2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Resilient nonlinear state estimation using the median operation for a network of droop-controlled power inverters

Anne van der Horst, Michelle S. Chong, Junsoo Kim, Henrik Sandberg

***Abstract*— We consider the problem of estimating the states of a continuous-time nonlinear dynamical system, where a subset of sensors can be maliciously corrupted using a potentially unbounded additive signal. The proposed estimation scheme employs a bank of observers which are robust with respect to disturbances and attacks, in conjunction with median operation to build the state estimate. The median operation is the key ingredient which guarantees that the state estimate is constructed using sensor(s) which are not under attack. A standing assumption in this scheme is that the system has to be observable from each sensor. We provide a constructive design method for the state observers for a class of nonlinear systems and illustrate the efficacy of the resilient median-based state estimation scheme using real data on an inverter-based power distribution network.**

I. Introduction

Cyber-physical systems (CPS) permeates multiple facets of the industry and our day-to-day life. Thus, the security of CPS has gained increasing interest over the last years [1]. Due to the integrated cyber- and physical dynamical nature of CPS, these systems are especially vulnerable to cyber-attacks ([2], [3]). Examples of CPS are critical infrastructure, like power grids [4]. Therefore, it is imperative that different mitigation strategies to minimize the impact of a possible attack and/or to reduce the likelihood of such attacks have been proposed, of which an overview is given by [5].

A popular point of attack has been the sensors and mitigation strategies have been widely studied. A common approach is to use the redundancy of sensing information. When less than half of the system's sensors are attacked, an accurate reconstruction of a system can be possible [6]. This has been shown for continuous LTI systems [7], discrete LTI systems [8], and for nonlinear systems [9], [10], to name a few.

We introduce a resilient state estimation method using a median operation for a class of nonlinear systems against malicious attacks on sensors. It is assumed that a nonlinear dynamical system has N sensors, of which at most M are compromised. The goal of the resilient state estimation method is to correctly estimate the true state of the system, given that it is unknown which sensors have been compromised. To do so, observers use a median operation to construct the correct state estimate. This method has been employed for linear [11] and linear distributed systems [12] in earlier works. However, it has not been used for nonlinear systems. Assuming that the plant is observable from each sensor, the median based method proposed for nonlinear systems has an advantage that it is computationally favorable compared to existing results employing multiple observers (c.f. the methods in [10] and [9]) as it employs only N observers, where N is the number of sensors present in the system.

In this paper, we employ a multiple observer architecture in constructing our state estimate, see Figure 1. Under the assumption that the system with N sensors is observable via each sensor, we design an observer using data from each sensor which is robust with respect to system disturbances. We perform a median operation in building our state estimate. We show that this design is constructive on Lur'e systems with slope-restricted and bounded nonlinearities. The robustness property imposed on each state observer in our framework is constructive, as existing observer designs such as in [13], [14], possess this property. Its design is formulated in the form of a linear matrix inequality. Finally, we apply the proposed median-based resilient state estimation algorithm to a model of an inverter-based power distribution network and verified our algorithm with benchmark data from a low voltage power distribution network of a residential zone.

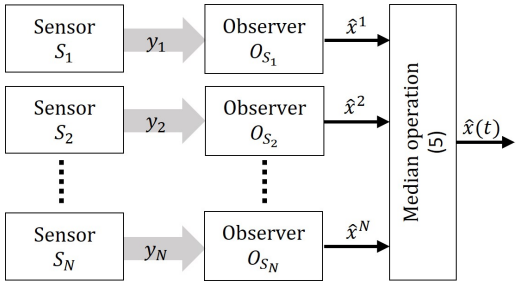

Fig. 1. Configuration of the proposed resilient state estimation.

The contributions of this paper are:

- A resilient state estimator using the median operation for nonlinear systems which are observable via every output. When strictly less than half of the sensors have been compromised, the estimation error is independent of the additive sensor attack.
- A constructive design for Lur'e-like systems.

A. Horst and M. Chong are with the Department of Mechanical Engineering, Eindhoven University of Technology. {a.v.d.horst1, m.s.t.chong}@tue.nl

J. Kim is with the Department of Electrical and Information Engineering, Seoul National University of Science and Technology, Korea. junsookim@seoultech.ac.kr

H. Sandberg is with the Division of Decision and Control, KTH Royal Institute of Technology, Stockholm, Sweden. hsan@kth.se

J. Kim and H. Sandberg were supported in part by the Swedish Energy Agency and ERA-Net Smart Energy Systems (grant agreement No 883973) and the Swedish Research Council (project 2016-00861).

- The efficacy of the algorithm is verified with benchmark residential data of a power distribution network.

The paper is structured as follows. In Section II, we formulate the problem and introduce the preliminaries used in building the resilient state estimator in Section III. We then show that our framework is constructive on a class of nonlinear systems in Section IV. The efficacy of the algorithm is then illustrated on a smart grid network in Section V and we validated the algorithm on real data in Section VI. Lastly, Section VII concludes the paper. All proofs are provided in the appendix.

Notation:

- Let $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_{\geq 0} = [0, \infty)$, $\mathbb{R}_{>0} = (0, \infty)$. The set of integers $\{i, i+1, i+2, ..., i+k\}$ is denoted by $\mathbb{N}_{[i,i+k]}$.
- Let (u, v), where $u \in \mathbb{R}^{n_u}$ and $v \in \mathbb{R}^{n_v}$, denote the column vector $(u^T, v^T)^T$.
- The cardinality of a set $\mathcal{J}$ is denoted as $|\mathcal{J}|$.
- The identity matrix of dimension n is denoted by $\mathbb{I}_n$ and a matrix of dimension m by n with all elements 1 is denoted by $\mathbf{1}_{m\times n}$.
- A diagonal matrix with elements e_i, $i \in \mathbb{N}_{[1,n]}$ is denoted by $\text{diag}(e_1, e_2, ..., e_n)$.
- Given a symmetric matrix P, its maximum (minimum) eigenvalue is denoted by $\lambda_{\max}(P)$ ($\lambda_{\min}(P)$).
- The infinity norm of a vector $x \in \mathbb{R}^n$, is denoted by $|x| := \max_{i\in\mathbb{N}_{[1,n]}} |x_i|$, and for a matrix $A \in \mathbb{R}^{n\times n}$, we define $|A| := \max_{i\in\mathbb{N}_{[1,n]}} \sum_{j\in\mathbb{N}_{[1,n]}} |a_{ij}|$, where a_{ij} is the i-th row and j-th column element of matrix A.
- A continuous function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a class $\mathcal{K}$ function, if it is strictly increasing and $\alpha(0) = 0$; additionally, if $\alpha(r) \to \infty$ as $r \to \infty$, then α is a class $\mathcal{K}_\infty$ function.
- A continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a class $\mathcal{KL}$ function, if: (i) $\beta(., s)$ is a class $\mathcal{K}$ function for each $s \geq 0$; (ii) $\beta(r, .)$ is non-increasing and (iii) $\beta(r, s) \to 0$ as $s \to \infty$ for each $r \geq 0$.

