State Estimation in Electric Power Grids: Meeting New Challenges Presented by the Requirements of the Future Grid

Article *in* IEEE Signal Processing Magazine · September 2012

DOI: 10.1109/MSP.2012.2187037

5 authors, including:


Yih-Fang Huang
University of Notre Dame


Stefan Werner
Norwegian University of Science and Technology


Neelabh Kashyap
Aalto University


Vijay Gupta
University of Notre Dame

Yih-Fang Huang, Stefan Werner, Jing Huang,
Neelabh Kashyap, and Vijay Gupta

State Estimation in Electric Power Grids

Meeting new challenges presented by the requirements of the future grid

ISTOCKPHOTO.COM/© SIGAL SUHLER MORAN

This article provides a survey on state estimation (SE) in electric power grids and examines the impact on SE of the technological changes being proposed as a part of the smart grid development. Although SE at the transmission level has a long history, further research and development of innovative SE schemes, including those for distribution systems, are needed to meet the new challenges presented by the requirements of the future grid. This article also presents some example topics that signal processing (SP) research can contribute to help meet those challenges.

INTRODUCTION

Since the pioneering work of F.C. Schweppe in 1970 [1], SE has become a key function in supervisory control and planning of electric power grids. It serves to monitor the state of the grid and enables energy management systems (EMS) to perform various important control and planning tasks such as establishing near real-time network models for the grid, optimizing power flows, and bad data detection/analysis (see, e.g., [2] and [3] and the references therein). Another example of the utility of SE is the SE-based reliability/security assessment deployed to analyze contingencies and determine necessary corrective actions against possible failures in the power systems.

In view of the ongoing development of a smarter grid, more research on SE is needed to meet the challenges that the envisioned smart grid functionalities present. Among others, environmental compliance, energy conservation, and improved dependability, reliability, and security will impose additional constraints on SE and require improved performance in terms of response time and robustness [4]. In this article, we provide a brief survey of some SE technologies developed over the last four decades and examine the challenges and opportunities presented by the evolution of the legacy grid into a smarter grid, within a framework relevant to SP research.

Digital Object Identifier 10.1109/MSP.2012.2187037
Date of publication: 20 August 2012

There are at least three major aspects in the future power grid that will directly impact SE research. First, more advanced measurement technologies like phasor measurement units (PMUs) have offered hope for near real-time monitoring of the power grid; see, e.g., [5]. Typically, a PMU takes 30 measurements/s, thereby offering the possibility of a much more timely view of the power system dynamics than conventional measurements. More importantly, all PMU measurements are synchronized, as they are time-stamped by the global positioning system's (GPS's) universal clock. However, PMUs with their higher measurement frequency put enormous strain on the communication and data processing infrastructure of the grid. This drives the need for resource-efficient, event-triggered SE solutions that employ on-demand (event-triggered) sensing, estimation, and communication.

Second, new regulations and market pricing competition may require utility companies to share more information and monitor the grid over large geographical areas. This calls for distributed control, and hence, distributed SE to facilitate interconnection-wide coordinated monitoring [6]. Recent advances made by the SP and automatic control communities in the field of distributed estimation would be particularly beneficial in achieving this.

Finally, to facilitate smart grid features such as demand response (DR) and two-way power flow, utility companies will need to have more timely and accurate models for their distribution systems. This calls for SE at the distribution level, which places more stringent requirements on SE algorithms. So far, utility companies have done little in implementing SE in distribution systems, even though SE has been deployed extensively in transmission systems for decades. However, as the electric power grid becomes smarter, more distribution automation (DA) will be needed and SE at the distribution level will become more important. The control mechanism in the distribution system will most likely be distributed and active in nature, so will be the corresponding SE functions. This necessitates the development of new distributed SE algorithms that avail themselves of the substantially increased number of real-time measurements.

Discussions in this article are, thus, motivated by these three aspects, and are organized in a way that is compatible with the hierarchy of the power grid, particularly, the transmission level, the subtransmission level, and the distribution level (see Figure 1). We envision that SE in the future grid would likely be carried out at different levels, specifically, the transmission system operator (TSO) level, the local level or subtransmission level, and the distribution level; see, e.g., the multilevel SE paradigm presented in [7]. The TSO is an entity that operates the transmission grid to supply electricity from the generating companies (GENCOs) [8] to the utility companies and then to the consumer. Substations are a vital link between the transmission and distribution networks and are responsible for converting voltage and current levels. The trend of deregulation of vertically integrated utilities, particularly in the United States, would mean that market forces would play an increasing role in the future grid.

[STATE ESTIMATION HAS BECOME A KEY FUNCTION IN SUPERVISORY CONTROL AND PLANNING OF ELECTRIC POWER GRIDS.]

EVOLUTION OF STATE ESTIMATION

The state of a power system can be described by the voltage magnitudes and phase angles at every bus. This information, along with the knowledge of the topology and impedance parameters of the grid, can be used to characterize the entire system. The EMS/supervisory control and data acquisition (SCADA) system is a set of computational tools used to monitor, control, and optimize the performance of a power system. SE is a vital component here; the relationship between SE and the SCADA system is shown in Figure 2. The data acquisition system obtains measurement from devices like remote terminal units (RTUs) and, more recently, phasor data concentrators (PDCs). The state estimator calculates the system state and provides the necessary information to the supervisory control system, which then takes action by sending control signals to the switchgear (circuit breakers).

The conventional state estimator built into the EMS consists of four main processes as shown in Figure 2. The topology processor tracks the network topology and maintains a real-time database of the network model. Observability analysis is a process that is run to ensure the measurement set is sufficient to perform SE. Next, the bad-data processor identifies any gross errors in the measurement set and eliminates bad measurements. The state estimator operates on the set of good measurements to calculate the system state. Finally, the bad-data processing identifies any gross errors in the measurement set and eliminates the bad measurements.

Depending on the timing and evolution of the estimates, SE schemes may be classified into two basic distinct paradigms: static SE (SSE) and forecasting-aided SE (FASE). We will provide a brief overview of the formulation, development, and evolution of those two SE paradigms. Additionally, we will discuss multiarea SE (MASE), which may become a fruitful area of research as the distributed approach is showing more promises for the future grid.

STATIC STATE ESTIMATION

For the last four decades, much of the research on SE has been focused on SSE, primarily due to the fact that the traditional monitoring technologies, such as those implemented in the SCADA system, can only take nonsynchronized measurements once every two to four seconds. Furthermore, to reduce the computational complexity required in implementing SE, the estimates are usually updated only once every few minutes. Hence, the usefulness of SSE as a means to provide real-time monitoring of the power grid is quite limited in practice.

In an N-bus system, the $(2N-1) \times 1$ state vector has the form $\mathbf{x} = [\theta_2, \theta_3, \ldots, \theta_N, |V_1|, \ldots, |V_N|]^{\mathrm{T}}$ where θ_i denote the phase angles and $|V_i|$ the magnitudes of the voltages at the ith bus. The phase angle θ_1 at the reference bus is assumed known and is normally set to zero radians. To estimate the state $\mathbf{x}$, a set of measurements $\mathbf{z} \in \mathbb{R}^{L \times 1}$, $L > 2N-1$, is collected. These measurements consist of nonsynchronized active and reactive power flows in network elements, bus injections and voltage magnitudes at the buses. The measurements are typically obtained within SCADA systems, and are related to the state vector by an overdetermined system of nonlinear equations, specifically,

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{n}, \tag{1}$$

where $\mathbf{h}(\cdot)$ is a set of L nonlinear functions of the state vector (determined by Kirchhoff's laws and the power network admittance matrix) and $\mathbf{n}$ is a zero-mean Gaussian measurement noise vector with covariance matrix $\mathbf{C}_n \in \mathbb{R}^{L \times L}$.

