N\times1}$, $\mathbf{u} \in \mathbb{R}^{N\times1}$, $\mathbf{w} \in \mathbb{R}^{5N\times1}$ is the process noise that can follow the Gaussian distribution incorporating zero mean and $\mathbf{Q}$ covariance, i.e., N($\mathbf{0}$, $\mathbf{Q}$), $\mathbf{A}^c \in \mathbb{R}^{5N\times5N}$ and $\mathbf{B}^c \in \mathbb{R}^{5N\times N}$ are given by:

$$\mathbf{A}^c = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_{12} & \cdots \mathbf{A}_{1N} \\ \mathbf{A}_{21} & \mathbf{A}_2 & \cdots \mathbf{A}_{2N} \\ \vdots & \vdots & \vdots \\ \mathbf{A}_{N1} & \mathbf{A}_{N2} & \cdots \mathbf{A}_N \end{bmatrix}$$

$$\mathbf{B}^c = \text{diag}(\mathbf{B}_1 \cdots \mathbf{B}_N).$$

Now, it can be written as a discrete-time form as follows:

$$\mathbf{s}(t+1) = \mathbf{A}\mathbf{s}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{w}(t) \quad (13)$$

where $\mathbf{A} = \mathbf{I} + \mathbf{A}^c\Delta t$, Δt is the sampling time, and $\mathbf{B} = \mathbf{B}^c\Delta t$.

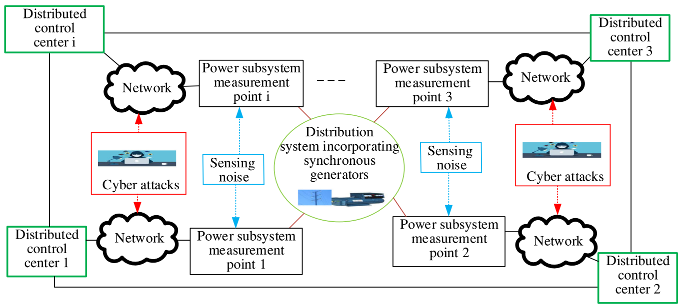

Fig. 3. Interconnected distribution power subsystems incorporating synchronous generators.

III. Measurement and Cyber Attack Frameworks

The distributed control centers are interconnected through communication links as shown in Fig. 3. In this figure, there are ith distribution subsystems, which are connected to the neighbors units. These control centers can share information with their neighbors in a distributed way. The sensors are installed into subsystem units to obtain distributed measurements. The sensing information is telemetered to the control centres to estimate system states such as rotor angle. The measurements are obtained as follows:

$$\mathbf{z}_i(t) = \mathbf{C}_i\mathbf{s}(t) + \mathbf{v}_i(t) \tag{14}$$

here, $\mathbf{z}_i(t) \in \mathbb{R}_i^p$ is the measurement, and $\mathbf{v}_i \backsim N(\mathbf{0}, \mathbf{R}_i)$ is the measurement noise, and $\mathbf{C}_i$ is the sensing matrix.

When sensing information is transmitted to the control center, the attacker can hack communication network and manipulate measurements. There are different kinds of attacks such as FDIA and replay attack [41]. For FDIA, an attacker is added intended information to the actual measurement over time, then report it to the control center for misleading. In latter case, the adversary records the normal measurements over time [9]. During attack, the actual measurements are replaced to be recoded one and thereby moving the system into an incorrect state [42]. The detailed attack templates are described in [43]. Mathematically, when there is attack then the system measurement can be written as follows:

$$\mathbf{z}_i^a(t) = \mathbf{C}_i\mathbf{s}(t) + \mathbf{v}_i(t) + \mathbf{a}_i(t) \tag{15}$$

where, $\mathbf{a}_i(t)$ is the cyber attack. We consider that the attack vector $\mathbf{a}_i$ is a Gaussian distribution with mean $\mu_\mathbf{i}$ and covariance $\bar{\mathbf{R}}_i^a$, i.e., $\mathbf{a}_i \backsim N(\mu_\mathbf{i}, \bar{\mathbf{R}}_i^a)$ [25], [28]. It assumes that the attack sequence is uncorrelated to each measurement [44], [43]. Let define the system model parameters $\varphi_i = (\mu_\mathbf{i}, \hat{\mathbf{R}}_i)$, where $\hat{\mathbf{R}}_i = \mathbf{R}_i + \mathbf{R}_i^a$ is the combined covariance of noise and cyber attack. Based on this noisy and corrupted version of measurements, the proposed state estimation algorithm is developed in the following section.

IV. Proposed Distributed Smart Grid State Estimation Algorithm and Convergence Condition

Using MSE principle, the distributed dynamic state estimation algorithm is designed where local and neighboring gains are obtained using optimal filter and graph theory. Afterward, the convergence condition of the developed approach is derived.

A. Proposed Algorithm

The proposed distributed state estimation algorithm is obtained using the optimal filter and Bayesian learning approaches [43]. Based on the interconnected structure in Fig. 3, the designed scheme is mathematically written as follows:

$$\hat{\mathbf{s}}_i(t+1) = \mathbf{A}\hat{\mathbf{s}}_i(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{G}_i(t)[\mathbf{z}_i^a(t) - \mathbf{C}_i\hat{\mathbf{s}}(t)] + \mathbf{L}_i(t)\sum_{j\in N_i}[\hat{\mathbf{s}}_j(t) - \hat{\mathbf{s}}_i(t)] \tag{16}$$

here, $\hat{\mathbf{s}}_i(t+1)$ is the posterior estimated system state, $\hat{\mathbf{s}}_i(t)$ is the previous estimated state, $\mathbf{G}_i(t)$ and $\mathbf{L}_i(t)$ are the local and consensus gains that can minimize the residual error dynamic, $\mathbf{z}_i^a(t) - \mathbf{C}_i\hat{\mathbf{s}}(t)$, and neighboring estimation mismatch, $\hat{\mathbf{s}}_j(t) - \hat{\mathbf{s}}_i(t)$, over time. Basically, the last term of the distributed scheme (16) is used for neighboring connections in Fig. 3, while the third term is included for self-estimation unit. The following theorem is used to compute these gains for distributed smart grid state estimation.