II. Problem formulation

We consider the following nonlinear system with N sensors in the presence of disturbances d and sensor attacks a_i.

$$\begin{aligned} \dot{x} &= f(x, z, w, d), \quad z = (z_1, z_2, ..., z_N), \\ z_i &= h_i(x, w, d), \qquad i \in \mathbb{N}_{[1,N]}, \\ y_i &= z_i + a_i, \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^{n_x}$ is the state, $y_i \in \mathbb{R}^{n_i}$ is the measured output at the i-th sensor, $w \in \mathbb{R}^{n_u}$ is a measured input, $d \in \mathbb{R}^{n_u}$ is the system disturbance, f and h_i are locally Lipschitz functions and $a_i : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_i}$ is a possibly unbounded attack signal that cannot be measured. We assume the following about the sensor attack model.

Assumption 1: Sensors $i \in \mathbb{N}_{[1,N]}$ which are not under attack satisfy $a_i(t) = 0$, for all $t \in \mathbb{R}_{\geq 0}$. The index set $\mathcal{I} \subseteq \mathbb{N}_{[1,N]}$ of attacked sensors is unknown and remains constant, i.e., the attack vector $a = (a_1, a_2, ..., a_N) \in \mathcal{N}_\mathcal{I}$, where $\mathcal{N}_\mathcal{I} := \{(a_1, a_2, ..., a_N) : a_i(t) = 0,\ \forall t \in \mathbb{R}_{\geq 0},\ \forall i \notin \mathcal{I}\}$. □

The objective of this paper is to estimate the states x of system (1) *under Assumption 1.*

We will build upon the multiple observer approach towards resilient state estimation proposed in [9]. The main contribution of this paper is in employing the median-based operation in building our state estimate from the estimates provided by multiple observers, which we will present in the next section.

III. Resilient median-based multi-observer

We first describe the median operation. The median of N values $y_1, y_2, ..., y_N$, denoted by $\text{med}(y_1, y_2, ..., y_N)$ is defined by the $((N+1)/2)^{\text{th}}$ largest value of $y_1, y_2, ..., y_N$ if N is odd, and defined by the average of the $(N/2)^{\text{th}}$ and the $(N/2+1)^{\text{th}}$ largest values of $y_1, y_2, ..., y_N$ if N is even. In the context of system (1), suppose there are N sensors measuring the same uncompromised sensor value y_0. Then, as long as $N > 2M$, where M is the number of compromised sensors, the median value of all estimates is equal to y_0, i.e.,

$$\text{med}(y_1, y_2, ..., y_N) = y_0. \tag{2}$$

This holds regardless of the values of a_i. This intriguing property of the median-based operation is used to choose an observer which receives attack-free sensor outputs. We assume that the system (1) is observable from each sensor as follows.

Assumption 2: There exists a function $\hat{f} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_i} \times \mathbb{R}^{n_i} \to \mathbb{R}^{n_x}$ such that the solution to

$$\dot{\hat{x}}^i = \hat{f}(\hat{x}^i, y_i, w), \qquad i \in \mathbb{N}_{[1,N]}, \tag{3}$$

and the solution to system (1), respectively satisfy

$$\begin{aligned} \left|x(t) - \hat{x}^i(t)\right| \leq & \hat{\beta}\left(\left|x(0) - \hat{x}^i(0)\right|, t\right) \\ & + \hat{\zeta}\left(\sup_{s\in[0,t)} |d(s)|\right) + \hat{\gamma}\left(\sup_{s\in[0,t)} |a_i(s)|\right), \end{aligned} \tag{4}$$

for all $t \in \mathbb{R}_{\geq 0}$ and initial conditions $x(0)$, $\hat{x}^i(0) \in \mathbb{R}^{n_x}$, where $\hat{\beta}$ is a $\mathcal{KL}$ function and $\hat{\zeta}$ and $\hat{\gamma}$ are $\mathcal{K}_\infty$ function. □

Assumption 2 requires the system (1) to be observable via each sensor y_i and condition (4) is an ISS (input-to-state stability) property of the estimation error system $x - \hat{x}^i$ with respect to the system disturbance d and the attack vector a_i. We employ the median operation to build a state estimate, which is defined as follows.

$$\begin{aligned} \hat{x} &= (\hat{x}_1, \hat{x}_2, ..., \hat{x}_{n_x}), \\ \hat{x}_j &= \text{med}(\hat{x}_j^1, \hat{x}_j^2, ..., \hat{x}_j^N), \qquad j \in \mathbb{N}_{[1,n_x]}. \end{aligned} \tag{5}$$

The multiple observer architecture using the median operation (5) is summarized in Fig. 1. With this framework, we guarantee that the state estimate converges to the true state up to a margin of error depending on the disturbance d, as follows.

Theorem 1: Consider the system (1) under Assumptions 1 and 2, with N outputs of which at most M are compromised,

i.e. the attack vector a belongs to $\mathcal{N}_{\mathcal{I}}$, for some set $\mathcal{I} \subset \mathbb{N}_{[1,N]}$ where $|\mathcal{I}| \leq M$. Suppose $N > 2M$, then there exist a class $\mathcal{KL}$ function β_m and a class $\mathcal{K}_\infty$ function ζ_m such that the solution to system (1) and the median-based observer (3), (5) satisfy

$$|x(t) - \hat{x}(t)| \leq \beta_m\left(|x(0) - \hat{x}(0)|, t\right) + \zeta_m\left(\sup_{s\in[0,t)} |d(s)|\right), \ \forall t \in \mathbb{R}_{\geq 0}, \quad (6)$$

for any initial conditions $x(0)$, $\hat{x}^i(0) \in \mathbb{R}^{n_x}$ and system disturbance d. □

Theorem 1 proves the existence of resilient observers. Next, we provide a constructive design methodology for a large class of systems, which include models of interconnected power inverters.