In the traditional SSE approach, the state vector is estimated from the measurement equation in (1) using the weighted least-squares (WLS) method; see, e.g., [1]. In particular, the SSE problem is solved by finding

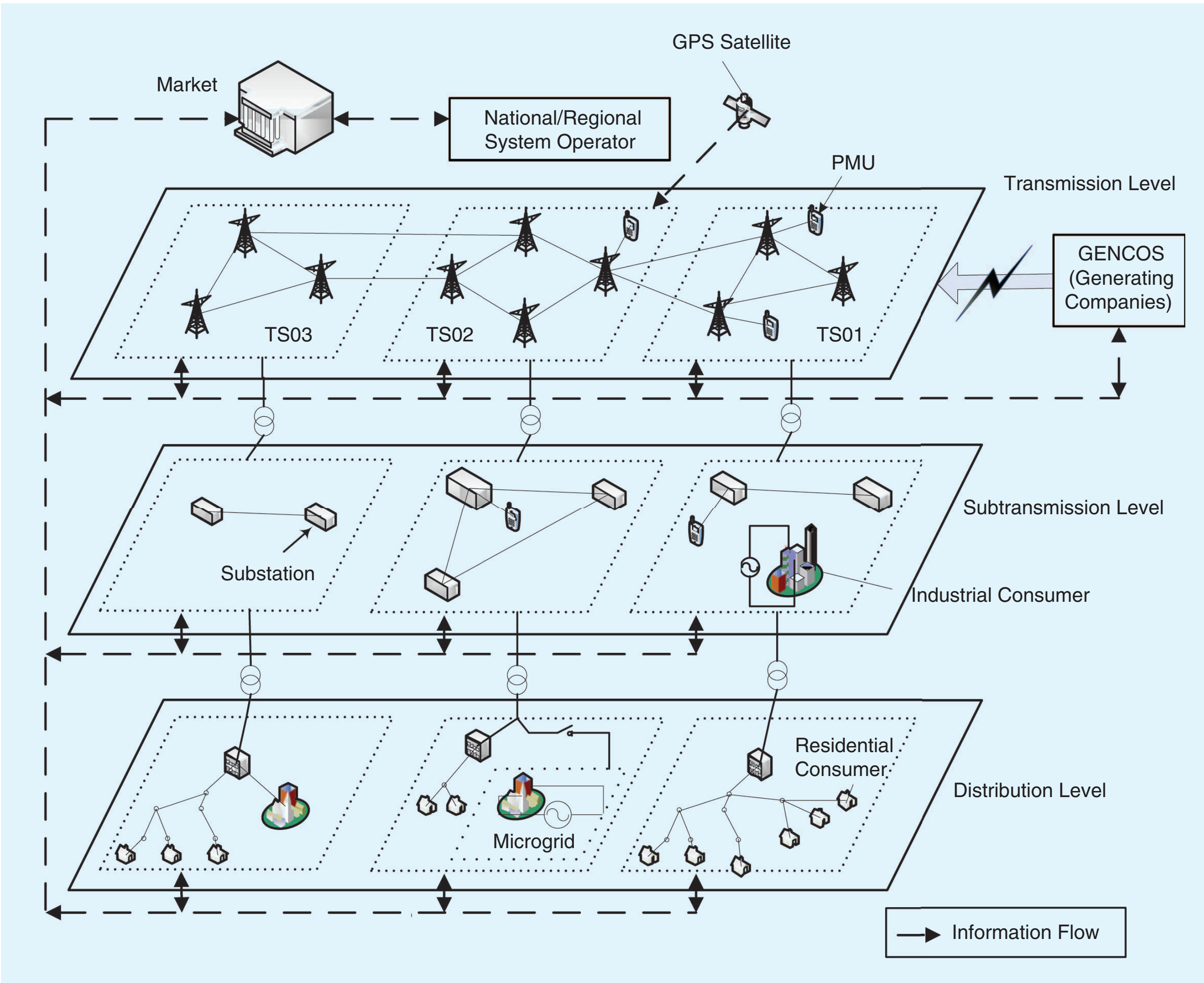
$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} [\mathbf{z} - \mathbf{h}(\mathbf{x})]^{\mathrm{T}} \mathbf{W}^{-1} [\mathbf{z} - \mathbf{h}(\mathbf{x})], \tag{2}$$

where weighting matrix $\mathbf{W}$ is commonly taken as diagonal with elements related to background noise covariance as $\mathbf{W} = \mathbf{C}_n$. The solution for $\hat{\mathbf{x}}$ is obtained in an iterative fashion by linearizing (1) around the available estimate (at iteration j) and applying the Gauss-Newton algorithm to improve the estimate, using the following equations:

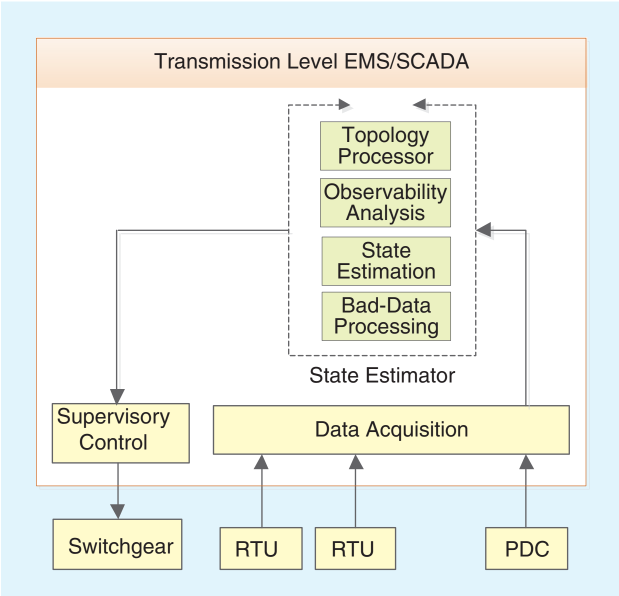
$$\mathbf{G}(j)\Delta\mathbf{x}(j) = \mathbf{H}^{\mathrm{T}}(j)\mathbf{W}^{-1}[\mathbf{z} - \mathbf{h}(\mathbf{x}(j))] \tag{3}$$

$$\hat{\mathbf{x}}(j+1) = \hat{\mathbf{x}}(j) + \Delta\mathbf{x}(j), \tag{4}$$

where $\mathbf{G}(j) = \mathbf{H}^{\mathrm{T}}(j)\mathbf{W}^{-1}\mathbf{H}(j)$ is the gain matrix at iteration j. Equation (3) is usually referred to as the normal equation. The



[FIG1] **Electricity ecosystem of the future grid featuring various players and levels of interaction.**



[FIG2] Relationship between different elements that collectively constitute the EMS/SCADA.

Jacobian matrix, $\mathbf{H}(j) \in \mathbb{R}^{L\times(2N-1)}$, needed at each iteration, is the first-order partial derivative of $\mathbf{h}(\mathbf{x})$, with respect to $\mathbf{x}$, evaluated at $\hat{\mathbf{x}}(j)$, i.e., $\mathbf{H}(j) = \{\partial \mathbf{h}(\mathbf{x})/\partial \mathbf{x}\}_{\mathbf{x}=\hat{\mathbf{x}}(j)}$. The iterative process is terminated when the norm of the residual falls below a predefined value, i.e., for some $\delta > 0$, $\|\mathbf{z} - \mathbf{h}(\hat{\mathbf{x}}(j))\|^2 \le \delta$, and the covariance matrix of the final estimate is given by $\mathbf{G}^{-1}(j) = [\mathbf{H}^{\mathrm{T}}(j)\mathbf{W}^{-1}\mathbf{H}(j)]^{-1}$.

