Real-Time Distribution System State Estimation With Asynchronous Measurements

Guido Cavraro, *Associate Member, IEEE*, Joshua Comden, Emiliano Dall'Anese, and Andrey Bernstein

***Abstract*—State estimation is a fundamental task in power systems. Although distribution systems are increasingly equipped with sensing devices and smart meters, measurements are typically reported at different rates and asynchronously; these aspects pose severe strains on workhorse state estimation algorithms, which are designed to process batches of data collected in a synchronous manner from all the measurement units. In this paper, we develop a novel state estimation algorithm to continuously update the estimate of the state based on measurements received in an asynchronous manner from measurement units. The synthesis of the algorithm hinges on a proximal-point type method, implemented in an online fashion, and capable of processing measurements received sequentially from sensors. A performance analysis is presented by providing bounds on the estimation error in terms of the mean and variance that hold at each iteration and asymptotically. The scheme is also compared with a more traditional Weighted Least Squares estimator that compensates for the lack of measurement data by using, as pseudo measurements, the measurement retrieved during a certain time window. Numerical simulations on the IEEE 37-bus feeder corroborate the analytical findings.**

***Index Terms*—State estimation, data fusion, asynchronous sensors, networked systems, sensor networks, stability.**

I. Introduction

THE INTEGRATION of renewable energy sources, electric vehicles, and other power electronics-interfaced distributed energy resources (DERs) is leading to net-loading conditions in distribution networks that are less predictable and highly variable [1]. Hence, it is fundamental for distribution system operators (DSOs) to estimate the system state at a timescale that matches the variability of the net-loading conditions in order to provide meaningful information for the control of the grid.

Manuscript received 13 May 2021; revised 23 September 2021 and 31 December 2021; accepted 23 April 2022. Date of publication 2 May 2022; date of current version 23 August 2022. This work was supported by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) through NREL Directed Research and Development Program under Contract DE-AC36-08GO28308. Paper no. TSG-00748-2021. *(Corresponding author: Guido Cavraro.)*

Guido Cavraro, Joshua Comden, and Andrey Bernstein are with the Power Systems Engineering Center, National Renewable Energy Laboratory, Golden, CO 80401 USA (e-mail: guido.cavraro@nrel.gov; joshua.comden@nrel.gov; andrey.bernstein@nrel.gov).

Emiliano Dall'Anese is with the Department of Electrical, Computer, and Energy Engineering, University of Colorado Boulder, Boulder, CO 80309 USA (e-mail: emiliano.dallanes@colorado.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2022.3171466.

Digital Object Identifier 10.1109/TSG.2022.3171466

Traditional state estimators for transmission systems are generally designed for the case in which the system operator has an overabundance of measurements [2]. Although distribution networks were historically undermetered [3], in recent years utilities have been installing a number of digital devices, such as smart meters, phasor measurement units (PMUs) and intelligent electronic devices throughout the distribution grid for sensing and control purposes. However, such devices are intrinsically heterogeneous and have different sampling rates. For instance, residential smart meters take measurements typically every 15–60 minutes [4], whereas PMUs can possibly take 30–60 measurements every second [5]. Moreover, sensors do not take and report measurements all at the same time [6], [7] to keep communication networks free from congestion [8], [9]. As a result, data available to utilities are *asynchronized* and not enough measurements are typically available to obtain a well-conditioned state estimation problem *at any given time*. Hence, a state estimator for distribution networks should be able to tackle the case in which the number of available measurements at any given moment is much smaller than the number of states, or, in other words, the case in which the state estimation problem is ill posed.

Pseudo measurements are typically used to compensate for the lack of measurements. The main source for pseudo measurements is the historical load data that utilities collect for billing purposes [10]; more advanced methods, e.g., involving neural networks, for modeling pseudo measurements have been developed, too [11]. Historical data or simulated samples can be used to train neural networks, which provide an effective tool for state estimation [12]. The Bayesian linear state estimator in [13] is suitable when probability distributions of load demands are available. If a dynamical model of the network is at hand, approaches based on Kalman filtering can be used [14], [15]. Other methods do not rely on pseudo measurements or a priori probabilistic models but only on measurement data. A matrix completion approach under low-observability conditions was proposed in [16]. Leveraging the communication, actuation, and sensing capabilities of smart inverters, the authors in [17] probe the grid by varying the power injections at selected buses, record the incurred voltage responses, and infer the complex loads at non-actuated buses.

In this paper, we start from the state estimator proposed in [18], designed to estimate the state variation in general time-varying linear systems, and provide the following contributions.

- We tailor the state estimator for applications in three-phase distribution systems in which noisy measurements

are both taken and sent to the system operator asynchronously, and formally analyze the estimation error.
- To benchmark the proposed algorithm, we formulate a Weighted Least Square (WLS) estimator that uses, as pseudo measurements, the measurement retrieved during a certain time window, and analyze its estimation error.

The rest of the paper is organized as follows. The three-phase distribution model is reported in Section II. In Section III, the online state estimation problem is formulated and the asynchronous state estimation algorithm is presented. In Section IV, we introduce a WLS estimator that uses, as pseudo measurements, the data gathered by the system operator during a time window big enough to have a full rank regression matrix. In Section V, we evaluate the performance of the two algorithms on a three-phase 37-bus distribution system under realistic measurement equipment data-reporting scenarios.

Notation: Lower- (upper-) case boldface letters denote column vectors (matrices). Calligraphic symbols are reserved for sets. Symbol $^\top$ stands for transposition. Vectors $\mathbf{0}$ and $\mathbf{1}$ are the all-zero and all-one vectors, while $\mathbf{e}_m$ is the m-th canonical vector. The identity matrix of appropriate dimension is denoted by $\mathbf{I}$. Symbol $\|\mathbf{x}\|$ denotes the Euclidean norm of the vector $\mathbf{x}$; $\|\mathbf{X}\|_F$ denotes the Frobenius norm of the matrix $\mathbf{X}$. Moreover, $\|\mathbf{x}\|_\mathbf{Q} \coloneqq \mathbf{x}^\top\mathbf{Q}\mathbf{x}$ for a positive definite matrix $\mathbf{Q}$. The diagonal matrix having the entries of the vector $\mathbf{x}$ on its diagonal is denoted as $\mathrm{diag}(\mathbf{x})$. Given a matrix $\mathbf{A}$, its kernel, namely the set of all vectors $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{0}$, is denoted as $\ker\mathbf{A}$. The expectation operator is defined as $\mathbb{E}[\cdot]$. The real part, the imaginary part, the complex conjugate of a complex vector (matrix) $\mathbf{x}$ ($\mathbf{X}$) are denoted by $\Re(\mathbf{x})$, $\Im(\mathbf{x})$, $\bar{\mathbf{x}}$ ($\Re(\mathbf{X})$, $\Im(\mathbf{X})$, $\bar{\mathbf{X}}$); by $|\mathbf{x}|$ we denote instead the vector whose entries are the absolute value of the entries of $\mathbf{x}$.

II. Three-Phase Distribution Network Model

We model a three-phase power distribution grid having $B+1$ buses with the graph $\mathcal{G} = (\mathcal{B}, \mathcal{E})$, where the nodes in the set $\mathcal{B} \coloneqq \{0, \dots, B\}$ represent the grid buses, and the edges in the set $\mathcal{E}$ correspond to the distribution lines. The substation bus is indexed by $i = 0$ and it is assumed to be an ideal voltage generator (slack bus) imposing the nominal voltage $\mathbf{v}_0 = [1, 1 - \jmath\frac{2\pi}{3}, 1 + \jmath\frac{2\pi}{3}]^\top$. Similar to prior works in the context of distribution system state estimation (see, e.g., [13], [19], [20]), in the rest of the paper we assume that the system operator knows the voltage $\mathbf{v}_0$.

At each multiphase bus, the model of the distribution system can have: (i) grounded wye-connected loads/sources; (ii) ungrounded delta connections; (iii) a combination of wye-connected and delta-connected loads/sources; or, (iv) a combination of line-to-line and line-to-grounded-neutral devices at the secondary of distribution transformers. For simplicity, we consider only three-phase wye-connected or delta-connected buses. The more general case can be found in [21]. Assume that in the network there are N^Y wye-connected buses and N^Δ delta-connected buses. Let bus n be a wye-connected bus; its power injections on each phase ϕ are denoted by $s_n^\phi \in \mathbb{C}$; collect them in the vector $\mathbf{s}_n^Y \coloneqq [s_n^a, s_n^b, s_n^c]^\top$. On the contrary, if bus n is a delta-connected bus, its power and current injections from phase ϕ' to ϕ are denoted by $s_n^{\phi\phi'} \in \mathbb{C}$ and $i_n^{\phi\phi'} \in \mathbb{C}$, respectively. Collect them in the vectors $\mathbf{s}_n^\Delta \coloneqq [s_n^{ab}, s_n^{bc}, s_n^{ca}]^\top$ and $\mathbf{i}_n^\Delta \coloneqq [i_n^{ab}, i_n^{bc}, i_n^{ca}]^\top$. The phase ϕ to ground voltages of bus n are $v_n^\phi \in \mathbb{C}$ and are collected in the vector $\mathbf{v}_n \coloneqq [v_n^a, v_n^b, v_n^c]^\top$. The phase net current injections on each phase ϕ of node n are denoted by $i_n^\phi \in \mathbb{C}$ and stacked in the vector $\mathbf{i}_n \coloneqq [i_n^a, i_n^b, i_n^c]^\top$.