IV. Constructive design for Lur'e-like system

Consider the following Lur'e-like system, which is a specific form of (1):

$$\begin{aligned} &\dot{x} = Ax + \phi(z), \quad \phi(z) = (\phi_1(z_1), \phi_2(z_2), ..., \phi_N(z_N)), \\ &z_i = H_i x + w_i + d_i, \quad i \in \mathbb{N}_{[1,N]}, \\ &y_i = z_i + a_i, \end{aligned} \quad (7)$$

where the nonlinearities $\phi_i : \mathbb{R}^{n_i} \to \mathbb{R}$ are slope-restricted as follows.

Assumption 3: For $i \in \mathbb{N}_{[1,N]}$, the nonlinearity ϕ_i satisfies

$$\underline{e}_i \leq \frac{\phi_i(\xi) - \phi_i(\psi)}{\xi - \psi} \leq \bar{e}_i, \qquad \forall \xi, \psi \in \mathbb{R}, \quad \xi \neq \psi. \quad (8)$$

For this class of nonlinear systems (7), the condition of Assumption 2 is satisfied by designing each observer using each sensor y_i in the following manner:

$$\begin{aligned} \dot{\hat{x}}^i &= A\hat{x}^i + \phi(\xi^i) + L_i(y_i - (H_i\hat{x}^i + w_i)), \\ \xi^i &= H\hat{x}^i + w + K_i(y_i - (H_i\hat{x}^i + w_i)), \end{aligned} \quad (9)$$

where L_i and K_i are the observer matrices to be designed according to the sufficient condition in Proposition 1. Note that the terms $H\hat{x}^i$ and w use the full H from system (7) and all known inputs w, respectively.

Proposition 1: Consider the system (7) under Assumption 3 and observers (9). If for $N > 2M$ and, for every $i \in \mathbb{N}_{[1,N]}$, there exist a matrix $P_i = P_i^T > 0$, a positive definite matrix $U_i = \text{diag}(u_1, ..., u_N)$, scalars $\nu_i \geq 0$, $\mu_{i,d} \geq 0$, $\mu_{i,a} \geq 0$ and observer matrices L_i and K_i such that the following is satisfied

$$\begin{bmatrix} \mathcal{A}(P_i, P_iL_i, \nu_i) & \mathcal{B}\left(P_i, U_i, K_i^TU_i\right) & P_i & -P_i \\ \mathcal{B}\left(P_i, U_i, K_i^TU_i\right)^T & \mathcal{D}(U_i, \bar{e}) & 0 & 0 \\ P_i & 0 & \mathcal{M}(\mu_{i,d}) & 0 \\ -P_i & 0 & 0 & \mathcal{N}(\mu_{i,a}) \end{bmatrix} \leq 0 \quad (10)$$

where

$$\begin{aligned} \mathcal{A}(P_i, P_iL_i, \nu_i) &:= P_i(A - L_iH_i) + (A - L_iH_i)^TP_i + \nu_i\mathbb{I}_{n_x}, \\ \mathcal{B}\left(P_i, U_i, K_i^TU_i\right) &:= P_i + (H - K_iH_i)^TU_i, \\ \mathcal{D}(U_i, \bar{e}) &:= -2U_i\text{diag}(\bar{e}_1^{-1}, ..., \bar{e}_N^{-1}), \\ \mathcal{M}(\mu_{i,d}) &:= -\mu_{i,d}\mathbb{I}_{n_i}, \\ \mathcal{N}(\mu_{i,a}) &:= -\mu_{i,a}\mathbb{I}_{n_i}. \end{aligned}$$

Then the observation error $x - \hat{x}^i$ for every $i \in \mathbb{N}_{[1,N]}$ satisfies Assumption 2. □

The inequality in (10) is linear in P_i, P_iL_i, ν_i, U_i, $K_i^TU_i$, $\mu_{i,d}$ and $\mu_{i,a}$ and hence, can be solved numerically using MATLAB's LMI Lab, for example. To minimize the effect of the system disturbance d on the observation error $x - \hat{x}_i$ of each observer i, we aim to minimize the parameter $\mu_{i,d}$ subject to (10) for each $i \in \mathbb{N}_{[1,N]}$.

Remark 1: The LMI-based observer design presented in Proposition 1 follows the same ideas in [9] to obtain robustness with respect to system disturbances d, in addition to the attack signal a_i.

V. Resilient state estimation for an inverter-based power distribution network

We consider an inverter-based power distribution network in a line configuration as shown in Fig. 2. The network consists of N customers that each have an inverter, which is connected to the distribution network, and a smart secondary substation at the head of the line. The substation functions as a monitoring center, it communicates a desired nominal reference voltage $\bar{v} \in \mathbb{R}$ to each inverter containing the local controller Σ_i, which is able to generate reactive power $q_{g,i}$ while producing an active power $\rho_{g,i}$. As such, the voltages received by the customers v_i are regulated to operate in a safe margin, i.e. for a given $\delta > 0$, $\bar{v} - \delta \leq v_i(t) \leq \bar{v} + \delta$, for all $t \in \mathbb{R}_{\geq 0}$.

The voltage level at the connection point between the customer and the distribution line is v_i', with line impedances $Z_i' = R_i' + jX_i'$ between the customer i and the distribution line, and line impedances $Z_i = R_i + jX_i$ in between the connection points on the distribution line.

The power flow in the distribution network consists of P_i and Q_i which are the total active and reactive powers flowing from node i to node $i+1$, respectively; $\rho_i := \rho_{g,i} - \rho_{c,i}$ is the net injected active power into the distribution line from customer i; $q_i := q_{g,i} - q_{c,i}$, where $q_{g,i}$ is the net injected generated reactive power and $q_{c,i}$ the consumed reactive power from customer i, respectively. We model the power flow with a linearized DistFlow model [15].