One of the main problems in solving the normal equation in (3) is computational complexity. An approach to reduce this complexity is to realize that $\mathbf{G}(j)$ is sparse and symmetric, then implement various iterative solutions, e.g., Krylov subspace methods, to find $\Delta\mathbf{x}(j)$ in (3). A more common approach in the literature is to take advantage of the sparseness of matrix $\mathbf{H}(j)$, which is in general even more sparse than $\mathbf{G}(j)$, and employ a robust and computationally efficient QR factorization of the weighted Jacobian, e.g., using a sequence of Givens rotations (or Householder reflections) of the weighted Jacobian matrix $\mathbf{W}^{-1/2}\mathbf{H}(j)$. A good treatment of this topic and the related important references can be found in [3]. Specifically, let $\mathbf{Q}(j) \in \mathbb{R}^{L\times L}$ be an orthogonal transformation that triangularizes the weighted Jacobian as follows:

$$\mathbf{Q}(j)\mathbf{W}^{-1/2}\mathbf{H}(j) = \begin{bmatrix}\mathbf{R}(j)\\ \mathbf{0}\end{bmatrix}, \tag{5}$$

where $\mathbf{R}(j) \in \mathbb{R}^{(2N-1)\times(2N-1)}$ is upper/lower triangular. We may now rewrite (3) as

$$\begin{aligned}[\mathbf{R}^{\mathrm{T}}(j) \quad 0]\begin{bmatrix}\mathbf{R}(j)\\ 0\end{bmatrix}\Delta\mathbf{x}(j) \\ = [\mathbf{R}^{\mathrm{T}}(j) \quad 0]\,\mathrm{Q}(j)\mathrm{W}^{-1/2}[\mathrm{z} - \mathrm{h}(\mathrm{x}(j))],\end{aligned} \tag{6}$$

which allows us to solve for $\Delta\mathbf{x}(j)$ in two stages

$$\begin{bmatrix}\mathbf{y}_1(j)\\ \mathbf{y}_2(j)\end{bmatrix} = \mathbf{Q}(j)\mathbf{W}^{-1/2}[\mathbf{z} - \mathbf{h}(\mathbf{x}(j))] \tag{7}$$

$$\mathbf{R}(j)\Delta\mathbf{x}(j) = \mathbf{y}_1(j), \tag{8}$$

where $\mathbf{y}_1(j) \in \mathbb{R}^{(2N-1)\times 1}$, as seen in (7), is formed by taking the $2N-1$ first element of the transformed (weighted) measurement error vector. The correction term $\Delta\mathbf{x}(j)$ is obtained via backward (or forward) substitution.

Alternatively, to overcome the computational cost associated with directly solving (3), it has often been argued that the gain matrix $\mathbf{G}(j)$ does not change considerably during several iterations, which implies that we can assume a piecewise constant Jacobian matrix [1]. This observation is exploited in the hybrid method [9] to reduce storage requirements when applying orthogonal transformations. In particular, the triangular matrix $\mathbf{R}$ in (5) remains constant for those iterations when the measurement Jacobian is not reevaluated. Thus, by only transforming the left-hand side of (3) [cf. (6)] we may acquire the correction term $\Delta\mathbf{x}(j)$ from $\mathbf{R}^{\mathrm{T}}\mathbf{R}\Delta\mathbf{x}(j) = \mathbf{H}^{\mathrm{T}}\mathbf{W}^{-1}[\mathbf{z} - \mathbf{h}(\mathbf{x}(j))]$. Compared with (7)–(8), we see that only $\mathbf{R}$ needs to be stored (and not factors of $\mathbf{Q}$) at the expense of an additional forward (or backward) substitution.

The computational complexity of the aforementioned SE approaches may be further reduced by assuming voltage magnitudes and phases to be independent [3]. The state estimate is then obtained by solving two decoupled WLS problems since the measurement Jacobian becomes block-diagonal. This approach renders a particularly efficient implementation of the hybrid method both in terms of storage requirements and computational cost.

A more recent approach for reducing the computational cost is to use a nested, or multilevel, formulation of the nonlinear measurement model [7]. This approach can sustain growth in size, complexity, and data. It is designed to function at different levels of the modeling hierarchy to accomplish very large-scale interconnection-wide monitoring. This method uses the same overdetermined set of measurement equations as in (1). The equations are then "unfolded" into K sequential WLS problems by introducing a set of intermediate variables $\mathcal{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K\}$ with the following nested structure:

$$\begin{aligned}\mathbf{z} &= \mathbf{f}_1(\mathbf{y}_1) + \mathbf{n}\\ \mathbf{y}_1 &= \mathbf{f}_2(\mathbf{y}_2) + \mathbf{n}_1\\ &\vdots\\ \mathbf{y}_{K-1} &= \mathbf{f}_K(\mathbf{y}_K) + \mathbf{n}_{K-1}\\ \mathbf{y}_K &= \mathbf{f}_{K+1}(\mathbf{x}) + \mathbf{n}_K.\end{aligned} \tag{9}$$

The set $\mathcal{Y}$ is chosen such that the solution of the nested system of equations (9) offers some desired advantage over solving (1), e.g., reduction of the computational complexity or the amount of information exchanged between different levels. This is a particularly appealing solution when the measurement model can be factorized into separate linear and nonlinear parts, e.g., a hierarchical structure that comprises a linear substation model and a nonlinear transmission level model.

FORECASTING-AIDED STATE ESTIMATION

Conventional SSE relies on a single set of measurements all taken at one snapshot in time. Hence, it disregards the evolution of the state over consecutive measurement instants. The basic idea of FASE is to provide a recursive update of the state estimate that can also track the changes occurring during normal system operation. One of the advantages of FASE is that it includes by design a forecasting feature that can get around the problem of missing measurements, as the predicted states may be used in lieu of those measurements. Note, however, that FASE is somewhat different from true dynamic SE since the transients in power systems usually occur at a much faster time scale than those considered in FASE.

The first step toward a dynamic state estimator was taken by Debs and Larson in 1970 [10]. A simple state transition model was developed assuming the system was in a quasi steady-state. Tracking state estimators [11] came next, but the problem here was that no time evolution model was assumed explicitly to follow the dynamics of the system. The next breakthrough in FASE came from [12] that introduced a more appropriate state transition model and used Kalman filtering and an exponential smoothing algorithm for forecasting. A robust FASE algorithm based on M-estimation was presented in [13] as an alternative to the Kalman filter-based approaches and more recently, a FASE algorithm was proposed based on unscented Kalman filter (UKF) [14]. A more extensive literature survey and related references may be found in [15].

A typical FASE is formulated with the following dynamic model [12]:

$$\mathbf{x}(k+1) = \mathbf{F}(k)\mathbf{x}(k) + \mathbf{g}(k) + \mathbf{w}(k), \tag{10}$$

where for time instant k, $\mathbf{F}(k) \in \mathbb{R}^{(2N-1)\times(2N-1)}$ is the state-transition matrix, vector $\mathbf{g}(k)$ is associated with the trend behavior of the state-trajectory, and $\mathbf{w}(k)$ is assumed to be zero-mean Gaussian noise with covariance matrix $\mathbf{C}_w$.

Using (10) and the measurements arriving at instant $k+1$, $\mathbf{z}(k+1) = \mathbf{h}(\mathbf{x}(k+1)) + \mathbf{n}(k+1)$, the majority of the FASE algorithms that appear in the literature are based on the extended Kalman filter (EKF), whose recursions are given by

$$\hat{\mathbf{x}}(k+1) = \tilde{\mathbf{x}}(k+1) + \mathbf{K}(k+1)[\mathbf{z}(k+1) - \mathbf{h}(\tilde{\mathbf{x}}(k+1))], \tag{11}$$

where

$$\begin{aligned}
\tilde{\mathbf{x}}(k+1) &= \mathbf{F}(k)\hat{\mathbf{x}}(k) + \mathbf{g}(k)\\
\mathbf{K}(k+1) &= \Sigma(k+1)\mathbf{H}^{\mathrm{T}}(k+1)\mathbf{C}_n^{-1}\\
\Sigma(k+1) &= [\mathbf{H}^{\mathrm{T}}(k+1)\mathbf{C}_n^{-1}\mathbf{H}(k+1) + \mathbf{M}^{-1}(k+1)]^{-1}\\
\mathbf{M}(k+1) &= \mathbf{F}(k)\Sigma(k)\mathbf{F}^{\mathrm{T}}(k) + \mathbf{C}_w
\end{aligned}$$

with $\mathbf{H}(k+1)$ being the measurement Jacobian evaluated at $\tilde{\mathbf{x}}(k+1)$. We note that matrix $\mathbf{F}(k)$ and vector $\mathbf{g}(k)$ in (10) are usually updated recursively using the classic Holt-Winters method [12]. This rather naive state-transition model appears to work quite well, although it ignores any coupling between the state variables.