Theorem 1: After defining the error, $\boldsymbol{\eta}_i(t) = \mathbf{s}(t) - \hat{\mathbf{s}}_i(t)$, between the true and estimated system states and using the optimal filter as well as graph theory, the designed gains are obtained as follows [43]:

$$\mathbf{G}_i(t) = [\mathbf{A}\mathbf{P}_i(t)\mathbf{C}_i' + \mathbf{L}_i(t)\sum_{r\in N_i}\{\mathbf{P}^{ri}(t) - \mathbf{P}_i(t)\}\mathbf{C}_i'] \times [\mathbf{C}^i\mathbf{P}^i(k)\mathbf{C}'^i + \hat{\mathbf{R}}_i]^{-1}. \tag{17}$$

Using MSE principle, the estimation error covariance $\mathbf{P}_i(t+1) = E[\boldsymbol{\eta}_i(t+1)\boldsymbol{\eta}_i'(t+1)]$ is determined by

$$\mathbf{P}_i(t+1) = \mathbf{A}\mathbf{P}_i(t)\mathbf{A}' - \mathbf{A}\mathbf{P}_i(t)\mathbf{C}_i'[\mathbf{C}_i\mathbf{P}_i(t)\mathbf{C}_i' + \hat{\mathbf{R}}_i]^{-1}\mathbf{C}_i\mathbf{P}_i(t)\mathbf{A}' + \mathbf{Q} \tag{18}$$

here, $\mathbf{P}_i(t)$ is the prior estimation error covariance. Using the Bayesian learning formula, the covariance $\hat{\mathbf{R}}_i$ is computed as follows [45]:

$$\hat{\mathbf{R}}_i = (\alpha_i\bar{\mathbf{R}}_i + \rho_i[\mathrm{diag}(\hat{\boldsymbol{\mu}}_i)]^2 - (\rho_i+1)[\mathrm{diag}(\hat{\boldsymbol{\mu}}_i)]^2 + [\mathrm{diag}(\mathbf{z}_i^a - \mathbf{C}_i\hat{\mathbf{s}}_i)]^2)/(\alpha_i+1) \tag{19}$$

$$\hat{\boldsymbol{\mu}}_i = (\rho_i\bar{\boldsymbol{\mu}}_i + \mathbf{z}_i^a - \mathbf{C}_i\hat{\mathbf{s}}_i)/(\rho_i+1) \tag{20}$$

where, $\bar{\mathbf{R}}_i$ and $\bar{\boldsymbol{\mu}}_i^a$ are the initial values, α_i and ρ_i are the hyperparameters.

For mathematical simplicity, we assume that neighboring gain $\mathbf{L}_i(t) = \upsilon\mathbf{I}$, where υ is the designed gain coefficient. Under a steady-state condition, it can be computed through the following convex optimization process [43]:

$$\upsilon = \underset{\upsilon}{\mathrm{argmax}}\begin{bmatrix}-\mathbf{I} & \boldsymbol{\Gamma}\\ \boldsymbol{\Gamma}' & -\mathbf{I}\end{bmatrix} < \mathbf{0} \tag{21}$$

here, $\boldsymbol{\Gamma} = \mathbf{I}\otimes\mathbf{A} - b\mathrm{diag}\{\mathbf{G}_i\mathbf{C}_i\} - b\mathrm{diag}\{\mathbf{L}_i\}(\mathbf{L}_p\otimes\mathbf{I})$ is the augmented state dynamic, $\mathbf{L}_p$ is the Laplacian operator which is obtained through the graph theory after combining all error dynamics in a compact form as shown in (32), $\otimes$ indicates the

Kronecker product. From (32), it can be seen that the overall system dynamic is written into two items: augmented state dynamic, $\mathbf{\Gamma}$, and augmented error, $\mathbf{W} = (\mathbf{1} \otimes \mathbf{I})\mathbf{w} - b\text{diag}\{\mathbf{G}_i\}\mathbf{v}$. The Proof is derived in Appendix B.

B. Convergence Condition

When the estimated grid state converges to the actual state over time then the convergence of the developed approach is guaranteed asymptotically [46]. In doing this, the designed local and neighbouring gains ($\mathbf{G}$ and $\mathbf{L}$) can play key rule to analyze the convergence of the algorithm. With these gains, the system is stable when the spectral radius of augmented state matrix is less than one, i. e., $\lambda(\mathbf{\Gamma}) < 1$. In this case, the term $E[\mathbf{W}(t)\mathbf{W}'(t)] = \tilde{\mathbf{Q}}_R(t)$ converges to

$$\tilde{\mathbf{Q}}_R(t) = \{\mathbf{I} \otimes \mathbf{G}_i(t)\}\text{diag}(\mathbf{R}_1, \ldots, \mathbf{R}_i)\{\mathbf{I} \otimes \mathbf{G}_i(t)\}' + \mathbf{1}\mathbf{1}'\mathbf{Q}. \tag{22}$$

The entry of vector $\mathbf{1}$ is one with appropriate dimension.

Lemma 1: If the eigenvalues of a matrix $\mathbf{H} \in \mathbb{R}^{r \times r}$ are increasing order, then the following inequality holds [47, p. 235], [46]:

$$\underset{\mathbf{b}\neq\mathbf{0}}{\text{minimum}} \left|\frac{\mathbf{b}'\mathbf{H}\mathbf{b}}{\mathbf{b}'\mathbf{b}}\right| \leq |\lambda_l(\mathbf{H})| \leq \underset{\mathbf{b}\neq\mathbf{0}}{\text{maximum}}\left|\frac{\mathbf{b}'\mathbf{H}\mathbf{b}}{\mathbf{b}'\mathbf{b}}\right|. \tag{23}$$

Here, $\lambda(\mathbf{H})$ denotes the eigenvalues of $\mathbf{H}$, $\lambda_l(\mathbf{H})$ is the spectral radius of $\mathbf{H}$ with $l = 1, 2, \ldots, r$ and $\mathbf{b} \neq \mathbf{0}$ is the nonzero vector. The smallest and largest eigenvalues of $\mathbf{H}$ can be characterized as solution of the min–max problems involving the *Rayleigh quotient* $\mathbf{b}'\mathbf{H}\mathbf{b}/\mathbf{b}'\mathbf{b}$. The Rayleigh quotient is used in the min–max theorem to get the actual values of all eigenvalues.