To conveniently state the power flow equations, collect the aforesaid nodal quantities, except the ones associated with the substation, in the vectors $\mathbf{s}^Y, \mathbf{s}^\Delta, \mathbf{v}, \mathbf{i}, \mathbf{i}^\Delta \in \mathbb{C}^{3B}$:

$$\mathbf{s}^Y \coloneqq \left[\left(\mathbf{s}_1^Y\right)^\top, \dots, \left(\mathbf{s}_B^Y\right)^\top\right]^\top,\ \mathbf{s}^\Delta \coloneqq \left[\left(\mathbf{s}_1^\Delta\right)^\top, \dots, \left(\mathbf{s}_B^\Delta\right)^\top\right]^\top$$
$$\mathbf{v} \coloneqq \left[(\mathbf{v}_1)^\top, \dots, (\mathbf{v}_B)^\top\right]^\top,\quad \mathbf{i} \coloneqq \left[(\mathbf{i}_1)^\top, \dots, (\mathbf{i}_B)^\top\right]^\top$$
$$\mathbf{i}^\Delta \coloneqq \left[\left(\mathbf{i}_1^\Delta\right)^\top, \dots, \left(\mathbf{i}_B^\Delta\right)^\top\right]^\top$$

Let $\mathbf{Y}$ be the three-phase bus admittance matrix and partition it as

$$\mathbf{Y} = \begin{bmatrix}\mathbf{Y}_{00} & \mathbf{Y}_{0L}\\ \mathbf{Y}_{L0} & \mathbf{Y}_{LL}\end{bmatrix} \in \mathbb{C}^{3(B+1)\times 3(B+1)}.$$

where the 0-th block-row and block-column are associated with the substation and $\mathbf{Y}_{00} \in \mathbb{C}^{3\times 3}$, $\mathbf{Y}_{0L} \in \mathbb{C}^{3\times 3B}$, $\mathbf{Y}_{L0} \in \mathbb{C}^{3B\times 3}$, $\mathbf{Y}_{LL} \in \mathbb{C}^{3B\times 3B}$. The power flow equations read [21]

$$\mathrm{diag}\left(\mathbf{H}^\top\overline{\mathbf{i}^\Delta}\right)\mathbf{v} + \mathbf{s}^Y = \mathrm{diag}(\mathbf{v})\bar{\mathbf{i}} \tag{1a}$$
$$\mathrm{diag}(\mathbf{H}\mathbf{v})\overline{\mathbf{i}^\Delta} = \mathbf{s}^\Delta \tag{1b}$$
$$\mathbf{Y}_{L0}\mathbf{v}_0 + \mathbf{Y}_{LL}\mathbf{v} = \mathbf{i} \tag{1c}$$

where $\mathbf{H}$ is a block diagonal matrix

$$\mathbf{H} \coloneqq \begin{bmatrix}\mathbf{\Gamma} & & \\ & \ddots & \\ & & \mathbf{\Gamma}\end{bmatrix},\quad \mathbf{\Gamma} \coloneqq \begin{bmatrix}1 & -1 & 0\\ 0 & 1 & -1\\ -1 & 0 & 1\end{bmatrix}.$$

By eliminating the currents $(\mathbf{i}, \mathbf{i}^\Delta)$ in Equation (1) and assuming that $\mathbf{Y}_{LL}$ is invertible [21], we have the following fixed-point equation around the zero-load voltage $\mathbf{w} \coloneqq -\mathbf{Y}_{LL}^{-1}\mathbf{Y}_{L0}\mathbf{v}_0$:

$$\mathbf{v} = \mathbf{w} + \mathbf{Y}_{LL}^{-1}\left(\mathrm{diag}(\overline{\mathbf{v}})^{-1}\overline{\mathbf{s}^Y} + \mathbf{H}^\top\mathrm{diag}(\mathbf{H}\overline{\mathbf{v}})^{-1}\overline{\mathbf{s}^\Delta}\right). \tag{2}$$

The nonlinear equation (2) can be linearized around the zero-load voltage $\mathbf{w}$ to make the voltage deviation phasors

$$\tilde{\mathbf{v}} \coloneqq \mathbf{v} - \mathbf{w} \tag{3}$$
$$|\tilde{\mathbf{v}}| \coloneqq |\mathbf{v}| - |\mathbf{w}| \tag{4}$$

a linear function of the power injections (see [21] for details)

$$\begin{bmatrix}\tilde{\mathbf{v}}\\ |\tilde{\mathbf{v}}|\end{bmatrix} \approx \begin{bmatrix}\mathbf{M}^Y & \mathbf{M}^\Delta\\ \mathbf{K}^Y & \mathbf{K}^\Delta\end{bmatrix}\begin{bmatrix}\mathbf{p}^Y\\ \mathbf{q}^Y\\ \mathbf{p}^\Delta\\ \mathbf{q}^\Delta\end{bmatrix} \tag{5}$$

where the power injections $(\mathbf{s}^Y, \mathbf{s}^\Delta)$ are split into their active parts $\mathbf{p}^Y \coloneqq \Re(\mathbf{s}^Y)$, $\mathbf{p}^\Delta \coloneqq \Re(\mathbf{s}^\Delta)$ and reactive parts $\mathbf{q}^Y \coloneqq \Im(\mathbf{s}^Y)$, $\mathbf{q}^\Delta \coloneqq \Im(\mathbf{s}^\Delta)$. The matrices $\{\mathbf{M}^Y, \mathbf{M}^\Delta\} \in \mathbb{C}^{3B\times 6B}$ are defined as

$$\mathbf{M}^Y \coloneqq \left[\mathbf{Y}_{LL}^{-1}\mathrm{diag}(\overline{\mathbf{w}})^{-1}, -\jmath\mathbf{Y}_{LL}^{-1}\mathrm{diag}(\overline{\mathbf{w}})^{-1}\right]$$

$$\mathbf{M}^{\Delta} \coloneqq \left[\mathbf{Y}_{LL}^{-1}\mathbf{H}^{\top}\mathrm{diag}(\mathbf{H}\overline{\mathbf{w}})^{-1}, -\jmath\mathbf{Y}_{LL}^{-1}\mathbf{H}^{\top}\mathrm{diag}(\mathbf{H}\overline{\mathbf{w}})^{-1}\right]$$
$$\mathbf{K}^{Y} \coloneqq \mathrm{diag}(|\mathbf{w}|)\Re\Big(\mathrm{diag}(\mathbf{w})^{-1}\mathbf{M}^{Y}\Big)$$
$$\mathbf{K}^{\Delta} \coloneqq \mathrm{diag}(|\mathbf{w}|)\Re\Big(\mathrm{diag}(\mathbf{w})^{-1}\mathbf{M}^{\Delta}\Big).$$

Two kinds of metering devices are considered in this work: smart meters, able to measure power injections and voltage magnitudes; and phasor measurement units (PMUs), able to measure power injections and complex voltages [22], [23]. Buses hosting measurement devices are collected in the set $\mathcal{M}$ of cardinality $|\mathcal{M}| = M$. Buses endowed with smart meters are collected in the set $\mathcal{M}_{\mathrm{SM}}$, whereas buses endowed with PMUs are collected in the set $\mathcal{M}_{\mathrm{PMU}}$ so that $\mathcal{M} = \mathcal{M}_{\mathrm{SM}} \oplus \mathcal{M}_{\mathrm{PMU}}$. Using (5), the measurable quantities can be approximated by a function of power injections as:

$$\begin{bmatrix}\Re(\tilde{\mathbf{v}})\\ \Im(\tilde{\mathbf{v}})\\ |\tilde{\mathbf{v}}|\\ \mathbf{p}^{Y}\\ \mathbf{q}^{Y}\\ \mathbf{p}^{\Delta}\\ \mathbf{q}^{\Delta}\end{bmatrix} = \begin{bmatrix}\Re(\mathbf{M}^{Y}) & \Re(\mathbf{M}^{\Delta})\\ \Im(\mathbf{M}^{Y}) & \Im(\mathbf{M}^{\Delta})\\ \mathbf{K}^{Y} & \mathbf{K}^{\Delta}\\ \mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{p}^{Y}\\ \mathbf{q}^{Y}\\ \mathbf{p}^{\Delta}\\ \mathbf{q}^{\Delta}\end{bmatrix} = \boldsymbol{\Phi}\begin{bmatrix}\mathbf{p}^{Y}\\ \mathbf{q}^{Y}\\ \mathbf{p}^{\Delta}\\ \mathbf{q}^{\Delta}\end{bmatrix} \quad (6)$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{21B\times 12B}$. The system state is the real vector $\mathbf{x} \coloneqq [(\mathbf{p}^{Y})^{\top}, (\mathbf{q}^{Y})^{\top}, (\mathbf{p}^{\Delta})^{\top}, (\mathbf{q}^{\Delta})^{\top}]^{\top}$ of dimension $N \coloneqq 12B$, which denotes the nodal power injections vector.