MULTIAREA STATE ESTIMATION

MASE traces its origins back to the late 1970s, when microprocessor technology was not mature enough to handle the computational load of SE in very large interconnections and SE was implemented on multiprocessor computing architectures. Since the power grid is inevitably a large network, a centralized solution to the associated SE problem poses tremendous computational complexity. An alternative is to divide the large power system into smaller areas, each equipped with a local processor to provide a local SE solution. As compared with a centralized SE approach, MASE reduces the amount of data that each state estimator needs to process (hence reduces complexity) and it improves the robustness of the system by distributing the knowledge of the state. However, its implementation requires additional communication overhead and it comes with the time-skewness problem that results from asynchronous measurements obtained in different areas.

In MASE, each area has local measurements formulated by

$$\mathbf{z}_m = \mathbf{h}_m(\mathbf{x}_m) + \mathbf{n}_m,\ \ m = 1, \ldots, M, \tag{12}$$

where $\mathbf{x}_m = [\mathbf{x}_{im}^{\mathrm{T}}\mathbf{x}_{bm}^{\mathrm{T}}]^{\mathrm{T}}$ is the local state vector of area m, which is further partitioned into internal state variables, $\mathbf{x}_{im}^{\mathrm{T}}$, and border state variables, $\mathbf{x}_{bm}^{\mathrm{T}}$. Internal variables are those state variables that are observable for the particular area while border variables are states of those buses with lines connecting two areas (so-called tie-lines).

A local estimate can be obtained from (12) using the techniques outlined above with the difference that the measurement Jacobian is derived from the local estimate. Taking into account the coupling between areas located in close proximity, improved state estimates can be obtained by combining local estimates using either a hierarchical structure, a decentralized structure, or a combination of both. In the hierarchical scheme, a central computer controls the local processors which may be either located in disparate geographical areas (distributed architecture) or in the same area (parallel architecture). The local state estimators communicate only with the central computer. In a fully decentralized architecture, there is no central computer, and each local state estimator communicates only with its neighbors. The amount of data exchange of the solution depends on whether local estimates (or measurements) are transmitted at every iteration of the local estimation algorithm or upon convergence. A survey of various MASE methods is given in [16] along with a good treatment of a two-level hierarchical MASE example. More recently, MASE was introduced in [17], which performs SE in a fully distributed manner.

IMPACT OF PHASOR MEASUREMENT UNITS ON STATE ESTIMATION

Conventional SCADA measurements are obtained too infrequently to fully capture the dynamics of the power system. Practically, when faults occur, there is usually little time for the controller to respond, and this presents a serious challenge to operators. Integration of renewable energy sources in distributed generation (DG) may also increase the chance of sudden unpredictable changes in the system. Consequently, it is necessary to track these changes in a timely manner to ensure the dependability and reliability of the power system. Thus, SE schemes that are capable of capturing and tracking the near real-time dynamics of the power system are needed.

Recently, synchronized phasor measurement units (PMUs) have been increasingly deployed in power systems. These devices can directly measure bus voltage magnitudes and phase angles, because they are synchronized by the GPS universal clock. The PMUs also sample at a much higher frequency (roughly two orders of magnitude faster) compared to the traditional sensors in the SCADA system. In essence, PMUs provide more accurate and more timely measurements with many more samples. The main challenges faced by engineers today include 1) combining those PMU measurements with conventional measurements to obtain an optimal state estimate, and 2) dealing with the large number of data rendered by PMUs. This section examines the impact that the synchrophasor technology has had on the three SE paradigms described in the previous section.

PHASOR MEASUREMENT UNITS

PMUs measure not only voltage phasors at buses where they are installed but also current phasors through all incident buses. Since the current phasor on a line between two buses is linearly related to the two voltage phasors at those two buses, the PMU measurements across the system, aggregated into vector $\mathbf{z}_2$, are linearly related to the voltage phasors, $\mathbf{v} = [\Re\{V_1\}, \ldots, \Re\{V_N\} \Im\{V_2\}, \ldots, \Im\{V_N\}]^{\mathrm{T}}$, where $\Re\{V\}$ and $\Im\{V\}$ are the real and imaginary parts of the voltage phasor, respectively [5]. In particular, the measurements from the PMUs satisfy

$$\mathbf{z}_2 = \mathbf{A}\mathbf{v} + \mathbf{u}_2 = \begin{bmatrix} \mathbf{B} \\ \mathbf{Y} \end{bmatrix} \mathbf{v} + \mathbf{u}_2 \tag{13}$$

where each row of matrix $\mathbf{B}$ is a unit vector of appropriate dimension with a "1" placed in the column associated with a particular voltage phasor, $\mathbf{Y}$ is an admittance matrix of appropriate dimension corresponding to the current phasors, and $\mathbf{u}_2$ is a zero-mean Gaussian measurement noise vector.

PMUs provide synchronized local measurements with global time stamps. In other words, with local cooperation (sharing of PMU measurements) between substations, SE can be made more dynamic and reactive to local disturbances before effects cascade through the system. The fact that utility companies are open to sharing PMU data with each other [18] makes distributed SE relevant. When a sufficient number of PMUs are deployed on the grid, the system is fully observable and iterative solutions are avoided as the measurement equation becomes linear as seen in (13). Even though making the system fully observable using PMUs is not yet realizable due to financial constraints, it seems likely that in the near future, we could see large-scale deployment of PMUs in power grid as the deployment costs decrease. However, presently, there is a need for state estimators that combine conventional SCADA and PMU measurements.

FUSING CONVENTIONAL MEASUREMENTS WITH PMU MEASUREMENTS

SSE using both PMU and traditional SCADA measurements has been studied extensively. There are two ways to include PMU measurements in the SE process [5]:

1) A single state estimator, where PMU measurements are mixed with the traditional power flow measurements;
2) A two-stage scheme, where the state estimate obtained from the traditional SCADA measurements in (3) is improved by using a second estimator that employs PMU measurements only.

The latter method has the advantage of leaving the existing SCADA software intact.

Let us consider first the approach when conventional SCADA measurements as formulated by (1) are mixed with PMU measurements as formulated by (13). To jointly process the measurements we first need to relate the PMU state $\mathbf{v}$ of complex phasors (Cartesian coordinates) to the conventional state vector $\mathbf{x}$ (polar coordinates), through a simple nonlinear transformation of the type $\mathbf{v} = \mathbf{g}(\mathbf{x})$. Thus, a single estimator, static or dynamic, that incorporates both conventional and PMU measurements can be derived based on the following augmented measurement model:

$$\begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h}(\mathbf{x}) \\ \mathbf{A}\mathbf{g}(\mathbf{x}) \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{bmatrix}, \tag{14}$$

where $\mathbf{z}_1$ and $\mathbf{n}_1$ are the conventional measurements and noise vectors, and $\mathbf{z}_2$ and $\mathbf{n}_2$ denote the PMU measurements and (transformed) noise vectors, respectively.