In order to apply the Lemma 1, one can ascending order the eigenvalues of $\mathbf{\Gamma}(t)$ as follows [46]:

$$\lambda_{\text{min}} = \lambda_1(\mathbf{\Gamma}(t)) \leq \lambda_2(\mathbf{\Gamma}(t)), \ldots, \leq \lambda_l(\mathbf{\Gamma}(t)) = \lambda_{\text{max}}. \tag{24}$$

Therefore, the maximum absolute eigenvalue of $\mathbf{\Gamma}(t)$ is

$$\rho(E[\mathbf{\Gamma}(t)]) \leq \underset{l}{\text{maximum}}|\lambda_l[\mathbf{I} \otimes \mathbf{A} - b\text{diag}\{\mathbf{G}_i(t)\mathbf{C}_i\} - b\text{diag}\{\mathbf{L}_i(t)\}(\mathbf{L}_p \otimes \mathbf{I})]|. \tag{25}$$

When the estimated states converge to the true states, the state error covariance $\mathbf{P}_i(t)$ converges to $\mathbf{P}_i$ [48].

V. Numerical Simulation Results and Analysis

To estimate the system state, the proposed algorithm is applied to the distribution power network as shown in Fig. 2. For simplicity, we assume that there are $i = 4$ interconnected distributed controllers as shown in Fig. 3. In this network, the step by step procedure of the developed algorithm is demonstrated in Fig. 4. First of all, the power system and measurement are modeled by (13) and (15), respectively. It can be seen that the sensing measurement is corrupted by noise and cyber attacks. For instance, the attacker is injected the malicious false data into the measurements based on the attack templates presented in Section III. Afterward, the distributed estimation structure (16) is formulated considering unknown filtering parameters such as local and consensus gains, which are obtained by (17) and (21).

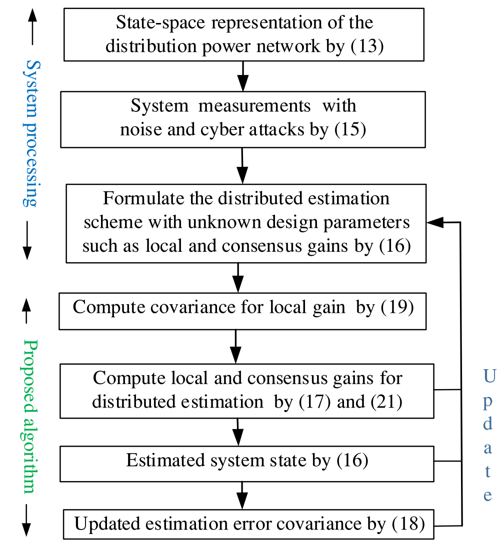

Fig. 4. Step-by-step procedure of the proposed algorithm.

TABLE I
Considered Generator Parameters

Parameters	G1	G2	G3	G4	G5
D_i	3.04	3.67	3.35	3.98	3.4
H_i	4.5	4.65	4.43	3.94	4.9
X'_{di}	0.0329	0.0329	0.3290	0.0239	0.2390
X_{di}	0.1016	0.1016	1.016	0.1016	1.016
T'_{doi}	5.57	5.57	5.57	5.57	5.57
b_{oi}	656	656	656	656	656
b_{1i}	1232	1232	1232	1232	1232
c_{oi}	3.23	3.23	3.23	3.23	3.23
c_{1i}	32.3	32.3	32.3	32.3	32.3
V	1.03	1.03	1.023	1.03	1.023
θ	0	0.1041	0.0933	0.0351	0.0607
Q	2.9141	1.3821	0.4187	2.1802	0.3469
P	3.1521	4.1016	0.4608	4.0578	0.1547

TABLE II
Considered Line Parameters

Node i	Node j	R_{ij}	X_{ij}	B_{ij}
1	7	0.00335	0.01057	0.01436
2	6	0.00313	0.00368	0.00304
3	6	0.03004	0.05242	0.06305
4	8	0.00514	0.01074	0.01654
5	6	0.00701	0.02231	0.02632
6	7	0.04022	0.12685	0.15848
7	8	0.01714	0.04143	0.05013

For local gain, the covariance is computed by (18). Technically, the local and consensus gains can minimize the residual error and neighboring estimation mismatch over time leads to an accurate estimation result. Finally, the estimated system state is updated by (16), and it error covariance is determined by (18).

The simulation is conducted through MATLAB and YALMIP environments [49]. The simulation parameters are described in Tables I and II [36], [39]. Basically, the process noise covariance is followed by Gaussian distribution with covariance is $10^{-8}\mathbf{I}$. In addition, the measurement noise covariance for four estimators are followed by Gaussian distributions with respective covariances are $1 * 10^{-7}\mathbf{I}$, $2 * 10^{-7}\mathbf{I}$, $3 * 10^{-7}\mathbf{I}$, and $4 * 10^{-7}\mathbf{I}$.

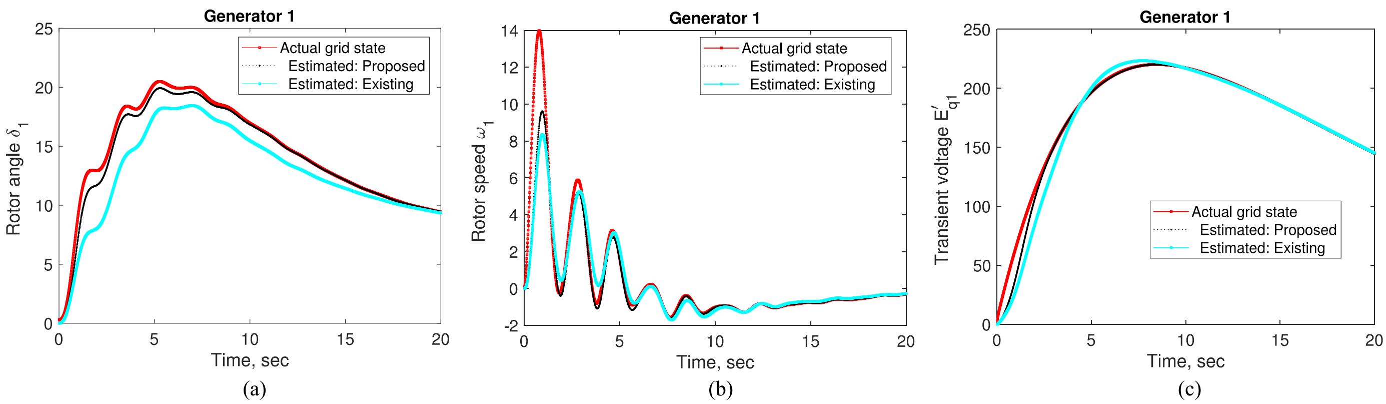

Fig. 5. Generator 1. (a) Actual rotor angle and its estimated state with FDIA. (b) Actual rotor speed and its estimated state with FDIA. (c) Actual transient voltage and its estimated state with FDIA.