Data from PMUs, DERs, and smart meters are generally not synchronized and have significant gaps in time between measurements [6], [24]. Hence, we consider the case in which only a subset of sensors send data to the system operator at times $t = t_1, t_2, \ldots$, potentially not equally spaced. However, to keep the notation simple and without loss of generality, in the following we assume that the measurements are collected at times $t = 1, 2, \ldots$, from a time-varying subset of buses, denoted $\mathcal{S}(t)$, and are collected in the vector $\mathbf{y}(t)$. The vector $\mathbf{y}(t)$ can be written as a function of the system state as:

$$\mathbf{y}(t) = \mathbf{S}(t)\Big[\Re(\tilde{\mathbf{v}})^{\top}(t), \Im(\tilde{\mathbf{v}})^{\top}(t), |\tilde{\mathbf{v}}|^{\top}(t), (\mathbf{p}^{Y})^{\top}(t), (\mathbf{q}^{Y})^{\top}(t), (\mathbf{p}^{\Delta})^{\top}(t), (\mathbf{q}^{\Delta})^{\top}(t)\Big]^{\top} + \mathbf{n}(t) \quad (7)$$

where $\mathbf{S}(t)$ is a selection matrix that picks the quantities measured at time t and $\mathbf{n}(t)$ represents the measurement noise.

Before formally defining $\mathbf{S}(t)$, assume, for simplicity but without loss of generality, that the system operator gathers the same number of measurements at each time step, i.e., $|\mathcal{S}(t)| = S$ for all t. Matrix $\mathbf{S}(t)$ can be constructed by simple blocks that depend on the list of buses measured $\mathcal{S}(t) \coloneqq \{s_1, \ldots, s_S\}$ and the type of measurement device at each bus

$$\mathbf{S}(t) = \begin{bmatrix}\mathbf{S}_{s_1}^{\top} & \cdots & \mathbf{S}_{s_S}^{\top}\end{bmatrix}^{\top}$$

with every $\mathbf{S}_{s_i}$ is defined by one of two ways:

- if $s_i \in \mathcal{M}_{SM}$, then $\mathbf{S}_{s_i} \in \{0,1\}^{15\times 21B}$

$$\mathbf{S}_{s_i} = \begin{bmatrix}\mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i}\end{bmatrix} \quad (8)$$

- if $s_i \in \mathcal{M}_{PMU}$, then $\mathbf{S}_{s_i} \in \{0,1\}^{18\times 21B}$

$$\mathbf{S}_{s_i} = \begin{bmatrix}\mathbf{E}_{s_i} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{E}_{s_i} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E}_{s_i}\end{bmatrix} \quad (9)$$

and where $\mathbf{E}_{s_i} = (\mathbf{e}_{s_i}^{\intercal} \otimes \mathbf{I}) \in \{0,1\}^{3\times 3B}$. Moreover, we collect all the matrices $\mathbf{S}_{s_i}$'s in the matrix

$$\mathbf{S} = \begin{bmatrix}\mathbf{S}_{s_1}^{\top} & \cdots & \mathbf{S}_{s_M}^{\top}\end{bmatrix}^{\top}.$$

Remark 1: The more general case in which line currents are also measured can be straightforwardly included in our setup. Define as $\mathbf{i}_{mn}$ the vector collecting the line currents between bus m and bus n. There exists a matrix $\mathbf{A}_{mn}$ such that [21]

$$\begin{bmatrix}\Re(\mathbf{i}_{mn})\\ \Im(\mathbf{i}_{mn})\end{bmatrix} = \mathbf{A}_{mn}\begin{bmatrix}\Re(\mathbf{v})\\ \Im(\mathbf{v})\end{bmatrix}.$$

The linear model (6) can be then straightforwardly augmented by adding, for each line current sensors, terms of the form

$$\begin{bmatrix}\Re(\mathbf{i}_{mn})\\ \Im(\mathbf{i}_{mn})\end{bmatrix} = \mathbf{A}_{mn}\begin{bmatrix}\Re(\mathbf{M}^{Y}) & \Re(\mathbf{M}^{\Delta})\\ \Im(\mathbf{M}^{Y}) & \Im(\mathbf{M}^{\Delta})\end{bmatrix}\begin{bmatrix}\mathbf{p}^{Y}\\ \mathbf{q}^{Y}\\ \mathbf{p}^{\Delta}\\ \mathbf{q}^{\Delta}\end{bmatrix}.$$

However, to have a lighter notation and facilitate readability, in the following we will outline our framework for sensors taking only measurements on nodes.

III. An Online State Estimation Algorithm

Here, the online state estimator designed in [18] is presented and formally characterized. Let the state of the network at time t be denoted as $\mathbf{x}(t)$. Denote the state variation at time t as $\boldsymbol{\delta}(t) \coloneqq \mathbf{x}(t) - \mathbf{x}(t-1)$. We make the following assumption.

Assumption 1: There exists a scalar Δ_x, $\Delta_x < \infty$, such that, for every t:

$$\|\boldsymbol{\delta}(t)\| \leq \Delta_x. \quad (10)$$

Roughly speaking, this assumption states that the state variation is finite and hence two consecutive states cannot be arbitrarily different. By combining (6) with (7), we obtain the following linear measurement model

$$\mathbf{y}(t) = \mathbf{S}(t)\boldsymbol{\Phi}\mathbf{x}(t) + \mathbf{n}(t). \quad (11)$$

Notice that the size of $\mathbf{y}(t)$ and $\mathbf{n}(t)$, denoted by M_t, varies as a function of the type of reporting metering devices. For instance, if at time t the system operator gathers measurements from C buses in $\mathcal{M}_{\mathrm{SM}}$ and $S-C$ buses in $\mathcal{M}_{\mathrm{PMU}}$, then $M_t = 15C + 18(S-C)$. We also make the following standard assumption on the measurement error

Assumption 2: For every time t, $\mathbf{n}(t)$ is zero mean, and with positive definite covariance $\mathbf{N}_t \coloneqq \mathbb{E}[\mathbf{n}(t)\mathbf{n}(t)^{\top}]$.

At each time step t, given the measurement vector $\mathbf{y}(t)$, the estimate $\hat{\mathbf{x}}(t)$ of the true state $\mathbf{x}(t)$ is obtained by solving the following time-varying *regularized WLS problem* [7]

$$\hat{\mathbf{x}}(t) = \arg\min_{\mathbf{x}} \|\mathbf{y}(t) - \mathbf{S}(t)\boldsymbol{\Phi}\mathbf{x}\|^2_{\mathbf{N}_t^{-1}} + \gamma\left\|\mathbf{x} - \hat{\mathbf{x}}(t-1)\right\|^2 \quad (12)$$

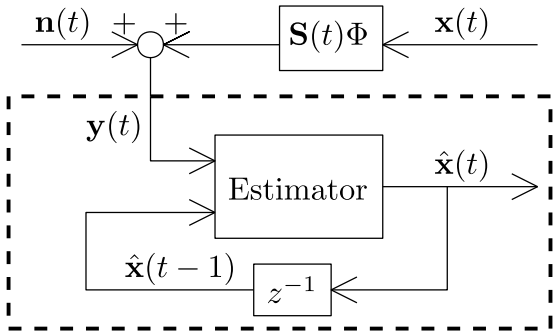

Fig. 1. Block scheme of the dynamical system described by equation (13).

where $\hat{\mathbf{x}}(0)$ is a given (arbitrary) initial state. The first term in (12) is a classical WLS cost. However, because the number of measurements available at any time t is assumed to be smaller than the number of system states (namely $M_t < \mathbf{x} = 12B$), the first term is convex *but not strictly convex*. The second term in (12) acts as a regularizer that penalizes the Euclidean distance of the new estimate from the older one and makes (12) a strongly convex problem having a unique solution. The real scalar $\gamma > 0$ will be referred to as the *inertia parameter* and has a straightforward interpretation: the smaller γ is, the farther the new estimate $\hat{\mathbf{x}}(t)$ is allowed to be from $\hat{\mathbf{x}}(t-1)$; the bigger γ is, the closer the new estimate $\hat{\mathbf{x}}(t)$ will be from the previous one. Since (12) is a quadratic unconstrained problem, its solution $\hat{\mathbf{x}}(t)$ is

$$\hat{\mathbf{x}}(t) = \mathbf{\Lambda}(t)\hat{\mathbf{x}}(t-1) + \frac{1}{\gamma}\mathbf{\Lambda}(t)\mathbf{\Phi}^\top \mathbf{S}(t)^\top \mathbf{N}_t^{-1}\mathbf{y}(t) \tag{13}$$

where

$$\mathbf{\Lambda}(t) \coloneqq \gamma\Big(\mathbf{\Phi}^\top \mathbf{S}(t)^\top \mathbf{N}_t^{-1}\mathbf{S}(t)\mathbf{\Phi} + \gamma\mathbf{I}\Big)^{-1}; \tag{14}$$

Equation (13) represents the sought online asynchronous state estimator (OASE) and constitutes a linear dynamical closed-loop system, whose block scheme is reported in Figure 1. The new estimate $\hat{\mathbf{x}}(t)$ can be recursively computed given the previous estimate $\hat{\mathbf{x}}(t-1)$, the new measurement $\mathbf{y}(t)$, and the set of reporting sensors $\mathcal{S}(t)$. The inverse on the right hand side of (14) always exists and $\mathbf{\Lambda}(t) \in \mathbb{R}^{N\times N}$ is a symmetric positive definite matrix for every t [18].

Recall that the measurements are processed as they come in, and that $\mathbf{y}(t)$ carries information of a limited number of buses. We make the following assumption.