Instead of mixing the measurements, we may instead use a two-step approach where the conventional state estimate $\hat{\mathbf{x}}$ from (3) is converted into voltage phasors, i.e., $\hat{\mathbf{v}}_1 = \mathbf{g}(\hat{\mathbf{x}})$, and then used as additional measurements in an augmented form of the linear measurement model (13)

$$\begin{bmatrix} \hat{\mathbf{v}}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{B}} \\ \mathbf{Y} \end{bmatrix} \mathbf{v} + \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}, \tag{15}$$

where $\widetilde{\mathbf{B}}$, like $\mathbf{B}$ in (13), simply sifts out the relevant phasors, and $\mathbf{u}_1$ and $\mathbf{u}_2$ are the noise vectors of (transformed) conventional and PMU measurements, respectively. We may now solve for the unknown phasors $\mathbf{v}$ using a linear WLS approach [5].

The problem of distributed SE incorporating phasor measurements was first introduced by Zhao and Abur [19] who presented a hierarchical scheme for distributed SE using PMU measurements. Jiang et al. [20] used PMU measurements in each region to obtain a hierarchical state estimator that functions in three steps. Inclusion of PMU measurements in multilevel state estimators has also been considered in [7].

In spite of these promising works, many well-known challenges remain in combining PMU measurements with conventional measurements to obtain an optimal state estimate. We outline some of those challenges below:

- *Significantly increased computational burden*: The dimensions of the vectors and matrices involved in the SE process are increased due to the inclusion of PMU measurements.
- *Data tsunami*: The sampling rate of PMUs is around two orders of magnitude higher than the conventional measurements. Novel techniques need to be developed to extract relevant state information from this tidal wave of measurement data.
- *Degraded numerical stability*: Since PMU measurements are significantly more accurate than traditional measurements, inclusion of those measurements in the estimation process often results in ill-conditioned gain or measurement noise covariance matrices.
- *Time skewness*: Synchronized PMU measurements are sampled much faster than nonsynchronized conventional measurements. These two sets of measurements have significantly different sampling rates and are not synchronized with each other.

Some of the above-mentioned challenges will be further discussed in the "Signal Processing and the Smart Grid" section.

DISTRIBUTION SYSTEM STATE ESTIMATION

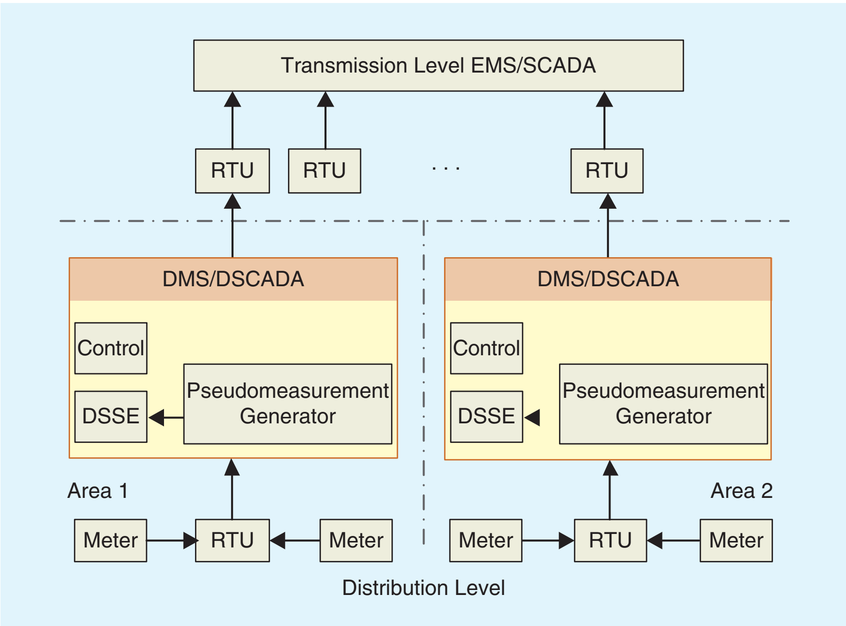
Research on distribution system SE (DSSE) dates back to the early 1990s; see, e.g., [21]. It is known that DSSE could play an important role in DA. However, DA has not been truly brought into fruition, probably due to the lack of proper infrastructure. The states of the future smart grid will undoubtedly be more dynamic, especially in the distribution system. The smart grid is envisioned to include the DG of various types of intermittent renewable sources, the integration of a potentially high level penetration of plug-in hybrid electric vehicles, and DR. While those changes can potentially increase the overall energy effectiveness of the distribution systems, they can also stress the grid, complicate the system operation, and introduce more possibilities for frauds, brownouts and even blackouts. Moreover, events occurring in the smart grid and the subsequent impacts can be too fast to be controlled by any human intervention. Therefore, it is important for the utility companies to have more efficient supervisory planning, enhanced DA, and improved situation awareness throughout the vast and complicated distribution systems. The system operators need to have more timely and reliable knowledge to properly monitor, proactively control, and economically dispatch power through the distribution system. This confers an instrumental role to SE in the development of future distribution systems.

This section first presents some key characteristics in the existing distribution systems that make DSSE different from SE at the transmission level. It then surveys the ongoing research efforts made to develop viable DSSE algorithms. We also discuss the need of DSSE algorithms suitable for active control of the future distribution systems.

DISTRIBUTION SYSTEM STATE ESTIMATION FEATURES AND ALGORITHMS

Figure 3 depicts the general relationship among the SE functions at both transmission and distributions levels as well as their corresponding data acquisition and management systems. The distribution SCADA (DSCADA) system is the counterpart at the distribution level of the SCADA system at the transmission level. At the transmission level, many functions of the EMS are based on the real-time modeling of the system generated by SE. One of the objectives in the development of DSSE is to make it comparable to the transmission level SE. However, the transmission level SE algorithms cannot be directly applied to distribution systems since the operation and planning philosophy of the distribution systems are quite different from those in the transmission systems. In the current distribution systems, the major distinct features for DSSE can be summarized as follows [22]:

- The number of existing telemetered devices that can provide real-time measurements is quite limited, and it is far from being sufficient to provide observability, not to mention bad data detection capability.
- The load data (also known as pseudomeasurements) obtained from historical load profiles and existing automated meter readings (AMRs) devices have limited accuracy.
- Many of the telemetered measurements at the feeders are current, rather than power, which also complicates the measurement functions.



[FIG3] Relationship between the transmission level SE and distribution SE.

- Three-phase imbalance and low reactance/resistance (X/R) ratios further complicate the measurement functions, making the decoupled WLS algorithms for SE at the transmission level not suitable for DSSE.

Additional features may arise as the power grid continues to evolve, since the distribution systems will undergo vast changes with the development of the smart grid. For example, as DG becomes more common with a larger number of microgrids integrated into the grid, the distribution system will be more like a meshed network as opposed to a radial network, like in the traditional grid. As compared to the research and development that have been done for SE at the transmission level, much less has been done for SE at the distribution level. This is perhaps due to 1) DA has not been brought into fruition due to insufficient infrastructure; and 2) SE at the distribution level has far fewer measurements as compared to the number of states to be estimated. The conventional method for solving the DSSE problem is based on the WLS algorithm according to (2), similar to SE at the transmission level. However, DSSE differs from SE at transmission level in many other aspects. Most of the existing DSSE algorithms use nodal voltages as state variables. A typical formulation of three-phase nodal voltage-based SE can be found in [21]. The authors of [21] also employed the same measurement equations as defined in (1), while specifying the corresponding Jacobian elements for branch currents, real and reactive power flow, real and reactive power injections, and nodal voltage magnitudes measurements. A good survey paper on DSSE algorithms and the choice of estimators for DSSE is given in [23].