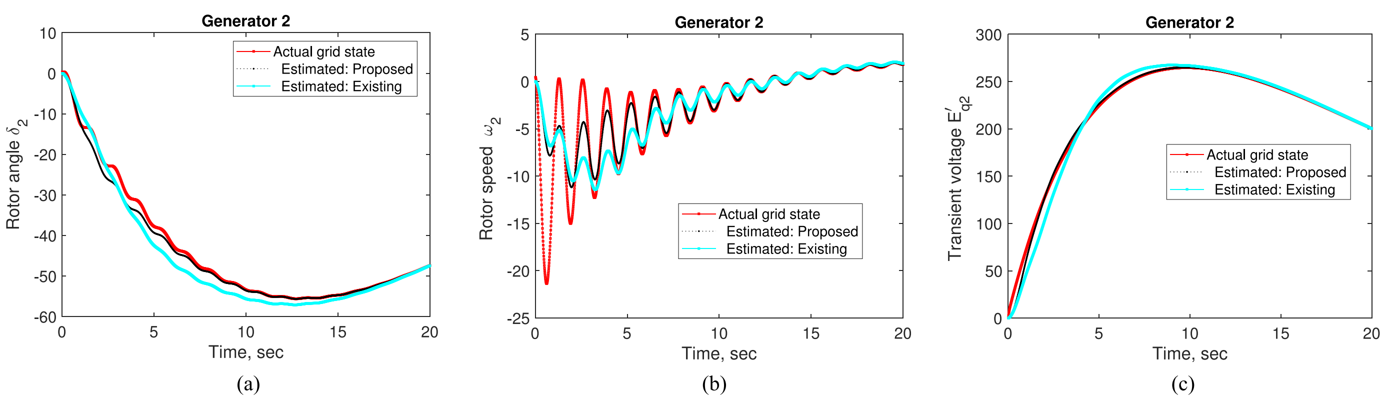

Fig. 6. Generator 2. (a) Actual rotor angle and its estimated state with FDIA. (b) Actual rotor speed and its estimated state with FDIA. (c) Actual transient voltage and its estimated state with FDIA.

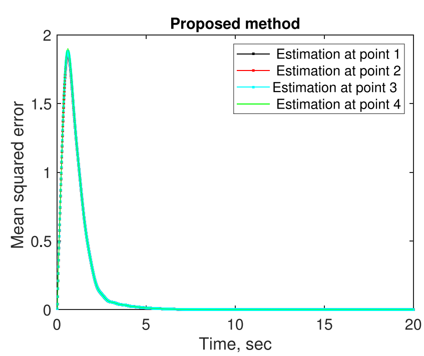

Fig. 7. MSE between true and estimated states with FDIA.

Moreover, the sampling period is 0.01 s, and there are five synchronous generators connected to the 8-bus distribution network as shown in Fig. 2. The simulation is conducted considering FDIA and replay attacks.

First of all, it assumes that the attacker is added FDIA into measurement during 0.05–0.25 s. In this case, the simulation results are illustrated in Figs. 5–8. Basically, Fig. 5(a)–(c) shows the generator 1 true states and their estimation results. The proposed algorithm can properly estimate the system states within 10 s, whereas the existing method requires 20 s [34]. Similarly, high estimation accuracy is illustrated in Fig. 6(a)–(c) for generator 2. This is due to the fact that the proposed algorithm can find the optimal gains, so the estimated states converge to the actual states within a short period of time. On the other hand, the existing method cannot determine the desired gains as they are suboptimal in nature. In this case [34], we assume the neighboring gain $\mathbf{L}$ is $0.026 \times \mathbf{I}$. The MSE between true and estimated system state is illustrated in Fig. 7. It can be seen that all estimators reach consensus on estimation. In summary, it can be observed that the proposed technique requires almost half time to estimate the system states compared with the exiting method.

When there is replay attack during 0.05–0.25 s, the simulation results are presented in Figs. 8–10. It can be seen that the proposed algorithm require more time to estimate the system states with this attack. Most importantly, the algorithm provides consistent results in all cases. The proposed algorithm can properly estimate the system states within 11 s while the existing method requires 20 s [34]. Fig. 8(a)–(c) shows the generator 4 true states and their estimation results. The developed scheme can accurately estimate the grid states within 11 s, whereas the existing method requires 20 s [34]. Similarly, high estimation accuracy is described in Fig. 9(a)–(c) for generator 5. This is

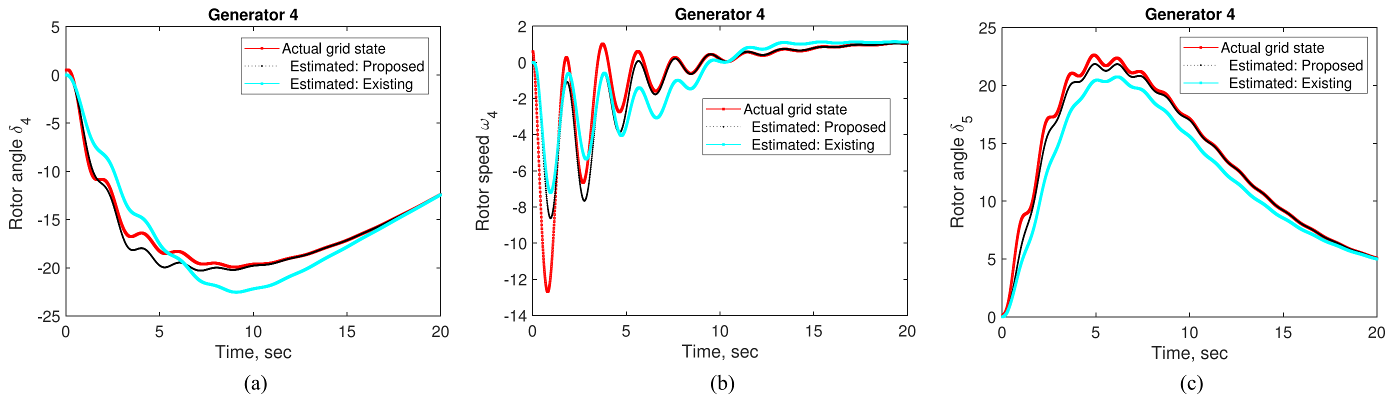

Fig. 8. Generator 4. (a) Actual rotor angle and its estimated state with replay attack. (b) Actual rotor speed and its estimated state with replay attack. (c) Actual transient voltage and its estimated state with replay attack.