Assumption 3: There exists a constant $\tau > 0$ such that the system operator gathers measurements from every sensor at least once in the τ-long interval $[t, t+1, \ldots, t+\tau-1]$. Moreover, the matrix $\mathbf{S\Phi}$ is a full column rank matrix, namely,

$$\mathbf{S\Phi x} = \mathbf{0} \Leftrightarrow \mathbf{x} = \mathbf{0}. \tag{15}$$

Notice that Assumption 3 is in line with the current practice. In fact, sensors report measurement data periodically to system operators, e.g., smart meters send data on an hourly basis [25]. Roughly speaking, equation (15) means that the system would be observable if all the sensors were taking measurements at the same time. Unfortunately, this is not the case for most distribution networks where measurements devices are heterogeneous and non-synchronized. Equation (15) holds when every bus is endowed with a measurement device, but also in the case in which some buses are not monitored, provided that enough measurements come from other locations.

Define the estimation error $\boldsymbol{\xi}(t) \coloneqq \hat{\mathbf{x}}(t) - \mathbf{x}(t)$. The following result holds true.

Lemma 1: The estimation error obeys the following dynamic system [18]

$$\boldsymbol{\xi}(t) = \mathbf{\Lambda}(t)\boldsymbol{\xi}(t-1) - \mathbf{\Lambda}(t)\boldsymbol{\delta}(t) + \frac{1}{\gamma}\mathbf{\Lambda}(t)\mathbf{\Phi}^\top \mathbf{S}(t)^\top \mathbf{N}_t^{-1}\mathbf{n}(t). \tag{16}$$

Proof: Simple manipulations yield, for every $\mathbf{x}$

$$(\mathbf{\Lambda}(t) - \mathbf{I})\mathbf{x} = -\frac{1}{\gamma}\mathbf{\Lambda}(t)\mathbf{\Phi}^\top \mathbf{S}(t)^\top \mathbf{N}^{-1}\mathbf{S}(t)\mathbf{\Phi}\mathbf{x}. \tag{17}$$

Using the definitions of $\boldsymbol{\delta}(t)$ and $\boldsymbol{\xi}(t)$, and using equations (11), (13), and (17) we obtain

$$\begin{aligned}\boldsymbol{\xi}(t) &= \mathbf{\Lambda}(t)(\boldsymbol{\xi}(t-1) - \boldsymbol{\delta}(t)) + \frac{1}{\gamma}\mathbf{\Lambda}(t)\mathbf{\Phi}^\top \mathbf{S}(t)^\top \mathbf{N}_t^{-1}\mathbf{n}(t) \\ &\quad + (\mathbf{\Lambda}(t) - \mathbf{I})\mathbf{x} + \frac{1}{\gamma}\mathbf{\Lambda}(t)\mathbf{\Phi}^\top \mathbf{S}(t)^\top \mathbf{N}_t^{-1}\mathbf{x}(t) \\ &= \mathbf{\Lambda}(t)\boldsymbol{\xi}(t-1) - \mathbf{\Lambda}(t)\boldsymbol{\delta}(t) + \frac{1}{\gamma}\mathbf{\Lambda}(t)\mathbf{\Phi}^\top \mathbf{S}(t)^\top \mathbf{N}_t^{-1}\mathbf{n}(t).\end{aligned}$$

■

Equation (16) represents a closed loop system in which the previous estimation error is fed back and the input depends on $\boldsymbol{\delta}(t)$ and $\mathbf{n}(t)$ whose stability properties, summarized in the next result, are formally proved in [18].

Proposition 1 [18, Th. 2]: Let Assumption 1, 2, and 3 hold. Denote as $\bar{\lambda}$ the smallest among the nonzero eigenvalue of the matrices in the sequence $\{(\mathbf{S}(t)\mathbf{\Phi})^\top \mathbf{N}_t^{-1}(\mathbf{S}(t)\mathbf{\Phi})\}$, $t = 1, 2, \ldots$, the error mean as $\boldsymbol{\mu}(t) \coloneqq \mathbb{E}[\boldsymbol{\xi}(t)]$, and the error variance as $\mathbf{\Sigma}(t) \coloneqq \mathbb{E}[(\boldsymbol{\xi}(t) - \boldsymbol{\mu}(t))(\boldsymbol{\xi}(t) - \boldsymbol{\mu}(t))^\top]$. Define the variables $m(t) \coloneqq \|\mathbf{N}_t^{-1}\|_F$, $m = sup_t\{m(t)\}$, $\psi \coloneqq \frac{\gamma}{\gamma+\bar{\lambda}}$, $C(t) \coloneqq \|(\mathbf{S}(t)\mathbf{\Phi})^\top \otimes (\mathbf{S}(t)\mathbf{\Phi})^\top\|_F$, $C \coloneqq \sup_t\{C(t)\}$. At every time $t = 1, 2, \ldots$, we have

$$\|\boldsymbol{\mu}(t)\| \le \psi^{\left\lfloor \frac{t}{\tau} \right\rfloor}\|\boldsymbol{\xi}(0)\| + \sum_{k=1}^{t} \psi^{\left\lfloor \frac{t+1-k}{\tau} \right\rfloor}\Delta_x(k) \tag{18}$$

$$\|\mathbf{\Sigma}(t)\|_F \le \psi^{\left\lfloor \frac{t}{\tau} \right\rfloor}\|\mathbf{\Sigma}(0)\|_F + \sum_{k=1}^{t} \psi^{\left\lfloor \frac{t+1-k}{\tau} \right\rfloor}C(k)m(k). \tag{19}$$

Moreover, the error mean, the error covariance, and the average distance between the estimate $\hat{\mathbf{x}}$ and the true state $\mathbf{x}$ are asymptotically upper-bounded, i.e.,

$$\limsup_{t\to\infty}\|\boldsymbol{\mu}(t)\| \le \tau\Delta_x\left(1 + \frac{\gamma}{\bar{\lambda}}\right) \tag{20}$$

$$\limsup_{t\to\infty}\|\mathbf{\Sigma}(t)\|_F \le \frac{\tau C m}{\gamma^2}\left(1 + \frac{\gamma}{\bar{\lambda}}\right) \tag{21}$$

$$\limsup_{t\to\infty}\sqrt{\mathbb{E}\big[\boldsymbol{\xi}(t)^\top\boldsymbol{\xi}(t)\big]} \le \tau\sqrt{\frac{C^2m^2}{\gamma^4} + \Delta_x^2}\left(1 + \frac{\gamma}{\bar{\lambda}}\right). \tag{22}$$

In summary, $\boldsymbol{\xi}$ has a mean and a covariance that are finite at every time t and asymptotically upper bounded. Furthermore,

this property holds for *every* choice of γ. Although we are not able to provide a closed form expression for the inertia parameter γ^* that minimizes the upper bound in (22), namely,

$$\gamma^* = \arg\min_{\gamma} \tau \sqrt{\frac{C^2 m^2}{\gamma^4} + \Delta_x^2}\left(1 + \frac{\gamma}{\bar{\lambda}}\right)$$

like we did in the case of bounded noise [7], γ^* can be easily found numerically. Nevertheless, a good rule of thumb to choose the inertia parameter is the following. If the state variation dominates the measurement noise, i.e., $\Delta_x \gg m$, γ should be chosen small so that the state estimate is able to promptly chase the true state. On the other hand, if the error introduced by the sensors is bigger than the state variability, i.e., $\Delta_x \ll m$, γ should be chosen large in order to somehow filter the measurement noise.

Remark 2: A similar dynamic state estimator was adopted in [19], where a *prediction-correction* method is applied to the DSSE. The scheme proposed in this paper does not require a prediction step and can handle asynchronous measurements. Another similar approach can be found in [26]. However, authors of [26] are considering systems fully observable at any time and use a proximal point method to solve an optimization problem providing the state estimate at a certain time instant.

IV. A WLS State Estimator

Traditionally, state estimation is performed by solving a WLS problem. Because the number of measurements must be at least equal to the number of states, pseudo measurements are introduced so that the WLS problem has the form

$$\hat{\mathbf{x}}_{LS}(t) = \arg\min_{\mathbf{x}} \quad \|\mathbf{y}(t) - \mathbf{S}(t)\mathbf{\Phi}\mathbf{x}\|^2 + \|\mathbf{y}_{PM} - \mathbf{A}_{PM}\mathbf{x}\|^2. \quad (23)$$

where the vector $\mathbf{y}_{PM}$ collects all the pseudo measurements and the matrix $\mathbf{A}_{PM}$ links them to the system state. Similar to (12), the first term in the cost of (23) aims at finding the state that best matches the last retrieved measurements. The second term makes the WLS problem strictly convex and the estimate $\hat{\mathbf{x}}(t)$ well defined. Pseudo measurements are usually obtained from historical data and have much larger measurement errors than the telemetered real-time measurements [27]. For instance, pseudo measurements of load demands are expected to introduce a maximum error of more than 50% [28]. Given the high uncertainty introduced by pseudo measurements, and given that sensors frequently report to system operators their measurement data, e.g., on a hourly basis, we propose directly using the last retrieved measurement for the buses not reporting at time t. Assumption 3 ensures that such measurements were taken within the last τ time instants. Hence, the vector $\mathbf{y}_{PM}$ can be written as

$$\mathbf{y}_{PM} = \mathbf{P}(t)\begin{bmatrix} \mathbf{y}(t-1) \\ \vdots \\ \mathbf{y}(t-\tau+1) \end{bmatrix} \quad (24)$$

with $\mathbf{P}(t)$ being a selection matrix opportunely selecting the measurements needed. Equation (11) and the definition of state variation $\boldsymbol{\delta}(t)$ yield

$$\begin{aligned}\mathbf{y}(t-k) &= \mathbf{S}(t-k)\mathbf{\Phi}\mathbf{x}(t-k) + \mathbf{n}(t-k) \\ &= \mathbf{S}(t-k)\mathbf{\Phi}\mathbf{x}(t) + \mathbf{n}(t-k) \\ &\quad - \mathbf{S}(t-k)\mathbf{\Phi}\sum_{\ell=1}^{k}\boldsymbol{\delta}(t-k+\ell).\end{aligned}$$