Various other SP techniques have been explored for studies on DSSE. For example, [24] proposes to model the distribution system as a Bayesian network and employs the belief propagation algorithm to solve a DSSE problem. As another example, DSSE is formulated as a constrained optimization problem which assumes an initial set of DG outputs to be modeled as equality constraints as follows [25]:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \left[\frac{1}{2}\mathbf{r}^{\mathrm{T}}\mathbf{C}_{\mathbf{r}}^{-1}\mathbf{r} + \frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^{\mathrm{T}}\mathbf{P}^{-1}(\mathbf{x}-\bar{\mathbf{x}})\right]$$
$$\text{such that: } \mathbf{r}_m = \mathbf{z}_m - \mathbf{h}_m(\mathbf{x})$$
$$\mathbf{h}_{os}(\mathbf{x}) = \mathbf{0}, \tag{16}$$

where the vectors with subscript m represent those subvectors that result from measurements; $\mathbf{h}_{os}(\hat{\mathbf{x}})$ denotes the operational and structural constraints, $\mathbf{r}$ is the residual vector for the measurements and constraints, i.e., $\mathbf{r} = [\mathbf{r}_m^{\mathrm{T}}\ \mathbf{0}^{\mathrm{T}}]^{\mathrm{T}}$ with $\mathbf{r}_m$ being the residual vector for the measurements, $\mathbf{C}_{\mathbf{r}}$ is the covariance matrix of the residual vector $\mathbf{r}$, $\bar{\mathbf{x}}$ is the vector of available a priori information on the states, and $\mathbf{P}$ is the corresponding covariance matrix. We refer the reader to [25] for detailed treatment of this approach. Some heuristic algorithms have also been proposed for DSSE by considering it as a nonlinear optimization problem. As an example, [26] formulates DSSE as a hybrid particle swarm optimization problem assuming that the telemetered devices can provide magnitude measurements of voltage and current at both the secondary side buses of substations and at RTUs.

DISTRIBUTION SYSTEM STATE ESTIMATION FOR ACTIVE CONTROL

The passive nature of control mechanism implemented in the existing distribution systems limits the number of distributed generators that can be connected. Coupled microgrids can potentially allow for a high penetration level of DGs into the distribution system [27]. A microgrid can be either connected to the main grid under normal conditions, or separated from the main grid during an emergency event or when the quality of the power from the main grid falls below certain standards. In practice, the topology changes can be too fast to be dealt with by any human initiated action. Therefore, supervisory control of microgrids should facilitate intelligent autonomous operation. Hence, active control systems and control functions are preferred for practical purposes. The proper operation of the active control systems calls for timely and accurate knowledge of the operating status of the entire system. Recent efforts in generalized SE [22] and autonomous SE [28] are good initial steps towards achieving this goal. The generalized SE algorithm presented in [22] integrates the estimation of the topology information with the SE process using real-time measurements by modeling parts of the distribution systems at the bus-section/switching-device level. In other words, some topology information is considered as part of the state being estimated, instead of being treated as known and fixed, during the estimation process. The autonomous SE algorithm proposed in [28] automatically identifies the network topology of the distribution system, and then extracts the operating status of the system. The implementation of this algorithm, however, requires upgrading and adding monitoring devices that can stream the necessary data to the control center.

SIGNAL PROCESSING AND THE SMART GRID

While some researchers may consider power grid SE research somewhat mature, new techniques for SE must be developed as the power grid becomes more complex, more interconnected, and more intelligent. Looking ahead, any progress that the SP community can make will greatly facilitate and benefit the development of the smart grid. In return, research on SE within the framework of one of the most complex man-made systems can invigorate the SP research community. As an example, a recent paper addressed the issue of malicious attack on the power grid using SP-based techniques [29]. This section presents some SP-related topics arising from the development of smart grid. These topics are general and apply to SE at both transmission and distribution levels.

FORECASTING-AIDED STATE ESTIMATION

As the power grid continues to evolve, it becomes necessary to closely track state changes to ensure the dependability, reliability, and security of the power system. To this end, there are some open research topics on FASE worth exploring that require application of innovative SP techniques.

Two specific examples are apparent: First, existing FASE algorithms in the literature are mostly implemented with EKF whose limitations are well known. Examples of such limitations include lack of optimality, slow convergence speed, and sensitivity to the linearization error. Fast converging algorithms with good tracking capabilities may be developed by exploiting the equivalence between the incremental Gauss-Newton algorithm and the EKF [30]. There are solutions, which are not yet applied to FASE, that aim at reducing the linearization errors of the EKF, e.g., the iterated EKF, where the point of linearization is reevaluated at the correction step. One may also consider the second-order EKF which performs a second-order Taylor expansion of the nonlinear measurement equation. Additionally, the UKF offers an interesting alternative to EKF-based methods, promising higher accuracy without significant additional computational overhead.

Second, the FASE realm lacks a complete analysis, including building adequate state evolution models, and obtaining proof of optimality. Assumptions about the state-space equations are often made without a rigorous justification. For example, almost all FASE algorithms proposed so far assume that there is no correlation between the state variables, making the state transition matrix $\mathbf{F}(k)$ simply diagonal. More accurate dynamic models are needed to incorporate coupling between state variables. Furthermore, efficient algorithms that are robust to model uncertainty, e.g., sudden topology changes, are of particular interest. For example, H_∞ filters in power system SE deserve further investigation. Computational intelligence tools such as artificial neural networks (ANNs) and fuzzy logic-based approaches have also gained increasing popularity in the field of FASE; see, e.g., [15].

State estimators employing PMU measurements can also be cast into the FASE framework, where either the PMU measurements are mixed with conventional measurements or included in a post-processing step. In the mixed approach, extended or UKFs could be derived for the mixed data model similar to (14). If the PMU measurements are included in a postprocessing step, a linear Kalman filter can be used. However, mixing measurements of different qualities into a single state estimator may cause the covariance matrix of the combined noise vector $[\mathbf{n}_1^{\mathrm{T}}(k)\,\mathbf{n}_2^{\mathrm{T}}(k)]^{\mathrm{T}}$ in (14) to become ill conditioned. In addition, the dimensions of the vectors and matrices involved in the SE process are increased due to the additional PMU measurements, which may lead to significantly increased computational complexity. This problem can be cast into a constrained Kalman filtering problem, where high-quality measurements are employed as deterministic equality or inequality constraints. For example, more robust FASEs with a reduced order KF may be derived by applying the ideas proposed in [31].

Furthermore, as mentioned previously, there is a need for improved state transition models, which are specifically designed with the future power grid in mind. Environmental awareness and climate change have also led to the emergence of renewable and sustainable methods of electricity generation which comes with the challenge of variability. However, there are statistical models available to predict the behavior of these sources. It is an open research problem to develop techniques to incorporate these into new state transition models, which in turn, leads to improved state estimators.

DISTRIBUTED ESTIMATION AND MASE

Electricity market deregulation may require utility companies to monitor the grid over a very large geographical area. Meanwhile, the number of monitoring devices may also grow significantly making the number of measurements prohibitively large for the centralized SE techniques to be effective. A more feasible approach would be to distribute the SE function throughout the interconnection. This kind of distributed SE facilitates interconnection-wide coordinated monitoring as well as the development of many other smart grid functions like self healing [32].

Distributed approaches can enhance the computational performance and the reliability of SE algorithms. Efficient and reliable communication is the backbone for the distributed SE algorithms. However, various challenges in communication prevent the realization of such approaches. Communication delay consists of a great portion of the response time, which has to be shortened to compute state estimates that are meaningful for those time-sensitive functions like self healing. The problems of optimizing the locations of distributed processors to minimize the communication delays while keeping the communication overhead within the practical constraints need to be addressed. The time skewness among measurements is another issue which can potentially be detrimental to the control decisions made based on the state estimates. One way to tackle this is to enforce all measurements synchronized by GPS, which requires significant investment in upgrading the infrastructure. Another, more economical, way of handling this is to design filters to mitigate the time skewness effects by utilizing the statistics of the delays [33]. Therefore, the development of new distributed algorithms is a challenging but essential task for the development of future smart grids.