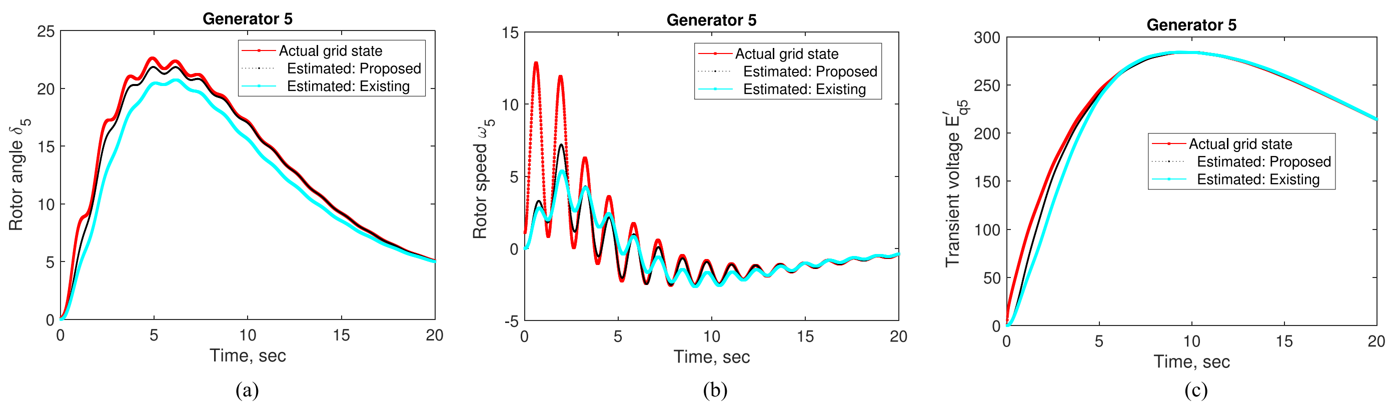

Fig. 9. Generator 5. (a) Actual rotor angle and its estimated state with replay attack. (b) Actual rotor speed and its estimated state with replay attack. (c) Actual transient voltage and its estimated state with replay attack.

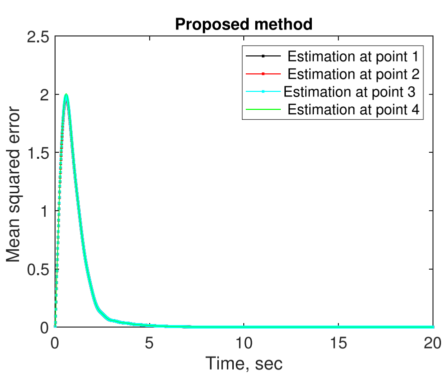

Fig. 10. MSE between true and estimated states with replay attack.

due to the fact that the designed scheme can find the optimal gains, so the estimated grid states converge to the true states within 11 s. However, the comparative scheme cannot compute the optimal gains as they are suboptimal in nature [34]. The MSE under this attack is illustrated in Fig. 10. It can be observed that all estimators reach consensus on estimation. Overall, it can be seen that the proposed method requires almost half time to estimate the grid states compared with the traditional approach.

VI. Conclusion

State estimation is the key task for power system operation and maintain stability as well as observability. However, the smart grid infrastructure is prone to cyber threats. In order to protect power network from cyber attacks, this article proposes a distributed state estimation algorithm. Specifically, we have made three main contributions to enhance the cyber security and resiliency of smart grids. First, the 8-bus distribution grid incorporating synchronous generators is modeled as a state-space framework, where measurement is obtained by a set of sensors. The measurement data are manipulated by cyber attacks such as FDIA. Second, we proposed an attack-resilient distributed state estimation algorithm based on the optimal filter and graph theory. Finally, the convergence condition of the proposed approach is derived. Extensive, simulation results show that the proposed algorithm can able to estimate system state within a short time. We will try to develop a data-driven distributed state estimation algorithm considering cyber attacks.

Appendix A

The network admittance is written by [39], [37]

$$\mathbf{Y} = \mathbf{Y}_{rr} - \mathbf{Y}_{re}\mathbf{Y}_{ee}^{-1}\mathbf{Y}'_{re} \qquad (26)$$

here, $\mathbf{Y}_{rr} = \text{diag}[Y_{17} + jB_{17}, Y_{26} + jB_{26}, Y_{36} + jB_{36}, Y_{46} + jB_{46}, Y_{56} + jB_{56}]$, where the mutual admittance is computed as follows as an example: $Y_{17} = 1/(R_{17} + jX_{17})$, R_{17} is the resistance between node 1 and 7, X_{17} and B_{17} are it's reactance and susceptance, respectively. The term $\mathbf{Y}_{re}$ is written as

$$\mathbf{Y}_{re} = \begin{bmatrix} 0 & -Y_{17} & 0 \\ -Y_{26} & 0 & 0 \\ -Y_{36} & 0 & 0 \\ 0 & 0 & -Y_{48} \\ -Y_{56} & 0 & 0 \end{bmatrix}. \tag{27}$$

The term $\mathbf{Y}_{ee}$ is written as

$$\mathbf{Y}_{ee} = \begin{bmatrix} Y_{66} & -Y_{67} & 0 \\ -Y_{67} & Y_{77} & -Y_{78} \\ 0 & -Y_{78} & Y_{88} \end{bmatrix} \tag{28}$$

where Y_{ii} is the self-admittance.