The local controller Σ_i actuated by the inverter is able to generate reactive power $q_{g,i}$ at each customer i as follows

$$\dot{q}_{g,i} = -\frac{1}{\tau_i}q_{g,i} + \frac{1}{\tau_i}\Psi_i(\bar{v}^2 - v_i^2), \quad (11)$$

where $\tau_i > 0$ is the time constant of the inverter's response, $\bar{v} \in \mathbb{R}$ is the reference voltage and the droop function $\Psi_i : \mathbb{R} \to \mathbb{R}$ is a static mapping from the difference of the squared voltages w to the set-point for the reactive power. The droop function $\Psi_i(w)$ is chosen to be a piecewise

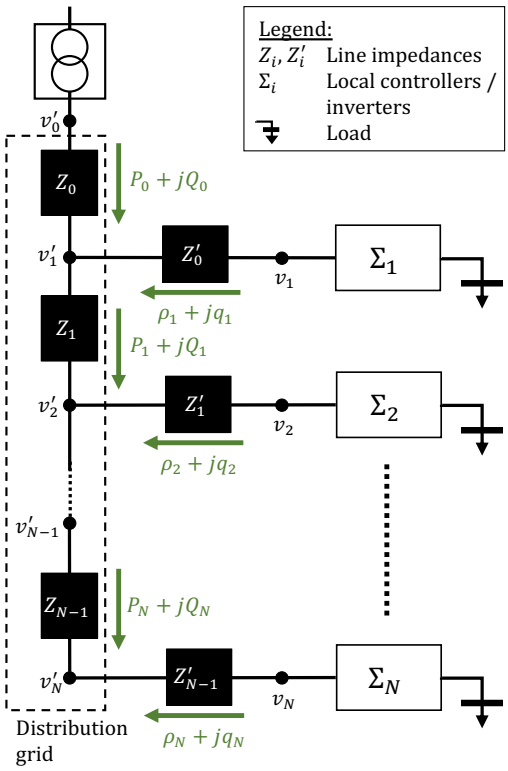

Fig. 2. Infrastructure of the low-voltage distribution grid.

saturation function considered in [16] of the following form:

$$\Psi_i(w) := \begin{cases} -\bar{Q}_i, & w \le w_{\min,i} \\ -\left(1 - \frac{w - w_{\min,i}}{w_{m,i} - w_{\min,i}}\right)\bar{Q}_i & w \in (w_{\min,i}, w_{m,i}] \\ 0, & w \in (w_{m,i}, w_{n,i}] \\ \left(\frac{w - w_n}{w_{\max,i} - w_{n,i}}\right)\bar{Q}_i & w \in (w_{n,i}, w_{\max,i}] \\ \bar{Q}_i, & w > w_{\max,i}, \end{cases} \tag{12}$$

where $w_{\min,i} \le w_{m,i} \le 0 \le w_{n,i} \le w_{\max,i}$ are design parameters, $\bar{Q}_i \in \mathbb{R}_{\ge 0}$ is the saturation limit of the i-th inverter, satisfying $\bar{Q}_i = \sqrt{\bar{s}_i^2 - \rho_{g,i}^2}$. Here, $\bar{s}_i \in \mathbb{R}$ is the maximum apparent power of the i-th inverter. The design parameters $w_{\min,i}$, $w_{m,i}$, $w_{n,i}$, $w_{\max,i}$ are chosen such that $\bar{e}_i := \min\left\{\frac{\bar{Q}_i}{w_{\max,i} - w_{n,i}}, \frac{\bar{Q}_i}{w_{m,i} - w_{\min,i}}\right\}$ satisfies [17, Theorem 6].

We consider the scenario where the measurements of the voltages v_i received at the substation may possibly be corrupted. This corruption is modelled using an additive attack signal $\alpha_i : \mathbb{R}_{\ge 0} \to \mathbb{R}$, which is possibly unbounded, as described by

$$y_i = v_i + \alpha_i. \tag{13}$$

We employ the same change in state coordinates as done by [17] such that the linearized DisFlow model from [15], controllers (11), and measurements received at the monitoring center (13) can be written in the form of (7) by choosing:

- $x = (q_{g,1}, q_{g,2}, ..., q_{g,N})$
- $z_i := \bar{v}^2 - v_i^2$
- the known input $w_i = \phi_i(\rho, q_c) + \bar{v}^2 - v'^2_0$, where

$$\begin{aligned} \phi_i(\rho, q_c) := & \sum_{j \in \mathbb{N}_{[0,i-1]}} \psi_j(\rho, q_c) \\ & + \sum_{j \in \mathbb{N}_{[0,i-2]}} 2\beta_j'(\rho_{j+1}, q_{c,j+1}), \\ \psi_j(\rho, q_c) := & \; 2X_j \sum_{k \in \mathbb{N}_{[j+1,N]}} q_{c,k} - 2R_j \sum_{k \in \mathbb{N}_{[j+1,N]}} \rho_k \\ & - 2\beta_j'(\rho_{j+1}, q_{c,j+1}), \end{aligned}$$

  where $\beta_i'(r,s) := R_i' r + X_i' s$, with $\beta_{-1}'(r,s) = 0$ for all $r, s \in \mathbb{R}$,
- the attack signal $a_i = 2v_i\alpha_i - \alpha_i^2$, where α_i is the additive attack signal from (13),
- H_i from (7) form the rows of the matrix

$$H = -2\begin{bmatrix} X_0 & X_0 & \dots & X_0 \\ X_0 & X_0 + X_1 & \dots & X_0 + X_1 \\ \vdots & \vdots & \ddots & \vdots \\ X_0 & X_0 + X_1 & \dots & \sum_{i \in \mathbb{N}_{[0,N-1]}} X_i \end{bmatrix} - 2\text{diag}(X_0', ..., X_{N-1}'),$$

- $A = \text{diag}(-1/\tau_1, -1/\tau_2, ..., -1/\tau_N)$, and
- $\phi_i(z_i) = \tau_i^{-1}\Psi_i(z_i)$, which satisfies Assumption 3, with $\underline{e}_i = 0$ and $\bar{e}_i = e_i/\tau_i$.

In order to estimate the voltages v_i at customer i, first the states x of all the controllers Σ_i, $i \in \mathbb{N}_{[1,N]}$ are estimated, after which the voltages v_i are estimated using

$$\hat{v}_i^2(t) = H_i\hat{x}(t) - \psi_i(\rho, q_c) + v'^2_0, \tag{14}$$

where $\hat{x}$ is the state estimate provided by the secure estimation algorithm, ψ_i and $v_0'^2$ are known inputs. We now have the following bound on the error between the estimated $\hat{v}_i$ and true voltages v_i.