Hence, it holds

$$\begin{aligned}\begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}(t-1) \\ \vdots \\ \mathbf{y}(t-\tau+1) \end{bmatrix} &= \begin{bmatrix} \mathbf{S}(t) \\ \mathbf{S}(t-1) \\ \vdots \\ \mathbf{S}(t-\tau+1) \end{bmatrix}\mathbf{\Phi}\mathbf{x}(t) + \begin{bmatrix} \mathbf{n}(t) \\ \mathbf{n}(t-1) \\ \vdots \\ \mathbf{n}(t-\tau+1) \end{bmatrix} \\ &\quad - \begin{bmatrix} \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{S}(t-1)\mathbf{\Phi} & \dots & \mathbf{0} \\ \vdots & \ddots & \\ \mathbf{S}(t-\tau+1)\mathbf{\Phi} & \dots & \mathbf{S}(t-\tau+1)\mathbf{\Phi} \end{bmatrix} \\ &\quad \times \begin{bmatrix} \boldsymbol{\delta}(t) \\ \boldsymbol{\delta}(t-1) \\ \vdots \\ \boldsymbol{\delta}(t-\tau+2) \end{bmatrix}. \end{aligned} \quad (25)$$

Introducing the notation

$$\mathbf{y}_\tau(t) = \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}(t-1) \\ \vdots \\ \mathbf{y}(t-\tau+1) \end{bmatrix}, \quad \mathbf{\Phi}_\tau(t) = \begin{bmatrix} \mathbf{S}(t) \\ \mathbf{S}(t-1) \\ \vdots \\ \mathbf{S}(t-\tau+1) \end{bmatrix}\mathbf{\Phi}$$

$$\boldsymbol{\delta}_\tau(t) = \begin{bmatrix} \boldsymbol{\delta}(t) \\ \boldsymbol{\delta}(t-1) \\ \vdots \\ \boldsymbol{\delta}(t-\tau+2) \end{bmatrix}, \quad \mathbf{n}_\tau(t) = \begin{bmatrix} \mathbf{n}(t) \\ \mathbf{n}(t-1) \\ \vdots \\ \mathbf{n}(t-\tau+1) \end{bmatrix}$$

$$\mathbf{N}_\tau(t) = \begin{bmatrix} \mathbf{N}_t & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{N}_{t-\tau+1} \end{bmatrix}$$

$$\mathbf{\Xi}_\tau(t) = -\begin{bmatrix} \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{S}(t-1)\mathbf{\Phi} & \dots & \mathbf{0} \\ \vdots & \ddots & \\ \mathbf{S}(t-\tau+1)\mathbf{\Phi} & \dots & \mathbf{S}(t-\tau+1)\mathbf{\Phi} \end{bmatrix}$$

we can compactly rewrite (25) as

$$\mathbf{y}_\tau(t) = \mathbf{\Phi}_\tau(t)\mathbf{x}(t) + \mathbf{\Xi}_\tau(t)\boldsymbol{\delta}_\tau(t) + \mathbf{n}_\tau(t). \quad (26)$$

From (24) and (25), $\mathbf{y}(t)$ and $\mathbf{y}_{PM}$ can be written as

$$\mathbf{y}(t) = \begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{y}_\tau(t) \quad (27)$$

$$\mathbf{y}_{PM} = \begin{bmatrix}\mathbf{0} & \mathbf{P}(t)\end{bmatrix}\mathbf{y}_\tau(t). \quad (28)$$

Using (23), (24), (27), and (28), and defining $\mathbf{\Pi}(t) := \begin{bmatrix}\mathbf{I} & \mathbf{P}(t)\end{bmatrix}$, problem (23) can be expressed as

$$\hat{\mathbf{x}}_{LS}(t) = \arg\min_{\mathbf{x}} \quad \|\mathbf{\Pi}(t)(\mathbf{y}_\tau(t) - \mathbf{\Phi}_\tau(t)\mathbf{x})\|^2_{\mathbf{N}_\tau(t)};$$

and the state estimate has the closed form

$$\begin{aligned}\hat{\mathbf{x}}_{LS}(t) &= \left(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\right)^{-1} \\ &\quad \times \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{y}_\tau(t). \end{aligned} \quad (29)$$

Since $\mathbf{y}(t)$ and $\mathbf{y}_{PM}$ collect measurements coming from all the sensors deployed in the grid, even if possibly taken at

different times in the interval $[t-\tau+1, t]$, the matrix $\mathbf{S\Phi}$ can be obtained by permutation of rows of $\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)$. Thus $\text{rank}(\mathbf{S\Phi}) = \text{rank}(\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t))$. Assumption 3 ensures then that $\text{rank}(\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t))$ is full column rank and that the inverse in (29) exists.

Define now the estimation error of the WLS estimator $\boldsymbol{\xi}_{LS}(t) \coloneqq \hat{\mathbf{x}}_{LS}(t) - \mathbf{x}(t)$. Plugging (26) into (29) yields

$$\hat{\mathbf{x}}_{LS}(t) = \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \times \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)(\mathbf{\Xi}_\tau(t)\boldsymbol{\delta}_\tau(t) + \mathbf{n}_\tau(t)) + \mathbf{x}(t)$$

and thus it is straightforward to see that

$$\boldsymbol{\xi}_{LS}(t) = \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \times \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)(\mathbf{\Xi}_\tau(t)\boldsymbol{\delta}_\tau(t) + \mathbf{n}_\tau(t)). \tag{30}$$

Different from (16), equation (30) does not represent a dynamical system but rather a static map. Hence, the estimation error can be characterized much more easily.

Proposition 2: Let Assumption 1, 2, and 3 hold. At every time t, the error mean $\boldsymbol{\mu}_{LS}(t) \coloneqq \mathbb{E}[\boldsymbol{\xi}_{LS}(t)]$ and the error variance $\mathbf{\Sigma}_{LS}(t) \coloneqq \mathbb{E}[(\boldsymbol{\xi}_{LS}(t) - \boldsymbol{\mu}_{LS}(t))(\boldsymbol{\xi}_{LS}(t) - \boldsymbol{\mu}_{LS}(t))^\top]$ are:

$$\boldsymbol{\mu}_{LS}(t) = \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \times \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Xi}_\tau(t)\boldsymbol{\delta}_\tau(t) \tag{31}$$

$$\mathbf{\Sigma}_{LS}(t) = \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \times \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{N}_\tau(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)^\top(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau \times \Big(\Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1}. \tag{32}$$

Proof: Being $\mathbf{n}(t)$ a random vector with zero mean

$$\begin{aligned}\mathbb{E}\big[\boldsymbol{\xi}_{LS}(t)\big] &= \mathbb{E}\Big[\Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \times \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)(\mathbf{\Xi}_\tau(t)\boldsymbol{\delta}_\tau(t) + \mathbf{n}_\tau(t))\Big] \\ &= \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \times \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)(\mathbf{\Xi}_\tau(t)\boldsymbol{\delta}_\tau(t) + \mathbb{E}[\mathbf{n}_\tau(t)]) \\ &= \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \times \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Xi}_\tau(t)\boldsymbol{\delta}_\tau(t).\end{aligned}$$

Moreover, observe that

$$\boldsymbol{\xi}_{LS}(t) - \boldsymbol{\mu}_{LS}(t) = \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \times \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{n}_\tau(t).$$

Hence, (32) follows from

$$\begin{aligned}\mathbf{\Sigma}_{LS}(t) &= \mathbb{E}\Big[(\boldsymbol{\xi}_{LS}(t) - \boldsymbol{\mu}_{LS}(t))(\boldsymbol{\xi}_{LS}(t) - \boldsymbol{\mu}_{LS}(t))^\top\Big] \\ &= \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \times \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbb{E}\Big[\mathbf{n}_\tau(t)\mathbf{n}_\tau(t)^\top\Big]\mathbf{\Pi}(t)^\top \times \mathbf{N}_\tau^{-1}(t)^\top(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau \times \Big(\Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \\ &= \Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1} \mathbf{\Phi}_\tau^\top(t)\mathbf{\Pi}(t)^\top \times \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{N}_\tau(t)\mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)^\top(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau \times \Big(\Big(\mathbf{\Phi}_\tau(t)^\top \mathbf{\Pi}(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Pi}(t)\mathbf{\Phi}_\tau(t)\Big)^{-1}. \quad \blacksquare\end{aligned}$$

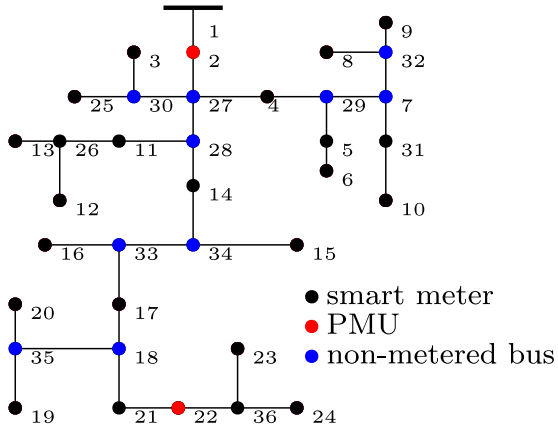

Fig. 2. The IEEE 37-bus feeder.