Appendix B (Proof of Theorem 1)

The estimation error $\boldsymbol{\eta}_i(t+1) = \mathbf{s}(t+1) - \hat{\mathbf{s}}_i(t+1)$ without first considering cyber attacks can be expressed as [43]

$$\boldsymbol{\eta}_i(t+1) = [\mathbf{A} - \mathbf{G}_i(t)\mathbf{C}_i]\boldsymbol{\eta}_i(t) + \mathbf{L}_i(t)\sum_{r\in N_i}[\boldsymbol{\eta}_r(t) - \boldsymbol{\eta}_i(t)] + \mathbf{w}(t) - \mathbf{G}_i(t)\mathbf{v}_i(t). \tag{29}$$

The estimation error covariance $\mathbf{P}_i(t+1) = E[\boldsymbol{\eta}_i(t+1)\boldsymbol{\eta}_i'(t+1)]$ is expressed as follows:

$$\begin{aligned}\mathbf{P}_i(t+1) &= [\mathbf{A} - \mathbf{G}_i(t)\mathbf{C}_i]\mathbf{P}_i(t)[\mathbf{A} - \mathbf{G}_i(t)\mathbf{C}_i]' \\ &+ [\mathbf{A} - \mathbf{G}_i(t)\mathbf{C}_i]\sum_{s\in N_i}[\mathbf{P}_{is}(t) - \mathbf{P}_i(t)]\mathbf{L}_i'(t) + \mathbf{L}_i(t)\sum_{r\in N_i} \\ &\times [\mathbf{P}_{ri}(t) - \mathbf{P}_i(t)][\mathbf{A} - \mathbf{G}_i(t)\mathbf{C}_i]' + \mathbf{L}_i(t)\sum_{r,s\in N_i} \\ &\times [\mathbf{P}_{rs}(t) - \mathbf{P}_{ri}(t) - \mathbf{P}_{is}(t) + \mathbf{P}_i(t)]\mathbf{L}_i'(t) \\ &+ \mathbf{G}_i(t)\mathbf{R}_i\mathbf{G}_i' + \mathbf{Q}.\end{aligned} \tag{30}$$

After partial derivative of $\mathbf{P}_i(t+1)$ with respect to local gain $\mathbf{G}_i(t)$, and setting it derivative equal to zero, i.e., $\frac{\partial\{tr[\mathbf{P}_i(t+1)]\}}{\partial \mathbf{G}_i(t)} = 0$, the optimal local gain is expressed as follows [43]:

$$\mathbf{G}_i(t) = [\mathbf{A}\mathbf{P}_i(t)\mathbf{C}_i' + \mathbf{L}_i(t)\sum_{r\in N_i}\{\mathbf{P}_{ri}(t) - \mathbf{P}_i(t)\}\mathbf{C}_i'] \times [\mathbf{C}_i\mathbf{P}_i(k)\mathbf{C}_i' + \mathbf{R}_i]^{-1}. \tag{31}$$

Using the heuristic approach, $\mathbf{R}_i$ in (31) is replaced by $\mathbf{R}_i + \mathbf{R}_i^a = \hat{\mathbf{R}}_i$, which leads to the local gain expression in (17).

Using (29), the overall error dynamic $\boldsymbol{\eta}(t) = [\boldsymbol{\eta}_1(t)\ldots\boldsymbol{\eta}_n(t)]'$ can be written as follows:

$$\begin{aligned}\boldsymbol{\eta}(t+1) &= [\mathbf{I}\otimes\mathbf{A} - b\text{diag}\{\mathbf{G}_i(t)\mathbf{C}_i\} - b\text{diag}\{\mathbf{L}_i(t)\} \\ &\quad \times (\mathbf{L}_p\otimes\mathbf{I})]\boldsymbol{\eta}(t) + (\mathbf{1}\otimes\mathbf{I})\mathbf{w} - b\text{diag}\{\mathbf{G}_i(t)\}\mathbf{v}(t) \\ &= \boldsymbol{\Gamma}(t)\boldsymbol{\eta}(t) + \mathbf{W}(t)\end{aligned} \tag{32}$$

where, $\mathbf{L}_p$ is the Laplican operator [13], $\boldsymbol{\Gamma}(t) = \mathbf{I}\otimes\mathbf{A} - b\text{diag}\{\mathbf{G}_i(t)\mathbf{C}_i\} - b\text{diag}\{\mathbf{L}_i(t)\}(\mathbf{L}_p\otimes\mathbf{I})$ and $\mathbf{W}(t) = (\mathbf{1}\otimes\mathbf{I})\mathbf{w} - b\text{diag}\{\mathbf{G}_i(t)\}\mathbf{v}(t)$. The overall error covariance $\mathbf{P}(t+1) = E[\boldsymbol{\eta}(t+1)\boldsymbol{\eta}(t+1)'(t+1)]$ can be expressed as follows:

$$\mathbf{P}(t+1) = E[\boldsymbol{\Gamma}(t)\mathbf{P}(t)\boldsymbol{\Gamma}'(k)] + E[\mathbf{V}(k)\mathbf{V}'(k)]. \tag{33}$$

According to the matrix property of the Kronecker product $vec(\mathbf{XYX}) = (\mathbf{X}'\otimes\mathbf{X})\text{vec}(\mathbf{Y})$, (33) can be written as a vector form

$$\begin{aligned}\text{vec}[\mathbf{P}(t+1)] &= E[\boldsymbol{\Gamma}(t)\otimes\boldsymbol{\Gamma}'(t)]\text{vec}[E\{\mathbf{P}(k)\}] \\ &\quad + \text{vec}[E\{\mathbf{V}(k)\mathbf{V}'(k)\}].\end{aligned} \tag{34}$$

It can be seen that $\mathbf{G}_i(t)$ and $\mathbf{L}_i(t)$ are unknown but function of each other that makes difficult to get any one of their estimated values. To overcome this, setting $\mathbf{L}^i(t) = \mathbf{0}$, the local gain $\mathbf{G}_i(t)$ and error covariance (18) is derived. For mathematical simplicity, we assume that neighboring gain $\mathbf{L}_i = \upsilon\mathbf{I}$, where υ is the designed gain coefficient. This is because, the consensus gain is generally small. With the given steady-state local gain and corresponding error covariance, the system is stable when the spectral radius of $\lambda[\boldsymbol{\Gamma}]$ is less than one [43], i.e.,

$$\upsilon = \underset{\upsilon}{\text{argmax}}\lambda(\boldsymbol{\Gamma}) < 1 \Rightarrow \boldsymbol{\Gamma}\boldsymbol{\Gamma}' < \mathbf{1} \Rightarrow \begin{bmatrix} -\mathbf{I} & \boldsymbol{\Gamma} \\ \boldsymbol{\Gamma}' & -\mathbf{I}\end{bmatrix} < \mathbf{0}. \tag{35}$$

The aforementioned LMI is obtained using the Schur complement.