MASE can most likely benefit from recent advances in distributed estimation, which has recently been an active field of research in the SP community. In distributed estimation, several nodes (or areas in case of MASE) estimate a common parameter vector through local collaborations. In the case of MASE, the measurements of each area only relates to a small part of the whole state vector. Thus, the resulting computational and communication costs of a distributed estimation approach depend on whether local knowledge of the whole state vector is required or not. For example, by redefining the correction vector in (3) as $\Delta\mathbf{x} = [\Delta\mathbf{x}_1^{\mathrm{T}} \cdots \Delta\mathbf{x}_M^{\mathrm{T}}]^{\mathrm{T}}$, an iterative WLS solution for the MASE would take the form

$$\left[\sum_{m=1}^{M}\mathbf{H}_m^{\mathrm{T}}(j)\mathbf{W}_m^{-1}\mathbf{H}_m(j)\right]\Delta\mathbf{x}(j) = \sum_{m=1}^{M}\mathbf{H}_m^{\mathrm{T}}(j)\mathbf{W}_m^{-1}[\mathbf{z}_m - \mathbf{h}_m(\mathbf{x}_m(j))]$$

$$\hat{\mathbf{x}}(j+1) = \hat{\mathbf{x}}(j) + \Delta\mathbf{x}(j), \quad (17)$$

where $\mathbf{H}_m(j)$ is the measurement Jacobian of area m obtained with the local state estimate $\hat{\mathbf{x}}_m(j)$. The simplest method seems to be to express (17) in terms of averages $1/M \sum_m \mathbf{H}_m^{\mathrm{T}} \mathbf{W}_m^{-1} \mathbf{H}_m$ and $1/M \sum_m \mathbf{H}_m^{\mathrm{T}} \mathbf{W}_m^{-1} [\mathbf{z}_m - \mathbf{h}_m(\mathbf{x}_m)]$ across the areas. The so-called consensus algorithms (also related to gossip algorithms) can be expected to be of relevance here; see, e.g., [34] and [35]. Thus, two separate consensus algorithms can be used to compute these quantities, and in turn, the quantity $\Delta\mathbf{x}$ to be used at the jth iteration of the estimator. However, this method requires communication corresponding to two consensus algorithms being executed in parallel. Perhaps more importantly, it requires the two algorithms to converge before the quantity $\Delta\mathbf{x}$can be computed and hence the iteration of the estimation algorithm can be done. In other words, this approach requires the consensus algorithm to be executed at a much faster time scale than the consensus algorithm. While there has been some characterization of the performance loss when the time-scales do not separate smoothly [36]–[38], the general problem still remains open. Since the problem is reminiscent of the classical information fusion and the distributed Kalman filtering problems studied in SP literature, this problem may be of independent interest. Furthermore, taking into account the sparseness of the problem, dynamic Kalman filter-based solutions for sparse systems, e.g., [39], can be useful when the amount of information shared between neighbors is kept to a minimum to reduce the need for an excessive communication infrastructure.

In summary, the SP community can contribute to the research on MASE for the future grid by building upon recent advances in distributed estimation, e.g., by not only developing new resource efficient algorithms but also analyzing their behavior in terms of convergence speed and stability.

EVENT-TRIGGERED APPROACHES TO STATE ESTIMATON

To realize the envisioned functionalities, the future grid will be equipped with a myriad of smart meters which will collect and transmit massive amount of data, and the control center will need to process those data, convert data into information and transform information into actionable intelligence. In fact, the deployment of PMUs at the transmission level has already resulted in more data than the legacy grid's control center can handle. When the smart grid is fully deployed, there is a risk that the grid's operator will be drowned in data, a phenomenon that is termed *data tsunami*. In the development of smart grid, the designers must be mindful of preventing this effect, allowing the control center to "separate the wheat from the chaff" in a timely manner. Furthermore, it is desirable to make the communication infrastructure throughout the grid energy- and bandwidth-efficient. Hence, an event-triggered approach to sensing, communicating and information processing would be quite appealing. The challenge here is to provide analytical performance guarantees in a distributed event triggering algorithm in a dynamically changing environment.

The event-triggered approach can be adopted for SE. An example to consider is for MASE as described in the previous section, where distributed estimation can be employed. In the event-triggered MASE, the local areas update their state estimates only when needed and cooperate (transmit the estimates) only when such an action is informative. An adaptive estimation paradigm, referred to as set-membership adaptive filter (SMAF), offers a viable solution to this approach [40]. The SMAF algorithms feature selective update of estimates. This is in contrast with conventional adaptive estimation algorithms such as recursive least squares and least mean squares, which update parameter estimates continually regardless of the benefits of such updates. In SMAF, estimates are updated only when the measurements offer sufficient innovation, as measured by some function of estimation error. Accordingly, distributed estimation derived from SMAF communicates only when such an action is informative.

CONCLUSIONS

SE is a fundamental functionality to ensure smooth, reliable and secure operation of power grids. At the transmission level and, particularly in the context of static estimation, SE has a rich history. New developments that are in the offering for the smart grid render such existing methods inadequate. Of particular interest are the challenges introduced by new metering infrastructure such as PMUs, variable and distributed sources including renewables, and structural changes resulting from integration of microgrids. The demands on SE are also much more stringent now and concerns such as reliability, dependability, security, and distributed nature and dynamic SE necessitate a paradigm shift from the existing algorithms. The SP community has much to offer in meeting these exciting challenges.

ACKNOWLEDGMENTS

The work of Dr. Werner and N. Kashyap has been supported, in part, by the Academy of Finland, Centre of Excellence Smart Radios and Wireless Research at the Aalto University School of Electrical Engineering. The work of Dr. Huang and Jing Huang has been supported, in part, by GE Energy under contract 900134438 and, in part, by an initiation grant of the 2011 Faculty Research Scholarship Program from the University of Notre Dame.

AUTHORS

Yih-Fang Huang (huang@nd.edu) is a professor of electrical engineering at the University of Notre Dame, where he started as an assistant professor upon receiving his Ph.D. degree in 1982 from Princeton University. He served as chair of the department from 1998 to 2006. His research focuses on statistical and adaptive SP. He was a visiting professor at the Munich University of Technology, Germany, in 2007 and subsequently received a Fulbright-Nokia scholarship for lectures/research at Helsinki University of Technology in Finland. In 1999, he received the Golden Jubilee Medal of the IEEE Circuits and Systems Society, for which he served as vice president in 1997–1998 and was a Distinguished Lecturer in 2000–2001. He is a Fellow of the IEEE.

Stefan Werner (stefan.werner@aalto.fi) received the M.Sc. degree in electrical engineering from the Royal Institute of Technology (KTH), Stockholm, Sweden, in 1998 and the D.Sc. (electrical engineering) degree (with honors) from the Signal

Processing Laboratory, Helsinki University of Technology (TKK), Espoo, Finland, in 2002. He is currently an academy research fellow with the Department of Signal Processing and Acoustics, Aalto University, Finland, where he is also appointed as a Docent. His research interests include adaptive SP, SP for communications, and statistical SP. He is a member of the editorial board for EURASIP's *Signal Processing* journal.

Jing Huang (jhuang6@nd.edu) received the B.S. degree from Shanghai Jiao Tong University, Shanghai, in 2007 and the M.S. degree from the University of Notre Dame, Notre Dame, Indiana, in 2010, all in electrical engineering. He is currently a Ph.D. candidate at the University of Notre Dame. His research interests include SP with its application to communication systems and smart grids.

Neelabh Kashyap (neelabh.kashyap@aalto.fi) received the B.E. degree in electronics and communications engineering in 2009 from the Visvesvaraya Technological University, India. He is currently pursuing his M.Sc. degree at the Aalto University School of Electrical Engineering, Finland. His research interests are statistical SP and smart grids.