Roughly speaking, Proposition 2 states that the mean of the estimation error depends on both the state variability and the measurement noise, whereas its covariance depends mostly on the measurement noise. Moreover, since

$$\begin{aligned}\mathbb{E}\Big[\boldsymbol{\xi}^\top(t)\boldsymbol{\xi}(t)\Big] &= \boldsymbol{\mu}(t)^\top\boldsymbol{\mu}(t) + \mathbb{E}\Big[(\boldsymbol{\xi}(t) - \boldsymbol{\mu}(t))^\top(\boldsymbol{\xi}(t) - \boldsymbol{\mu}(t))\Big] \\ &= \boldsymbol{\mu}(t)^\top\boldsymbol{\mu}(t) + \mathbb{E}\Big[\text{Tr}\Big((\boldsymbol{\xi}(t) - \boldsymbol{\mu}(t))(\boldsymbol{\xi}(t) - \boldsymbol{\mu}(t))^\top\Big)\Big] \\ &= \|\boldsymbol{\mu}(t)\|^2 + \|\mathbf{\Sigma}(t)\|_F^2.\end{aligned} \tag{33}$$

the average distance between the estimate $\hat{\mathbf{x}}_{LS}(t)$ and $\mathbf{x}(t)$ is finite for every time t.

V. Numerical Evaluation

In this section, the OASE is used to solve the problem of state estimation on the three phase distribution power system shown in Figure 2, namely, the IEEE 37 bus test feeder [29]. There are three different type of buses: nonmetered buses, nodes endowed with smart meters (providing measurements of active power, reactive power, and voltage magnitude), and nodes endowed with PMUs (providing measurement of active power, reactive power, voltage magnitude and voltage angle).

The nodal power injections are obtained from real-world active load data measurements collected in the UMass Trace Repository Smart Dataset [30] 2017 Release. It contains the time-series load data of 114 apartments during the year-long time period from December 16, 2015, through December 14, 2016, at one minute granularity. Since the data set does not provide reactive loads, power factors were chosen for each bus to match the ratio between the nominal active and reactive powers provided in the testbed data sheet [29]. At each bus, we randomly aggregate the active power consumption of ten apartments and normalize the obtained value so that the average absolute power injection matches the default settings of the test feeder. Sensors are affected by Gaussian zero-mean measurement noise and have different reporting rates.

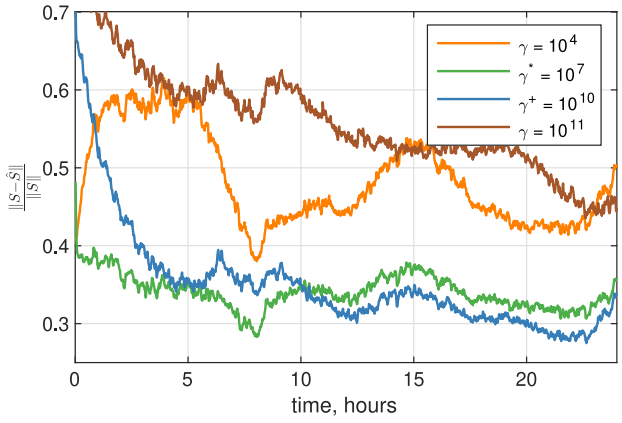

Fig. 3. Average power injection estimation error for the OASE under various inertia parameter settings.

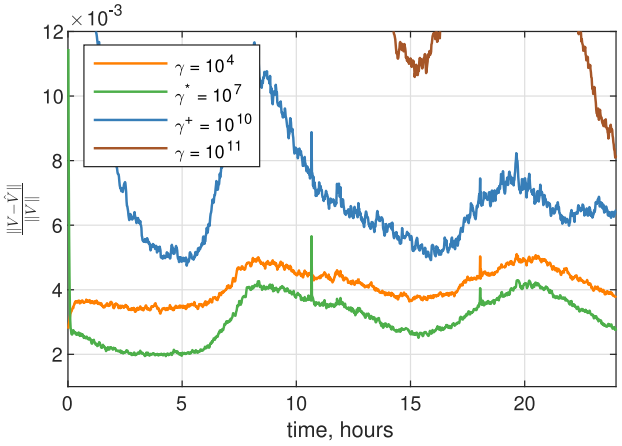

Fig. 4. Average voltage estimation error of the OASE under various inertia parameter settings.

- Smart meters report measurements once every hour asynchronously and introduce noise that is modeled as a zero mean Gaussian random variable with a relative standard deviation σ_{SM} and truncated outside $[-3\sigma_{\text{SM}}, 3\sigma_{\text{SM}}]$ to model a maximum error of 0.5% [31];
- PMUs report measurements every minute and introduce noise that is modeled as a zero mean Gaussian random variable with a relative standard deviation σ_{PMU} and truncated outside $[-3\sigma_{\text{PMU}}, 3\sigma_{\text{PMU}}]$ to model a maximum error of 0.05% [23].

Sensors provide $\mathbf{v}$ and $|\mathbf{v}|$ and the system operator computes the voltage deviations $\tilde{\mathbf{v}}$ and $|\tilde{\mathbf{v}}|$ that are used for estimating the state using (3) and (4). Since every sensor reports its measurement at least once every hour and the state estimation is performed every minute, we have $\tau \leq 60$. The OASE is compared with the WLS estimation algorithm (29). At any time t, the WLS algorithm uses both measurements from the reporting meters and the last retrieved measurements from the non-reporting sensors. The state estimation algorithms are tested on 365 Monte Carlo simulations, one for each day from the yearlong apartment load data set. Each Monte Carlo run randomly chooses which minute of the hour any particular smart meter reports its measurements. Although our approach relies on the approximate grid model of (5), voltages were calculated using the full ac grid model throughout our tests.

A. Algorithms' Performance Evaluation

First, the OASE's performance for different values of γ is studied. Figure 3 shows the average state estimation error among the Monte Carlo simulations over time under various settings of the inertia parameter γ. Given a state estimate, a voltage estimate can be found by computing the power flow equations. Figure 4 reports the average error on the computed voltages. Here, γ^* minimizes the right-hand side of equation (22), whereas γ^+, experimentally found, is the value that gives on average the minimum estimation error. Figure 3 shows that setting the inertia parameter to values that are much lower, e.g., 10^4, or much higher, e.g., 10^{10}, than γ^* provides worse performance: the estimation error settles around higher values or converges slowly. Nevertheless, for inertia parameters belonging to a wide interval centered around γ^*, the estimation error remains relatively close to the one associated with γ^+. Hence, the inertia parameter can be safely set to γ^*, even if this does not represent the optimal choice, without significantly deteriorating the estimator performance.

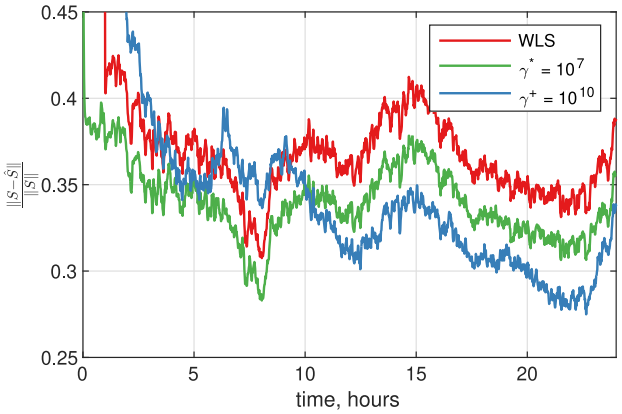

Fig. 5. Average power injection estimation error for the WLS estimator and the OASE.

Second, we compared the OASE with the WLS algorithm. Figure 5 reports the average relative estimation error in the power injections among the Monte Carlo runs; Figure 6 reports the average relative error in voltage estimation over time. For the first 60 minutes, we used the pseudoinverse $(\mathbf{\Phi}_\tau(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Phi}_\tau(t))^\#$ rather than the inverse $(\mathbf{\Phi}_\tau(t)^\top \mathbf{N}_\tau^{-1}(t)\mathbf{\Phi}_\tau(t))^{-1}$ because not enough measurement are available to make problem (23) strongly convex, making the WLS very inaccurate. Notably, the online state estimation algorithm, when the inertia parameter is set to γ^* or γ^+, outperforms the WLS estimator. We inspected the norm of the estimation error's mean and covariance, which are reported in Figure 7 and Figure 8, because the expectation of the estimator error squared norm can be related to these quantities, see (33). The results show that although WLS has almost the same average estimation error as the OASE under γ^+ or γ^*, it has a greater magnitude of the error covariance. This formally explains the results shown in Figures 5 and 6. Figure 9 shows estimates of the real power injection for phase a at bus 36. In general, the state estimate's volatility depends on either being too sensitive to the measurement error or being not very

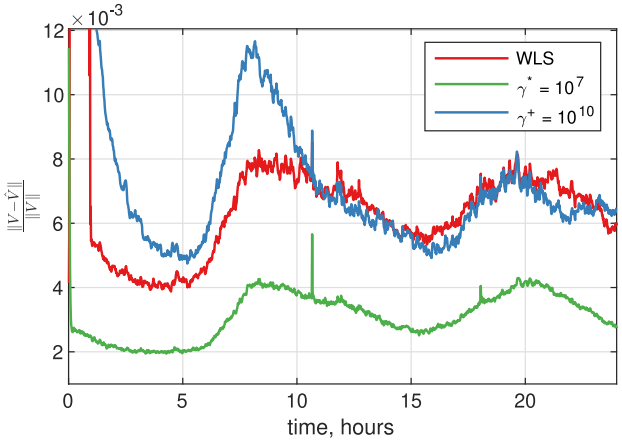

Fig. 6. Average voltage estimation error of WLS estimator and the OASE.