Inspired by the Bayesian learning approach, the covariance $\hat{\mathbf{R}}_i$ is computed for local gain computation. Let define the unknown model parameter $\boldsymbol{\varphi}_i = (\boldsymbol{\mu}_i, \hat{\mathbf{R}}_i)$. For simplicity, we are omitted time index. Using the variational Bayesian learning method [45], the likelihood function is written as follows:

$$f(\mathbf{z}_i^a) = \int f(\mathbf{z}_i^a|\mathbf{s}, \boldsymbol{\varphi}_i)f(\mathbf{s})f(\boldsymbol{\varphi}_i)dsd\varphi_i. \tag{36}$$

Based on the state-space model (13), and measurement (15), $f(\mathbf{s}) = N(\mathbf{As} + \mathbf{Bu}, \mathbf{Q})$ and $f(\mathbf{z}_i^a|\mathbf{x}, \boldsymbol{\varphi}_i) = N(\mathbf{C}_i\mathbf{s} + \boldsymbol{\mu}_i, \hat{\mathbf{R}}_i)$. It assumes that the model parameters $(\boldsymbol{\mu}_i, \hat{\mathbf{R}}_i)$ have conjugate priors, so the posterior model parameter $\hat{\boldsymbol{\varphi}}_i$ is almost same as the prior probability distribution. Mathematically, it can be written as follows:

$$\boldsymbol{\mu}_i|(\boldsymbol{\mu}_i, \rho_i, \hat{\mathbf{R}}_i) \backsim N(\boldsymbol{\mu}_i, \hat{\mathbf{R}}_i/\rho_i) \tag{37}$$

$$\mathbf{R}_i^{-1}|(\alpha_i, \zeta_i) \backsim W(\alpha_i, \zeta_i^{-1}/\alpha_i) \tag{38}$$

here, $W(\bullet)$ is the Wishart distribution, μ_i, ρ_i, α_i, and ζ_i are the hyperparameters. It can be seen that all distribution functions are belong to the exponential family, and using [45], [50], we can obtain the unknown estimation parameter $\hat{\boldsymbol{\mu}}_i = E(\boldsymbol{\mu}_i)$ and $\hat{\mathbf{R}}_i = E(\mathbf{R}_i)$ as shown in (19) and (20). This completes the derivation.

Acknowledgment

The authors would like to thank Prof. Dr. Bruce McMillin, Department of Computer Science, Missouri University of Science and Technology, USA, for providing comments, supports, and resources for this article.

REFERENCES

[1] A. Ashok, M. Govindarasu, and J. Wang, "Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid," *Proc. IEEE*, vol. 105, no. 7, pp. 1389–1407, Jul. 2017.

[2] Y. Saleem, N. Crespi, M. H. Rehmani, and R. Copeland, "Internet of things-aided smart grid: Technologies, architectures, applications, prototypes, and future research directions," *IEEE Access*, vol. 7, pp. 62 962–63 003, 2019.

[3] M. P. Coutinho, G. Lambert-Torres, L. B. da Silva, H. Martins, H. Lazarek, and J. C. Neto, "Anomaly detection in power system control center critical infrastructures using rough classification algorithm," in *Proc. Int. Conf. Digit. Ecosyst. Technol.*, 2009, pp. 733–738.

[4] X. Yu and Y. Xue, "Smart grids: A cyber–physical systems perspective," *Proc. IEEE*, vol. 104, no. 5, pp. 1058–1070, May 2016.

[5] A. Tajer, S. Kar, H. V. Poor, and S. Cui, "Distributed joint cyber attack detection and state recovery in smart grids," in *Proc. Int. Conf. Smart Grid Commun.*, 2011, pp. 202–207.

[6] S. K. Singh, K. Khanna, R. Bose, B. K. Panigrahi, and A. Joshi, "Joint-transformation-based detection of false data injection attacks in smart grid," *IEEE Trans. Ind. Inf.*, vol. 14, no. 1, pp. 89–97, Jan. 2018.

[7] S. Z. Yong, "Simultaneous input and state set-valued observers with applications to attack-resilient estimation," in *Proc. Ann. Amer. Control. Conf.*, 2018, pp. 5167–5174.

[8] M. Khalaf, A. Youssef, and E. El-Saadany, "Joint detection and mitigation of false data injection attacks in AGC systems," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 4985–4995, Sep. 2019.

[9] T. Huang, B. Satchidanandan, P. Kumar, and L. Xie, "An online detection framework for cyber attacks on automatic generation control," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6816–6827, Nov. 2018.

[10] M. Jin, J. Lavaei, and K. H. Johansson, "Power grid AC-based state estimation: Vulnerability analysis against cyber attacks," *IEEE Trans. Autom. Control.*, vol. 64, no. 5, pp. 1784–1799, May 2019.

[11] K. Pan, A. Teixeira, M. Cvetkovic, and P. Palensky, "Cyber risk analysis of combined data attacks against power system state estimation," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3044–3056, May 2019.

[12] J. Chen, C. Dou, L. Xiao, and Z. Wang, "Fusion state estimation for power systems under DoS attacks: A switched system approach," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 49, no. 8, pp. 1679–1687, Aug. 2019.

[13] M. M. Rana, L. Li, and S. W. Su, "Distributed state estimation of smart grids with packet losses," *Asian J. Control*, vol. 19, no. 4, pp. 1306–1315, 2017.

[14] J. Zhao, L. Mili, and A. Abdelhadi, "Robust dynamic state estimator to outliers and cyber attacks," in *Proc. Power Energy. Soc. Gen. Meet.*, 2017, pp. 1–5.

[15] Y. Wang, Y. Sun, and V. Dinavahi, "Robust forecasting-aided state estimation for power system against uncertainties," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 691–702, Jan. 2020.

[16] H. Karimipour and V. Dinavahi, "Robust massively parallel dynamic state estimation of power systems against cyber-attack," *IEEE Access*, vol. 6, pp. 2984–2995, 2017.

[17] G. Anagnostou, F. Boem, S. Kuenzel, B. C. Pal, and T. Parisini, "Observer-based anomaly detection of synchronous generators for power systems monitoring," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4228–4237, Jul. 2018.

[18] B. Chen, H. Li, and B. Zhou, "Real-time identification of false data injection attacks: A novel dynamic-static parallel state estimation based mechanism," *IEEE Access*, vol. 7, pp. 95 812–95 824, 2019.