Proposition 2: Consider the linearized Distflow model, controllers (11) with measurements (13), where α belongs to $\mathcal{N}_{\mathcal{I}}$ for some unknown set $\mathcal{I} \subset \mathbb{N}_{[1,N]}$ with at most M elements, received by the substation and the estimated voltages (14) at the substation. Suppose $N > 2M$ and for every $i \in \mathbb{N}_{[1,N]}$, there exist a matrix $P_i = P_i^T > 0$, a positive definite matrix $U_i = \text{diag}(u_1, ..., u_N)$, scalars $\nu_i \ge 0$, $\mu_{i,d} \ge 0$, $\mu_{i,a} \ge 0$ and observer matrices L_i and K_i such that (10) holds. Then, we have the following error bound on the resilient state estimator

$$\begin{aligned} |v_i(t)^2 - \hat{v}_i(t)^2| \le & \; \beta_v(|x(0) - \hat{x}(0)|, t) \\ & + \zeta_v\left(\sup_{s \in [0,t)} |d(s)|\right), \forall t \in \mathbb{R}_{\ge 0}, \end{aligned} \tag{15}$$

for all $i \in \mathbb{N}_{[1,N]}$, initial conditions $q_{g,i}(0) \in \mathbb{R}$, where $x(0) = (q_{g,1}(0), q_{g,2}(0), \dots, q_{g,N}(0))$, $\hat{x}(0) = (\hat{q}_{g,1}(0), \hat{q}_{g,2}(0), \dots, \hat{q}_{g,N}(0))$, β_v is a class $\mathcal{KL}$ function, and ζ_v is a class $\mathcal{K}_\infty$ function. □

In particular, Proposition 2 shows that if there are no disturbances, $d \equiv 0$, the estimates converge to the true values.

VI. Validation on a Benchmark Model

A. Load data

To validate the resilient state estimation method, a simulation of a benchmark residential European low-voltage

distribution network is performed [18, Table 7.26] with the model parameters in Table II. The residential subnetwork has $N = 5$ customers and is in a line configuration as shown in Fig. 2 with the mapping of the nodes in Table I.

TABLE I

MAPPING OF NODES FROM THE BENCHMARK TOPOLOGY TO OUR TOPOLOGY

Node label	
Benchmark topology [18, Fig. 7.7]	Our topology
$R1$	v_0'
$R3$	v_1'
$R11$	v_1
$R4$	v_2'
$R15$	v_2
$R6$	v_3'
$R16$	v_3
$R9$	v_4'
$R17$	v_4
$R10$	v_5'
$R18$	v_5

TABLE II

GRID PARAMETERS FROM TABLE 7.26 IN [18]

i	1	2	3	4	5
$R_{i-1}[\Omega]$	0.00343	0.00172	0.00343	0.00515	0.00172
$X_{i-1}[\Omega]$	0.04711	0.02356	0.04711	0.07067	0.02356
$R'_{i-1}[\Omega]$	0.00147	0.00662	0.00147	0.00147	0.00147
$X'_{i-1}[\Omega]$	0.02157	0.09707	0.02157	0.02157	0.02157
τ_i [s]	1	1	1	1	1
$\rho_{g,i}$ [W]	3500	5500	4000	4500	3000
$\rho_{c,i}$ [W]	2295	5440	5440	2295	2720
$q_{c,i}$ [VAr]	300	960	480	600	400

The setpoint voltage comminucated to each customer is $\bar{v} = 230$ V, and the nominal voltage at the head of the line $v_0(t) = 230 + 5\sin t$ V is a given function of time, to model harmonic perturbations. Due to the physical properties of power generation, the maximal reactive power generated satisfies the constraint $\bar{Q}_i = \sqrt{\bar{s}_i^2 - \rho_{g,i}^2}$, where $\bar{s}_i$ is a property of inverter i, such that we obtain

$$\begin{bmatrix}\bar{Q}_1 & \bar{Q}_2 & \bar{Q}_3 & \bar{Q}_4 & \bar{Q}_5\end{bmatrix} = \begin{bmatrix}2321.6 & 3464.1 & 2467.8 & 2800 & 1999\end{bmatrix} \text{VAr}$$

Considering the continuous droop function (12) and initializing the droop control law (11) at $q_{g,i} = 0$ VAr for $i \in \mathbb{N}_{[1,5]}$, then, according to [17, Theorem 6] and the model parameters in Table II, we obtain $\bar{R} = 0.0052\,\Omega$, $\bar{R}' = 0.0066\,\Omega$, $\bar{X} = 0.0707\,\Omega$, $\bar{X}' = 0.0971\,\Omega$, which leads to $\epsilon_y = 2325$ V^2, $\Delta_\rho = 2205$ W, and $\Delta_c = 627.04$ VAr. Further, we choose $\bar{e}_i = \bar{e}$ for i$\in \mathbb{N}_{[1,N]}$ where $\bar{e} := 0.2\,\frac{\text{A}}{\text{V}}$, which is achieved by setting $w_{n,i} = -w_{m,i} = 0$ V^2 and $w_{\max,i} = -w_{\min,i} = 17320.5$ V^2 for all $i \in \mathbb{N}_{[1,5]}$.

B. Conventional state estimation

The resilient state estimation algorithm proposed in this paper is compared to a conventional (*non-resilient*) state estimator as follows

$$\begin{aligned}\dot{\hat{x}} &= A\hat{x} + \phi(\xi) + L(y - (H\hat{x} + w)),\\ \xi &= H\hat{x} + w + K(y - (H\hat{x} + w)),\end{aligned} \tag{16}$$

where $w \in \mathbb{R}^{n_u}$ is a known input, and the observer matrices K and L are designed according to Proposition 1 with appropriate modification to the dimensions of the matrices. The main difference between a conventional *non-resilient* state estimator compared to the proposed *resilient* state estimation algorithm is in the number of sensor data that is being employed by the observer(s) in the respective algorithms. The conventional *non-resilient* state estimator employs the data from all sensors. This is in contrast to the *resilient* state estimation algorithm proposed in this paper, where each observer uses the data of only one sensor, and the algorithm then decides which sensors to use.

C. Results

To efficiently compare the performance of the conventional state estimator and the proposed resilient state estimator, the squared voltage estimation error $\tilde{v}_i^2(t) = v_i^2(t) - \hat{v}_i^2(t)$ is evaluated for both estimators. The results are shown in Fig. 3.