Vijay Gupta (vgupta2@nd.edu) received the B.Tech degree from the Indian Institute of Technology, Delhi, and the M.S. and Ph.D. degrees from the California Institute of Technology, Pasadena, all in electrical engineering. He is an assistant professor in the Department of Electrical Engineering, University of Notre Dame, Indiana. His research interests include networked control systems and sensor networks, distributed estimation and detection, and, in general, the interaction of communication, computation, and control. He received the NSF CAREER Award and the Ruth and Joel Spira Award for Excellence in Teaching in the Department of Electrical Engineering at Notre Dame.

REFERENCES

[1] F. C. Schweppe and J. Wildes, "Power system staticstate estimation—Part I, II & III," *IEEE Trans. Power App. Syst.*, vol. 89, no. 1, pp. 120–135, Jan. 1970.

[2] F. F. Wu, "Power system state estimation: A survey," *Int. J. Electr. Power Energy Syst.*, vol. 12, no. 2, pp. 80–87, 1990.

[3] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[4] P. Favre-Perrod, R. Critchley, E. Catz, and M. Bazargan, "New participants in smartgrids and associated challenges in the transition towards the grid of the future," in *Proc. 2009 IEEE Bucharest PowerTech*, July 2009, pp. 1–5.

[5] M. Zhou, V. A. Centeno, J. S. Thorp, and A. G. Phadke, "An alternative for including phasor measurements in state estimators," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1930–1937, Nov. 2006.

[6] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. Begovic, and A. Phadke, "Widearea monitoring, protection, and control of future electric power networks," *Proc. IEEE*, vol. 99, no. 1, pp. 80–93, Jan. 2011.

[7] A. Gómez-Expósito, A. Abur, A. de la Villa Jaén, and C. Gómez-Quiles, "A multilevel state estimation paradigm for smart grids," *Proc. IEEE*, vol. 99, no. 6, pp. 952–976, June 2011.

[8] M. Shahidehpour, H. Yamin, and Z. Li, Market *Operations in Electric Power Systems*. New York: Wiley, 2002.

[9] A. Monticelli and A. Garcia, "A hybrid state estimator: Solving normal equations by orthogonal transformations," *IEEE Trans. Power App. Syst.*, vol. 104, no. 12, pp. 3460–3468, Dec. 1985.

[10] A. S. Debs and R. E. Larson, "A dynamic estimator for tracking the state of a power system," *IEEE Trans. Power App. Syst.*, vol. 89, no. 7, pp. 1670–1678, Sept. 1970.

[11] F. C. Schweppe and R. D. Masiello, "A tracking static state estimator," *IEEE Trans. Power App. Syst.*, vol. 90, no. 3, pp. 1025–1033, May 1971.

[12] A. M. Leite da Silva, M. B. Do Coutto Filho, and J. F. de Queiroz, "State forecasting in electric power systems," *Inst. Elect. Eng. Proc., Gener., Trans. Distrib.*, vol. 130, no. 5, pp. 237–244, Sept. 1983.

[13] G. Durgaprasad and S. S. Thakur, "Robust dynamic state estimation of power systems based on M-estimation and realistic modeling of system dynamics," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1331–1336, Nov. 1998.

[14] G. Valverde and V. Terzija, "Unscented Kalman filter for power system dynamic state estimation," *IET Gener., Trans. Distrib.*, vol. 5, no. 1, pp. 29–37, Jan. 2011.

[15] M. Brown Do Coutto Filho and J. C. S. de Souza, "Forecasting-aided state estimation—Part I: Panorama," *IEEE Trans. Power Syst.*, vol. 24, no. 4, pp. 1667–1677, Nov. 2009.

[16] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, and T. Van Cutsem, "A taxonomy of multi-area state estimation methods," *Electric Power Syst. Res.*, vol. 81, no. 4, pp. 1060–1069, 2011.

[17] G. N. Korres, "A distributed multi-area state estimation," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 73–84, Feb. 2011.

[18] H. Wu and J. Giri, "PMU impact on state estimation reliability for improved grid security," in *Proc. IEEE Transmission and Distribution Conf. and Exhibition*, May 2006, pp. 1349–1351.

[19] L. Zhao and A. Abur, "Multi-area state estimation using synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 611–617, May 2005.

[20] W. Jiang, V. Vittal, and G. T. Heydt, "A distributed state estimator utilizing synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 563–571, May 2007.

[21] 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.

[22] G. N. Korres, N. D. Hatziargyriou, and P. J. Katsikas, "State estimation in multi-microgrids," *Eur. Trans. Electr. Power*, pp. 1178–1199, June 2010.

[23] R. Singh, B. C. Pal, and R. A. Jabr, "Choice of estimator for distribution system state estimation," *IET Gener., Trans. Distrib.*, vol. 3, no. 7, pp. 666–678, July 2009.

[24] Y. Hu, A. Kuh, T. Yang, and A. Kavcic, "A belief propagation based power distribution system state estimator," *IEEE Comput. Intell. Mag.*, vol. 6, no. 3, pp. 36–46, Aug. 2011.

[25] A. Souza, E. M. Lourenco, and A. S. Costa, "Real-time monitoring of distributed generation through state estimation and geometrically-based tests," in *Proc. iREP Symp. Power System Dynamics and Control*, Aug. 2010, pp. 1–8.

[26] S. Naka, T. Genji, T. Yura, and Y. Fukuyama, "A hybrid particle swarm optimization for distribution state estimation," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 60–68, Feb. 2003.

[27] R. H. Lasseter, "Smart distribution: Coupled microgrids," *Proc. IEEE*, vol. 99, no. 6, pp. 1074–1082, June 2011.

[28] S. Choi, B. Kim, G. J. Cokkinides, and A. P. S. Meliopoulos, "Feasibility study: Autonomous state estimation in distribution systems," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2109–2117, Nov. 2011.

[29] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attack on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[30] B. M. Bell and F. W. Cathey, "The iterated Kalman filter update as a Gauss–Newton method," *IEEE Trans. Automat. Contr.*, vol. 38, no. 2, pp. 294–297, Feb. 1993.

[31] E. Fogel and Y. Huang, "Reduced-order optimal state estimator for linear systems with partially noise corrupted measurement," *IEEE Trans. Automat. Contr.*, vol. 25, no. 5, pp. 994–996, Oct. 1980.

[32] K. Moslehi and R. Kumar, "A reliability perspective of the smart grid," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 57–64, June 2010.

[33] M. Choi, J. Choi, J. Park, and W. K. Chung, "State estimation with delayed measurements considering uncertainty of time delay," in *Proc. IEEE Int. Conf. Robotics and Automation*, May 2009, pp. 3987–3992.

[34] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. Int. Conf. Information Processing in Sensor Networks*, Los Angeles, CA, Apr. 2005, pp. 63–70.

[35] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Trans. Automat. Contr.*, vol. 55, no. 9, pp. 2069–2084, Sept. 2010.

[36] D. P. Spanos and R. M. Murray, "Distributed sensor fusion using dynamic consensus," in *Proc. IFAC World Congr.*, Paper We-E0-TP/3, 2005.

[37] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. Decision and Control*, Dec. 2007, pp. 5492–5498.

[38] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Approximate distributed Kalman filtering in sensor networks with quantifiable performance," in *Proc. Int. Conf. Information Processing in Sensor Networks (IPSN)*, Apr. 2005, pp. 133–139.

[39] U. A. Khan and J. M. F. Moura, "Distributing the Kalman filter for large-scale systems," *IEEE Trans. Signal Processing*, vol. 56, no. 10, pp. 4919–4935, Oct. 2008.

[40] S. Werner, M. Mohammed, Y.-F. Huang, and V. Koivunen, "Decentralized set-membership adaptive estimation for clustered sensor networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Apr. 2008, pp. 3573–3576.

[SP]