[19] S. Ahmed, Y. Lee, H. Seung-Ho, and I. Koo, "Unsupervised machine learning–Based detection of covert data integrity assault in smart grid networks utilizing isolation forest," *IEEE Trans. Inf. Forensics Secur.*, vol. 14, no. 10, pp. 2765–2777, Oct. 2019.

[20] S. Zhao, Y. S. Shmaliy, C. K. Ahn, and L. Luo, "An improved iterative FIR state estimator and its applications," *IEEE Trans. Ind. Inf.*, vol. 16, no. 2, pp. 1003–1012, Feb. 2020.

[21] M. V. Kulikova, "Chandrasekhar-based maximum correntropy Kalman filtering with the adaptive kernel size selection," *IEEE Trans. Autom. Control*, vol. 65, no. 2, pp. 741–748, Feb. 2020.

[22] A. Chattopadhyay and U. Mitra, "Security against false data injection attack in cyber-physical systems," *IEEE Trans. Cont. Net. Syst.*, vol. 7, no. 2, pp. 1015–1027, Jun. 2020.

[23] Q. Jiang, H. Chen, L. Xie, and K. Wang, "Real-time detection of false data injection attack using residual prewhitening in smart grid network," in *Proc. Int. Conf. Smart Grid Commun.*, 2017, pp. 83–88.

[24] V. Ugrinovskii, "Distributed h-∞ estimation resilient to biasing attacks," *IEEE Trans. Cont. Netw. Syst.*, vol. 7, no. 1, pp. 458–470, Mar. 2020.

[25] A. Minot, H. Sun, D. Nikovski, and J. Zhang, "Distributed estimation and detection of cyber-physical attacks in power systems," in *Proc. Int. Conf. Comm. World*, 2019, pp. 1–6.

[26] D. Du, X. Li, W. Li, R. Chen, M. Fei, and L. Wu, "ADMM-based distributed state estimation of smart grid under data deception and denial of service attacks," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 49, no. 8, pp. 1698–1711, Aug. 2019.

[27] A. Mustafa and H. Modares, "Analysis and detection of cyber-physical attacks in distributed sensor networks," in *Proc. Ann. Aller. Conf. Commun. Control. Comput.*, 2018, pp. 973–980.

[28] M. N. Kurt, Y. Yılmaz, and X. Wang, "Distributed quickest detection of cyber-attacks in smart grid," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 8, pp. 2015–2030, Aug. 2018.

[29] Y. Guan and X. Ge, "Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks," *IEEE Trans. Sig. Inf. Process. Over Netw.*, vol. 4, no. 1, pp. 48–59, Mar. 2018.

[30] F. Li and Y. Tang, "False data injection attack for cyber-physical systems with resource constraint," *IEEE Trans. Cybern.*, vol. 50, no. 2, pp. 729–738, Feb. 2020.

[31] X. Liu, Z. Li, and Z. Li, "Optimal protection strategy against false data injection attacks in power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1802–1810, Jul. 2017.

[32] S. Meliopoulos, G. Cokkinides, R. Fan, L. Sun, and B. Cui, "Command authentication via faster than real time simulation," in *Power Energy Soc. General Meet.*, 2016, pp. 1–5.

[33] U. A. Khan and A. M. Stanković, "Secure distributed estimation in cyber-physical systems," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, 2013, pp. 5209–5213.

[34] M. M. Rana, "Least mean square fourth based microgrid state estimation algorithm using the internet of things technology," *PLoS One*, vol. 12, no. 5, 2017, Art. no. e0176099.

[35] E. Ghahremani and I. Kamwa, "Online state estimation of a synchronous generator using unscented Kalman filter from phasor measurements units," *IEEE Trans. Enegry Convers.*, vol. 26, no. 4, pp. 1099–1108, Dec. 2011.

[36] J. Liu, A. Gusrialdi, S. Hirche, and A. Monti, "Joint controller-communication topology design for distributed wide-area damping control of power systems," *IFAC Proc. Vol.*, vol. 44, no. 1, pp. 519–525, 2011.

[37] M. M. Rana, L. Li, S. W. Su, and B. J. Choi, "Modelling the interconnected synchronous generators and its state estimations," *IEEE Access*, vol. 6, pp. 36 198–36 207, 2018.

[38] P. Kundur, N. J. Balu, and M. G. Lauby, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill, 1994, vol. 7.

[39] J. Machowski, J. Bialek, and J. Bumby, *Power System Dynamics: Stability and Control*. Hoboken, NJ, USAL: Wiley, 2011.

[40] A. Farraj, E. Hammad, and D. Kundur, "A distributed control paradigm for smart grid to address attacks on data integrity and availability," *IEEE Trans. Signal. Inf. Process. Over Netw.*, vol. 4, no. 1, pp. 70–81, Mar. 2018.

[41] S. Sridhar and M. Govindarasu, "Model-based attack detection and mitigation for automatic generation control," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 580–591, Mar. 2014.

[42] C. M. Ahmed, S. Adepu, and A. Mathur, "Limitations of state estimation based cyber attack detection schemes in industrial control systems," in *Proc. Smart City Secur. Privacy Workshop*, 2016, pp. 1–5.

[43] M. M. Rana, L. Li, S. W. Su, and W. Xiang, "Consensus-based smart grid state estimation algorithm," *IEEE Trans. Ind. Inf.*, vol. 14, no. 8, pp. 3368–3375, Aug. 2018.

[44] P. Ding, Y. Wang, G. Yan, and W. Li, "DoS attacks in electrical cyber-physical systems: A case study using truetime simulation tool," in *Proc. Chin. Automat. Congr.*, 2017, pp. 6392–6396.

[45] J. Bernardo *et al.*, "The variational Bayesian EM algorithm for incomplete data: With application to scoring graphical model structures," *Bayesian Statist.*, vol. 7, pp. 453–464, 2003.

[46] W. Yang, X. Wang, and H. Shi, "Optimal consensus-based distributed estimation with intermittent communication," *Int. J. Syst. Sci.*, vol. 42, no. 9, pp. 1521–1529, 2011.

[47] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[48] B. D. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Courier Corporation, 2012.

[49] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. Int. Conf. Robot. Autom.*, 2004, pp. 284–289.

[50] S. Zheng, T. Jiang, and J. S. Baras, "Robust state estimation under false data injection in distributed sensor networks," in *Proc. Global Telecomm. Conf.*, 2010, pp. 1–5.