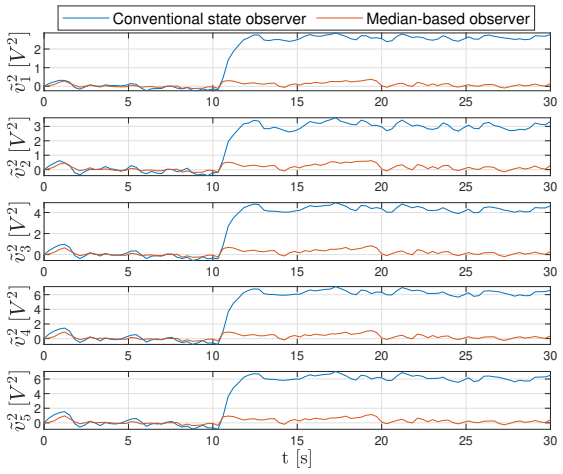

Fig. 3. The squared voltage estimation error $\tilde{v}_i^2(t) = v_i^2(t) - \hat{v}_i^2(t)$ for a conventional *non-resilient* state estimator and the proposed resilient state estimator with $a_1(t) = a_4(t) = 30$ V $\forall t \geq 10.5$ s and a random system disturbance with $|d(t)| \leq |d_{\max}| = 6$ V.

We launch sensor attacks on sensors 1 and 4 with $a_1(t) = a_4(t) = 30$ V, for all $t \geq 10.5$ s. Hence, we have $M = 2$, and $2M < N = 5$. We see that the resilient state estimation method outperforms the conventional state estimator. The squared voltage estimation error is small for the *resilient state estimator*, since the median operation excludes corrupted sensor data. The conventional *non-resilient* state estimator shows a jump in the squared voltage estimation error for $t \geq 10.5$ s as expected since it employs corrupted sensor data. It is interesting to note that all estimates show a jump even though only two sensors are corrupted. Hence, the non-resilient state estimator is sensitive to a few corrupted sensors.

VII. Conclusion and future work

We have proposed a resilient state estimation for continuous nonlinear systems with N outputs under adversarial sensor attacks using a multiple observer setup in lieu with a median operation to construct the state estimates. We require the system to be observable via each sensor (output) such that a robust observer employing each sensor can be constructed. Moreover, strictly less than half of the sensors can be compromised. The main feature that distinguishes the *resilient* algorithm proposed in this paper with existing works is the usage of a median operation in constructing each component of the state estimate from the state estimates provided by the bank of observers. We then applied the framework to a class of nonlinear systems and validated our results on an inverter-based power distribution network. In the future, we are interested in moving from a centralized to a distributed setup.

References

[1] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-Physical Systems Security - A Survey," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1802–1831, 2017.

[2] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[3] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[4] A. M. Grilo, J. Chen, M. Diaz, D. Garrido, and A. Casaca, "An integrated WSAN and SCADA system for monitoring a critical infrastructure," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 3, pp. 1755–1764, 2014.

[5] M. S. Chong, H. Sandberg, and A. M. Teixeira, "A tutorial introduction to security and privacy for cyber-physical systems," *18th European Control Conference (ECC 2019)*, pp. 968–978, 2019.

[6] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure Estimation and Control for Cyber-Physical Systems Under Adversarial Attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[7] M. S. Chong, M. Wakaiki, and J. P. Hespanha, "Observability of linear systems under adversarial attacks," *Proceedings of the American Control Conference*, pp. 2439–2444, 2015.

[8] C. H. Xie and G. H. Yang, "Secure estimation for cyber-physical systems with adversarial attacks and unknown inputs: An L2-gain method," *International Journal of Robust and Nonlinear Control*, vol. 28, no. 6, pp. 2131–2143, 2018.

[9] M. S. Chong, H. Sandberg, and J. P. Hespanha, "A secure state estimation algorithm for nonlinear systems under sensor attacks," *Proceedings of the 2020 IEEE Conference on Decision and Control*, pp. 5743–5748, 2020.

[10] J. Kim, C. Lee, H. Shim, Y. Eun, and J. H. Seo, "Detection of Sensor Attack and Resilient State Estimation for Uniformly Observable Nonlinear Systems having Redundant Sensors," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1162–1169, 2018.

[11] H. Jeon, S. Aum, H. Shim, and Y. Eun, "Resilient State Estimation for Control Systems Using Multiple Observers and Median Operation," *Mathematical Problems in Engineering*, 2016.

[12] J. G. Lee, J. Kim, and H. Shim, "Fully Distributed Resilient State Estimation Based on Distributed Median Solver," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3935–3942, 2020.

[13] M. Arcak and P. Kokotovic, "Observer-based control of systems with slope-restricted nonlinearities," *IEEE Transactions on Automatic Control*, vol. 46, no. 7, pp. 1146–1150, 2001.

[14] M. S. Chong, R. Postoyan, D. Nešić, L. Kuhlmann, and A. Varsavsky, "A robust circle criterion observer with application to neural mass models," *Automatica*, vol. 48, no. 11, pp. 2986–2989, 2012.

[15] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Power Engineering Review*, vol. 9, no. 4, pp. 101–102, 1989.

[16] F. Andrén, B. Bletterie, S. Kadam, P. Kotsampopoulos, and C. Bucher, "On the Stability of Local Voltage Control in Distribution Networks with a High Penetration of Inverter-Based Generation," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2519–2529, 2015.

[17] M. S. Chong, D. Umsonst, and H. Sandberg, "Local voltage control of an inverter-based power distribution network with a class of slope-restricted droop controllers," *IFAC-PapersOnLine*, vol. 52, no. 20, pp. 163–168, 2019. 8th IFAC Workshop on Distributed Estimation and Control in Networked Systems (NECSYS).

[18] K. Strunz, N. Hatziargyriou, C. Andrieu, and E. Al., "Benchmark systems for network integration of renewable and distributed energy resources," *CIGRE Task Force C*, vol. 6, no. 04-02, p. 78, 2009.