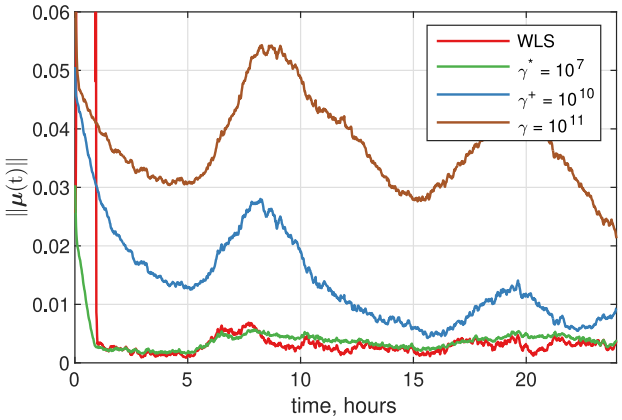

Fig. 7. Mean of the state estimation error for the WLS estimator and for the OASE.

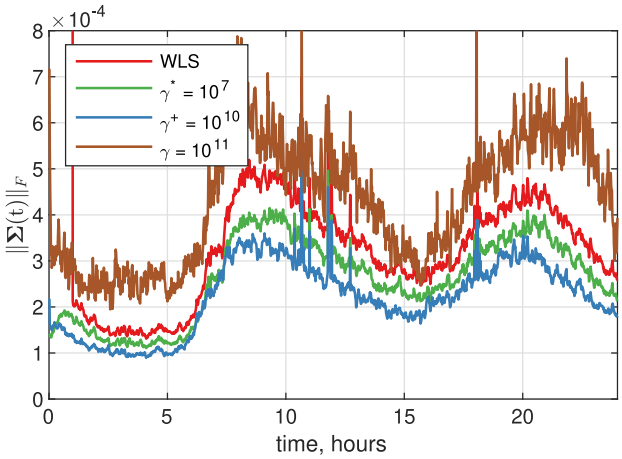

Fig. 8. Frobenius norm of the Covariance of the state estimation of the WLS estimator and of the OASE.

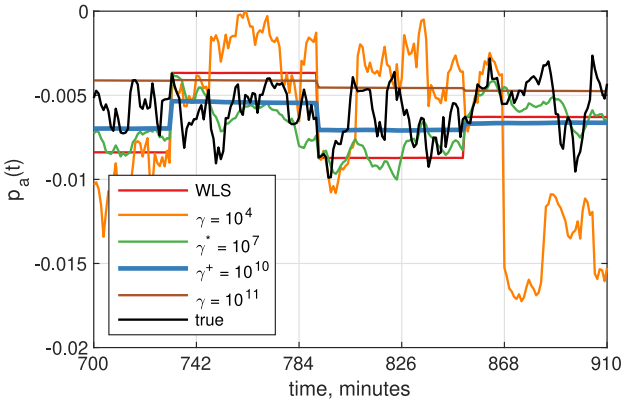

Fig. 9. Real power injection estimation over time for one phase at a particular bus.

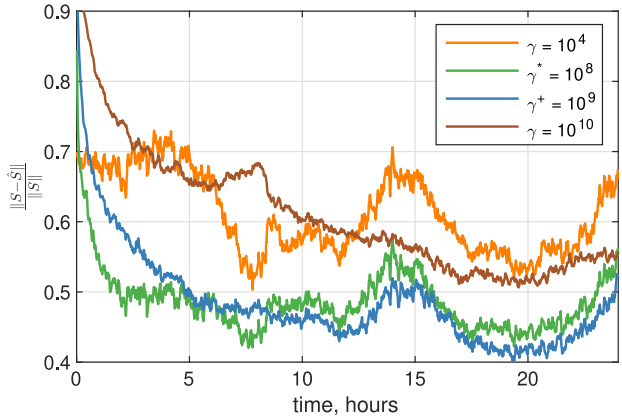

Fig. 10. Average power injection estimation error for the OASE under various inertia parameter settings and a different PMU placement.

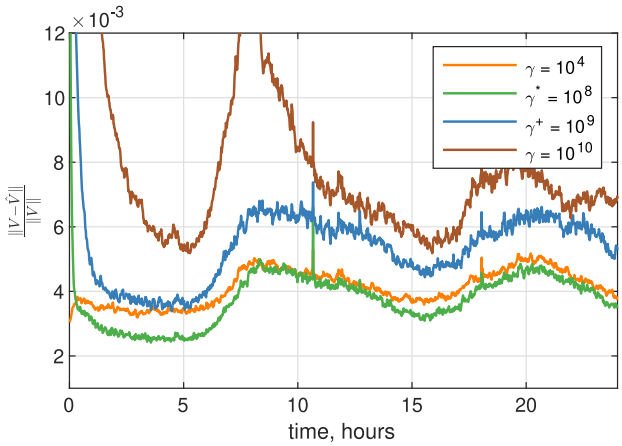

Fig. 11. Average voltage estimation error of the OASE under various inertia parameter settings and a different PMU placement.

responsive in tracking the true state when γ is either smaller or γ is bigger than the optimal parameter, respectively. It can be seen that γ^+ and γ^* track the true value reasonably well. When γ is too low ($\gamma = 10^4$), the state estimate fluctuates because of its sensitivity to measurements.

B. Effect of Sensor Placement

We tested how the placement of frequently reporting sensors (in our case, the PMUs) affects the performance of the OASE. In particular, we moved one PMU from Node 2 to Node 5. Figures 10 and 11 show the average state estimation error and the average error on the computed voltages, respectively. It is possible to notice that the new placement produced the following changes: first, γ^* and γ^+ are different from the one obtained before; values of the inertia parameter that were previously working well (e.g., $\gamma = 10^{10}$) are not providing good performance in the new setup. Second, the PMU placement affects the quality of the state estimates and of the voltages computed from them. Indeed, by comparing the curves in Figures 3 and 10 to those in Figures 4 and 11,

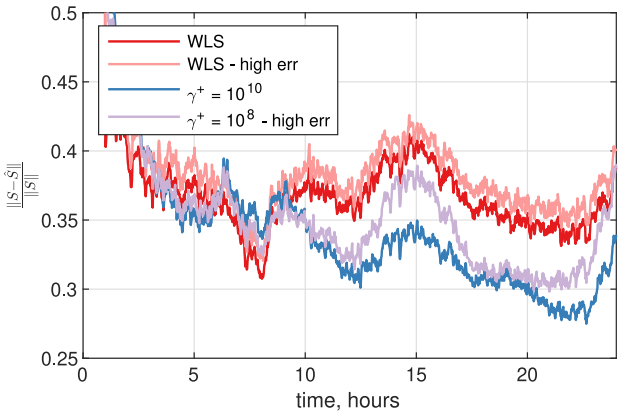

Fig. 12. Average power injection estimation error for the WLS estimator and the OASE under bad and noisy data.

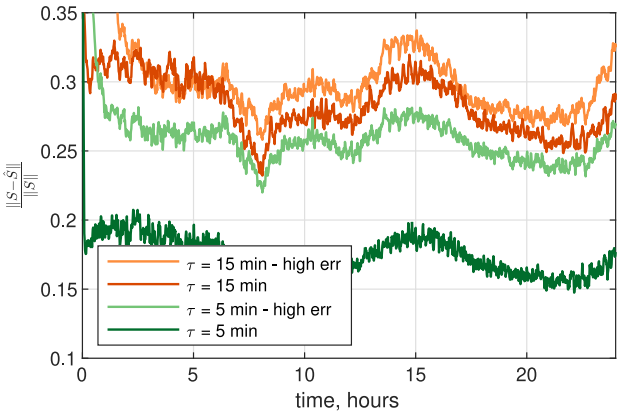

Fig. 13. Average power injection estimation error for the OASE under different smart meter measuring periods under bad and noisy data.

we can notice that the former PMUs placement gives better performance. Third, we notice that larger errors in the estimation of the power injections turn into smaller errors when looking at the associated voltage estimates; moreover, the ordering of the curves changes between the two scenarios, e.g., if the brown curve is the worst in both Figures 3 and 4, this is not the case anymore for Figures 10 and 11. This is because the computation of voltages entails solving highly nonlinear power flow equations; as experimentally shown, this is sensitive to the PMU placement. The study and the design of efficient algorithms for placing PMUs to minimize the estimation error and improve the quality of voltage estimates is an interesting direction that we will pursue in our future research efforts.

C. Algorithms' Robustness Analysis

We numerically analyzed the robustness of our algorithm by considering a scenario with increased measurement errors; this scenario is denoted hereafter as the *high err* case. In the *high err* case, the measurement noise affecting the sensors is larger than the one considered in the previous simulations; precisely, smart meters (PMUs) introduce a maximum error of 5% (0.5%). Moreover, sensors report corrupted data with a probability of 1% (i.e., the value zero for powers and voltages). Figure 12 compares the estimation errors in the *high err* case with the ones obtained before. The best results for the OASE were obtained by setting γ to 10^{10} rather than to γ^* or γ^+. Intuitively, having a large γ forces the new estimate to be close to the previous one, alleviating the effect of bad data on the estimation process. Increasing the measurement noise deteriorates the overall performance. Interestingly, the OASE outperforms even the WLS estimator using more precise measurements.