Appendix

A. Proof of Theorem 1

Since Assumption 2 hold, there exist a class $\mathcal{KL}$ function β_m^i and class $\mathcal{K}_\infty$ functions ζ_m^i and γ_m^i such that the solution to system (1) for every $i \in \mathbb{N}_{[1,N]}$ satisfies

$$\begin{aligned}\left|x(t)-\hat{x}^i(t)\right| \leq\ & \beta_m^i\left(\left|x(0)-\hat{x}_{\mathcal{J}}(0)\right|, t\right) \\ & + \zeta_m^i\left(\sup_{s\in[0,t)}|d(s)|\right) + \gamma_m^i\left(\sup_{s\in[0,t)}|a_i(s)|\right). \quad (17)\end{aligned}$$

In the case that $a_i \notin \mathcal{N}_I$, we simplify (17) to:

$$|x(t)-\hat{x}^i(t)| \leq \epsilon^i, \quad (18)$$

where the state estimation error is bounded by some upper bound ϵ^i, since $a_i(t) = 0$ for all $t \in \mathbb{R}_{\geq 0}$ and d is bounded. When using the median operation as presented in (5), it follows that for the final state estimate picked by the median operation,

$$|x(t) - \text{med}(\hat{x}_1, ..., \hat{x}_N)| = |x(t) - \hat{x}(t)| \leq \epsilon, \quad (19)$$

where $\epsilon := \max_{i\in\mathbb{N}_{[1,N]}} \epsilon^i$. In the case that $a_i \in \mathcal{N}_I$,

$$|x(t)-\hat{x}^i(t)| > \epsilon^i. \quad (20)$$

Let us, for a moment, consider the situation in which the median would contain an attack signal, then

$$|x(t)-\hat{x}(t)| > \epsilon. \quad (21)$$

For this to hold true, per definition of the median, $|\mathcal{I}| \geq N/2-1+1 = N/2$, or $M \geq N/2$, which violates our standing assumption. Thus, for $N > 2M$, $\text{med}(\hat{x}_1, \hat{x}_2, ..., \hat{x}_N) = \hat{x}_0$, where $\hat{x}_0$ is the attack-free value of the state estimate.

Now, we can rewrite (17) as

$$\begin{aligned}|x(t)-\hat{x}(t)| \leq\ & \beta_m\left(\left|x(0)-\hat{x}^i(0)\right|, t\right) \\ & + \zeta_m\left(\sup_{s\in[0,t)}|d(s)|\right), \quad (22)\end{aligned}$$

where $\beta_m := \max_{i\in\mathbb{N}_{[1,N]}} \beta_m^i$ is a class $\mathcal{KL}$ function and $\zeta_m := \max_{i\in\mathbb{N}_{[1,N]}} \zeta_m^i$ is a class $\mathcal{K}_\infty$ function. This concludes the proof.

B. Proof of Proposition 1

Let the state estimation error be denoted by $\tilde{x}^i := x - \hat{x}^i$ for every $i \in \mathbb{N}_{[1,N]}$. Then, the state estimation error system is

$$\dot{\tilde{x}}^i = (A - L_i H^i)\tilde{x}^i + \phi(z) - \phi(\xi^i) - L_i(d_i + a_i). \quad (23)$$

The nonlinearity ϕ satisfies Assumption 3, which leads to the conclusion that there exists a $\varepsilon_i(t) \in [\underline{e}_i, \bar{e}_i]$, for $i \in \mathbb{N}_{[1,N]}$ such that

$$\begin{aligned}\phi(z) - \phi(\xi^i) &= \varepsilon(t)(z - \xi^i) \\ &= \varepsilon(t)\eta^i + \varepsilon(t)d - \varepsilon(t)K_i d_i - \varepsilon(t)K_i a_i\end{aligned} \quad (24)$$

where $\varepsilon(t) = \text{diag}(\varepsilon_1(t), \varepsilon_2(t), ..., \varepsilon_N(t))$and $\eta^i := (H - K_i H_i)\tilde{x}^i$. Then,

$$\begin{aligned}\dot{\tilde{x}}^i =&(A - L_i H_i)\tilde{x}^i + \varepsilon(t)\eta^i + \varepsilon(t)d \\ &- (\varepsilon(t)K_i + L_i)d_i - (\varepsilon(t)K_i + L_i)a_i.\end{aligned} \quad (25)$$

To show that the state estimation error system (25) satisfies (4), we show that the time derivative of the candidate Lyapunov function $V_i(\tilde{x}^i) = (\tilde{x}^i)^T P_i \tilde{x}^i$ along the trajectories of the state estimation error system (25) is $V_i(\tilde{x}^i) < 0$. It is equal to $\dot{V}_i(\tilde{x}^i) = (\dot{\tilde{x}}^i)^T P_i \tilde{x}^i + (\tilde{x}^i)^T P_i \dot{\tilde{x}}^i$, or in matrix form:

$$\dot{V}(\tilde{x}_{\mathcal{J}}) = \chi_i^T \begin{bmatrix} \tilde{A}_i^T P_i + P_i \tilde{A}_i & P_i & P_i & -P_i \\ P_i & 0 & 0 & 0 \\ P_i & 0 & 0 & 0 \\ -P_i & 0 & 0 & 0 \end{bmatrix} \chi_i, \quad (26)$$

where $\chi_i := (\tilde{x}^i,\ \varepsilon(t)\eta^i,\ \varepsilon(t)d - (\varepsilon(t)K_i + L_i)d_i$ $-(\varepsilon(t)K_i + L_i)a_i)$, $P_i = P_i^T > 0$ satisfies (10), and $\tilde{A}_i := A - L_i H_i$.

Now, applying (10), the following is obtained

$$\begin{aligned}\dot{V}(\tilde{x}^i) \leq\ & -\nu|\tilde{x}^i|^2 - 2(\eta^i)^T U_i \varepsilon(t)\eta^i + \mu_{i,d}\varepsilon(t)^2|d|^2 \\ & +2(\eta^i)^T \varepsilon(t)^2 U_i \text{diag}(\bar{e}_1^{-1}, \ldots, \bar{e}_N^{-1})\eta^i \\ & +\mu_{i,d}|\varepsilon(t)K_i + L_i|^2\,|d_i|^2 \\ & +\mu_{i,a}|\varepsilon(t)K_i + L_i|^2\,|a_i|^2.\end{aligned} \quad (27)$$

By examining the second and fourth term of the right hand side of the inequality component-by-component, it becomes apparent that for $i \in \mathbb{N}_{[1,N]}$, $\varepsilon_i - \varepsilon_i^2/\bar{e}_i = \varepsilon_i(1 - \varepsilon_i/\bar{e}_i) \geq 0$, as $\varepsilon_i(t) > 0$ and $1 - \varepsilon_i(t)/\bar{e}_i \geq 0$, due to $\varepsilon_i \in [\underline{e}_i, \bar{e}_i]$. Next to that, since U_i is positive definite and a diagonal matrix, we have that $u_i > 0$. Therefore, $-2(\eta^i)^T U_i \varepsilon(t)\eta^i + +2(\eta^i)^T \varepsilon(t)^2 U_i \text{diag}(\bar{e}_1^{-1}, \ldots, \bar{e}_N^{-1})\eta^i \leq 0$ and we obtain

$$\begin{aligned}\dot{V}(\tilde{x}^i) \leq -\nu|\tilde{x}^i|^2 + \mu_{i,d}\varepsilon(t)^2|d|^2 + \mu_{i,d}|\varepsilon(t)K_i + L_i|^2\,|d_i|^2 \\ + \mu_{i,a}|\varepsilon(t)K_i + L_i|^2\,|a_i|^2.\end{aligned} \quad (28)$$

Since $|d_i|^2 \leq |d|^2$,

$$\begin{aligned}\dot{V}(\tilde{x}^i) \leq -\nu|\tilde{x}^i|^2 + \mu_{i,d}(\varepsilon(t)^2 - |\varepsilon(t)K_i + L_i|^2)\,|d|^2 \\ + \mu_{i,a}|\varepsilon(t)K_i + L_i|^2\,|a_i|^2.\end{aligned} \quad (29)$$

Recall that $\varepsilon_i(t) \in [\underline{e}_i, \bar{e}_i]$. Using Young's inequality, the derivative is bounded by

$$\begin{aligned}\dot{V}(\tilde{x}^i) \leq -\nu|\tilde{x}^i|^2 + \mu_{1i,d}(\bar{e}^2 + 2\bar{e}^2|K_i|^2 + 2|L_i|^2)|d|^2 \\ + \mu_{i,a}(2\bar{e}^2|K_i|^2 + 2|L_i|^2)|a_i|^2.\end{aligned} \quad (30)$$

where $\bar{e} = \max\{\bar{e}_1, \bar{e}_2, ..., \bar{e}_N\}$. Note that $V_i(\tilde{x}^i)$ can be sandwiched as follows

$$\lambda_{\min}(P)|\tilde{x}^i|^2 \leq V_i(\tilde{x}^i) \leq \lambda_{\max}(P)|\tilde{x}^i|^2, \quad (31)$$

and that using this in combination with the comparison principle, leads to the following from (30).

$$\begin{aligned}V(\tilde{x}^i(t)) \leq\ & e^{-\lambda_i t}V(\tilde{x}^i(0)) + \delta_i \int_0^t e^{-\lambda_i(t-s)}|d(s)|^2 ds \\ & + \alpha_i \int_0^t e^{-\lambda_i(t-s)}|a_i(s)|^2 ds,\end{aligned} \quad (32)$$

where $\lambda_i = \frac{\nu_i}{\lambda_{\max}(P_i)}$, $\delta_i = \frac{\mu_{i,d}}{\lambda_{\max}(P_i)}(\bar{e}^2 + 2\bar{e}^2|K_i|^2 + 2|L_i|^2)$ and $\alpha_i = 2\frac{\mu_{i,a}}{\lambda_{\max}(P_i)}(\bar{e}^2|K_i|^2 + |L_i|^2)$.

Since $\int_0^t e^{-\lambda_i(t-s)}ds = (1 - e^{-\lambda_i t})/\lambda_i \leq 1/\lambda_i$, we obtain

$$\begin{aligned}V(\tilde{x}^i) \leq\ & e^{-\lambda_i t}V(\tilde{x}^i(0)) + \frac{\delta_i}{\lambda_i}\left(\sup_{s\in[0,t)}|d(s)|^2\right) \\ & + \frac{\alpha_i}{\lambda_i}\left(\sup_{s\in[0,t)}|a_i(s)|^2\right).\end{aligned} \quad (33)$$

Now (31) is re-applied to (33), to obtain

$$\begin{aligned}\left|x(t) - \hat{x}^i(t)\right| \leq\ & \hat{\beta}\left(\left|x(0) - \hat{x}^i(0)\right|, t\right) \\ & + \hat{\zeta}\left(\sup_{s\in[0,t)}|d(s)|\right) + \hat{\gamma}\left(\sup_{s\in[0,t)}|a_i(s)|\right).\end{aligned} \quad (34)$$

where

$$\begin{aligned}\hat{\beta}(r,t) &= \sqrt{\frac{\lambda_{\max}(P_i)}{\lambda_{\min}(P_i)}}e^{-\frac{\lambda_i}{2}t}r,\ \hat{\zeta}(r) = \sqrt{\frac{\delta_i}{\lambda_i\lambda_{\min}(P_i)}}r, \\ \hat{\gamma}(r) &= \sqrt{\frac{\alpha_i}{\lambda_i\lambda_{\min}(P_i)}}r,\end{aligned} \quad (35)$$

and we note that $\hat{\beta}$ is a $\mathcal{KL}$ function; $\hat{\zeta}$ and $\hat{\gamma}$ are $\mathcal{K}_\infty$ functions, which concludes the proof.

C. Proof of Proposition 2

From the linear DistFlow model, (11), (13) and (14), it follows that for $i \in \mathbb{N}_{[1,N]}$ and $t \in \mathbb{R}_{\geq 0}$,

$$\begin{aligned}|v_i^2(t) - \hat{v}_i^2(t)| &\leq |H_i||x(t) - \hat{x}(t)|, \\ &\leq |H_i|\beta_e\left(|x(0) - \hat{x}(0)|, t\right) \\ &\quad + |H_i|\zeta_e\left(\sup_{s\in[0,t)}|d(s)|\right),\end{aligned} \quad (36)$$

where the second inequality is obtained according to Proposition 1 with a class $\mathcal{KL}$ function β_e and class $\mathcal{K}_\infty$ function ζ_e. Hence, we obtain $\beta_v(r,t) = |H_i|\beta_e(r,t)$ and $\zeta_v(r,t) = |H_i|\zeta_e(r,t)$, which are class $\mathcal{KL}$ and $\mathcal{K}_\infty$ functions respectively.