Further, we numerically studied the influence that the parameter τ has on the performance of the OASE. It is not a surprise that smaller values of τ lead to smaller errors. Intuitively, the estimation error depends both on the measurement noise and on the number of available measurements. Notably, the plots in Figure 13 show that smaller τ are associated with wider gaps between the errors in the *high err* scenario and the one considered before, i.e., when we consider the *high err* case, smaller values of τ are associated with bigger performance degradation. This shows that the estimation error mostly depend on (i) the scarcity of measurements when τ is big; (ii) the sensors' inaccuracy when τ is small.

VI. Conclusion

We have considered the problem of state estimation in a three-phase distribution network with asynchronized sensors. Two possible solutions have been studied. The first comes from our recent work [18] and it is a dynamic state estimator that has a recursive expression in which the new estimate is found as a function of the previous estimate, the gathered measurements, and of the inertia parameter. The second consists of a classic WLS estimator in which, as pseudo measurements, we used the measurement data retrieved in a certain finite time window. Both methods are able to estimate the true system state up to an error that depends on the state variability and the measurement error. Simulations on the standard IEEE-37 bus testbed with real-world load data traces have shown the effectiveness of the proposed strategies and that the dynamic estimator, for wise choices of the inertia parameter, outperforms the WLS estimator. Future research directions include considering a nonlinear measurement model instead of (5), and optimizing the placement of frequently reporting sensors (e.g., PMUs) to increase the estimation accuracy.

Acknowledgment

The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

References

[1] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, Sep. 2017.

[2] F. F. Wu and A. Monticelli, "Network observability: Theory," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 5, pp. 1042–1048, May 1985.

[3] M. E. Baran, "Challenges in state estimation on distribution systems," in *Proc. Power Eng. Soc. Summer Meeting Conf.*, vol. 1, 2001, pp. 429–433.
[4] B. Zhao, M. Ye, L. Stankovic, and V. Stankovic, "Non-intrusive load disaggregation solutions for very low-rate smart meter data," *Appl. Energy*, vol. 268, Jun. 2020, Art. no. 114949.
[5] L. Vanfretti, M. Baudette, and A. D. White, "Monitoring and control of renewable energy sources using synchronized phasor measurements," in *Renewable Energy Integration*, L. E. Jones, Ed., 2nd ed. Boston, MA, USA: Academic, 2017, pp. 419–434. [Online]. Available: https://doi.org/10.1016/B978-0-12-809592-8.00031-7
[6] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich, and E. Vaahedi, "Distribution system state estimation based on nonsynchronized smart meters," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2919–2928, Nov. 2015.
[7] G. Cavraro, E. Dall'Anese, and A. Bernstein, "Dynamic power network state estimation with asynchronous measurements," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Ottawa, ON, Canada, Nov. 2019, pp. 1–5.
[8] W. Luan, D. Sharp, and S. LaRoy, "Data traffic analysis of utility smart metering network," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2013, pp. 1–4.
[9] B. Karimi, V. Namboodiri, and M. Jadliwala, "Scalable meter data collection in smart grids through message concatenation," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1697–1706, Jul. 2015.
[10] A. K. Ghosh, D. L. Lubkeman, and R. H. Jones, "Load modeling for distribution circuit state estimation," *IEEE Trans. Power Del.*, vol. 12, no. 2, pp. 999–1005, Apr. 1997.
[11] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1888–1896, Nov. 2012.
[12] A. S. Zamzam and N. D. Sidiropoulos, "Physics-aware neural networks for distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4347–4356, Nov. 2020.
[13] L. Schenato, G. Barchi, D. Macii, R. Arghandeh, K. Poolla, and A. Von Meier, "Bayesian linear state estimation using smart meters and PMUs measurements in distribution grids," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2014, pp. 572–577.
[14] C. Carquex, C. Rosenberg, and K. Bhattacharya, "State estimation in power distribution systems based on ensemble Kalman filtering," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6600–6610, Nov. 2018.
[15] J. Zhao, M. Netto, and L. Mili, "A robust iterated extended Kalman filter for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205–3216, Jul. 2017.
[16] P. L. Donti, Y. Liu, A. J. Schmitt, A. Bernstein, R. Yang, and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2520–2530, May 2020.
[17] S. Bhela, V. Kekatos, and S. Veeramachaneni, "Smart inverter grid probing for learning loads: Part I—Identifiability analysis," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3527–3536, Sep. 2019.
[18] G. Cavraro, E. Dall'Anese, J. Comden, and A. Bernstein, "Online state estimation for time-varying systems," *IEEE Trans. Autom. Control*, early access, Oct. 15, 2021, doi: 10.1109/TAC.2021.3120679.
[19] J. Song, E. Dall'Anese, A. Simonetto, and H. Zhu, "Dynamic distribution state estimation using synchrophasor data," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 821–831, Jan. 2020.
[20] M. E. Baran and A. W. Kelley, "State estimation for real-time monitoring of distribution systems," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1601–1609, Aug. 1994.
[21] A. Bernstein and E. Dall'Anese, "Linear power-flow models in multiphase distribution networks," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Europe (ISGT-Europe)*, 2017, pp. 1–6.
[22] J. Zheng, D. W. Gao, and L. Lin, "Smart meters in smart grid: An overview," in *Proc. IEEE Green Technol. Conf. (GreenTech)*, Apr. 2013, pp. 57–64.
[23] A. von Meier, D. Culler, A. McEachern, and R. Arghandeh, "Micro-synchrophasors for distribution systems," in *Proc. IEEE Conf. Innov. Smart Grid Technol.*, Washington, DC, USA, Feb. 2014, pp. 1–5.
[24] A. Angioni, C. Muscas, S. Sulis, F. Ponci, and A. Monti, "Impact of heterogeneous measurements in the state estimation of unbalanced distribution networks," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, May 2013, pp. 935–939.
[25] N. Andreadou, E. Kotsakis, and M. Masera, "Smart meter traffic in a real LV distribution network," *Energies*, vol. 11, no. 5, p. 1156, 2018.
[26] G. Wang, G. B. Giannakis, and J. Chen, "Robust and scalable power system state estimation via composite optimization," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6137–6147, Nov. 2019.
[27] K. A. Clements, "The impact of pseudo-measurements on state estimator accuracy," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2011, pp. 1–4.
[28] A. Angioni, T. Schlösser, F. Ponci, and A. Monti, "Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 70–77, Jan. 2016.
[29] W. H. Kersting, *Distribution System Modeling and Analysis*. New York, NY, USA: CRC Press, 2001.
[30] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, J. Albrecht, "Smart*: An open data set and tools for enabling research in sustainable homes," in *Proc. SustKDD*, vol. 111, 2012, pp. 1–6.
[31] "Smart meters and smart meter systems: A metering industry perspective," Edison Elect. Inst., Washington, DC, USA, White Paper, 2011.

Guido Cavraro (Associate Member, IEEE) received the Ph.D. degree in information engineering from the University of Padova, Italy, in 2015. He was a Visiting Scholar with the California Institute for Energy and Environment, U.C. Berkeley in 2014. From 2015 to 2016, he was a Postdoctoral Associate with the Department of Information Engineering, University of Padova. From 2016 to 2018, he was a Postdoctoral Associate with the Bradley Department of Electrical and Computer Engineering, Virginia Tech, USA. He is currently a Senior Researcher with the Power Systems Engineering Center with National Renewable Energy Laboratory, USA. His research interests include control, optimization, and identification applied to power systems and smart grids.

Joshua Comden received the B.E. degree in chemical engineering from the University of Delaware, Newark, DE, USA, in 2009, and the M.S. and Ph.D. degrees in operations research with the Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, USA, in 2015 and 2019, respectively. He is a Postdoctoral Researcher with the National Renewable Energy Laboratory, Golden, CO, USA. His research interests include the design and analysis of online optimization and control algorithms for power systems. He won the Woo Jong Kim Dissertation Award in 2020.

Emiliano Dall'Anese received the Ph.D. degree in information engineering from the Department of Information Engineering, University of Padova, Italy, in 2011. He is an Assistant Professor with the Department of Electrical, Computer, and Energy Engineering, University of Colorado Boulder, and an Affiliate Faculty with the Department of Applied Mathematics. From January 2011 to November 2014, he was a Postdoctoral Associate with the Department of Electrical and Computer Engineering, University of Minnesota, and from December 2014 to July 2018 he was a Senior Researcher with the National Renewable Energy Laboratory. His research interests span the areas of optimization, control, and learning. He received the National Science Foundation CAREER Award in 2020 and the IEEE PES Prize Paper Award in 2021.

Andrey Bernstein received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology. From 2010 to 2011, he was a Visiting Researcher with Columbia University. From 2011 to 2012, he was a Visiting Assistant Professor with Stony Brook University. From 2013 to 2016, he was a Postdoctoral Researcher with the Laboratory for Communications and Applications of Ecole Polytechnique Federale de Lausanne, Switzerland. Since October 2016, he has been a Senior Researcher and Group Manager with National Renewable Energy Laboratory, Golden, CO, USA. His research interests are in the decision and control problems in complex environments and related optimization and machine learning methods, with application to power and energy systems.