A Hybrid Framework Combining Model-Based and Data-Driven Methods for Hierarchical Decentralized Robust Dynamic State Estimation

Preprint

Marcos Netto,¹ Venkat Krishnan,² Lamine Mili,¹ Yoshihiko Susuki,³ and Yingchen Zhang²

¹ Virginia Polytechnic Institute and State University
² National Renewable Energy Laboratory
³ Osaka Prefecture University

To be presented at the 2019 IEEE Power and Energy Society General Meeting (IEEE PES GM)
Atlanta, Georgia
August 4–8, 2019
A Hybrid Framework Combining Model-Based and Data-Driven Methods for Hierarchical Decentralized Robust Dynamic State Estimation

Preprint

Marcos Netto,¹ Venkat Krishnan,² Lamine Mili,¹ Yoshihiko Susuki,³ and Yingchen Zhang²

¹ Virginia Polytechnic Institute and State University
² National Renewable Energy Laboratory
³ Osaka Prefecture University

Suggested Citation
A Hybrid Framework Combining Model-Based and Data-Driven Methods for Hierarchical Decentralized Robust Dynamic State Estimation

Marcos Netto, Student Member, IEEE, Venkat Krishnan, Senior Member, IEEE, Lamine Mili, Life Fellow, IEEE, Yoshihiko Susuki, Member, IEEE, and Yingchen Zhang, Senior Member, IEEE

Abstract—This paper combines model-based and data-driven methods to develop a hierarchical, decentralized, robust dynamic state estimator (DSE). A two-level hierarchy is proposed where in the lower level comprises robust, model-based, decentralized DSEs. The state estimates sent from the lower level are received at the upper level, where they are filtered by a robust data-driven DSE after a principled sparse selection. This selection allows us to shrink the dimension of the problem at the upper level and hence significantly speed up the computational time. The proposed hybrid framework does not depend on the centralized infrastructure of the control centers; thus it can be completely embedded into the wide-area measurement systems. This feature will ultimately facilitate the placement of hierarchical decentralized control schemes at the phasor data concentrator locations. Also, the network model is not necessary, thus a topology processor is not required. Finally, there is no assumption on the dynamics of the electric loads. The proposed framework is tested on the 2,000-bus synthetic Texas system, and shown to be capable of reconstructing the dynamic states of the generators with high accuracy, and of forecasting in the advent of missing data.

Index Terms—Compressed sensing, data-driven dynamical systems, dynamic state estimation, Kalman filtering, Koopman mode decomposition, sparse selection.

I. INTRODUCTION

The large-scale deployment of phasor measurement units (PMUs) and other grid-edge metering devices, such as smart meters [1] and micro-PMUs [2], is enabling the application of data analytics in electric power systems. This is fueling interest in data-driven methods that can enhance power system's reliability and resilience [3]. Simultaneously, and perhaps most importantly, we are witnessing an unprecedented increase in the share of renewable energy sources [4], [5], which are inherently intermittent and uncertain. The consequent increase in the stochastic dynamics of the net load challenges the traditional deterministic model-based, real-time monitoring and control methods applied in the legacy systems [6], [7]. Therefore, power engineers are looking to exploit the fast sampled measurement data to advance grid modeling, state estimation, forecasting, and controls through data analytics [8], [9], including by proposing data-driven dynamic state estimators (DSEs) [10], [11], which can track system dynamic states [12].

In this paper, we develop a two-level, hierarchical, decentralized, robust DSE by combining model-based and data-driven methods. The model-based decentralized DSEs at the lower level provide the necessary data for performing data-driven model identification at the upper level by using the Koopman mode decomposition (KMD) [13], which further allows for dynamic stability assessment and modal analysis of nonlinear dynamical systems [14]. The paper also illustrates the application of the proposed framework for large-scale systems by using compressed sensing [15] to find sparse state estimate selection. Because the use of all state estimates at the upper level could be prohibitive for high-dimensional systems, we rely on compressed sensing to find a sparse selection of state estimates, following the work of Brunton et al. [16], [17]. The proposed hybrid framework provides the opportunity to devise powerful tools by combining concepts from dynamical systems, estimation, and control theory. Firstly, the use of decentralized DSEs and KMD makes it independent of the network model; hence, it does not require a topology processor. Secondly, by virtue of the data-driven KMD, the method does not need to make assumptions about the underlying load model dynamics. Thirdly, and most notably, it can be completely embedded into the wide-area measurement systems instead of being an add-on to the energy management systems installed at the control centers. This attribute will ultimately facilitate the hierarchical control design of electric power systems [18] with the placement of control schemes at the phasor data concentrator (PDC) location, thereby exploiting the synergies between agile, low-latency, decentralized and holistic centralized monitoring and control architectures.

The paper proceeds as follows. Section II summarizes relevant literature and motivation for this work. Section III briefly introduces the data-driven robust DSE, and Section IV presents the sparse selection approach. Numerical results are discussed in Section V. The conclusions and directions for future research are given in Section VI.

II. BACKGROUND AND MOTIVATION

Following the work of Modir and Schlueter [19], several DSEs [20]--[23] have been proposed based on the conceptual idea of transmitting all the PMU measurements to a centralized location, such as the control center, where the DSE is supposedly installed. We refer to this configuration as the centralized
DSE; see Fig. 1a for more details. The assumptions are that the system is fully observable by PMUs and that the Kron-reduced network model (KNM) [24] is accessible. Although this may be true for independent system operators (ISOs) and regional transmission organizations (RTOs), access to the KNM is a hurdle for local utilities and transmission operators because of limited data sharing across neighboring systems. One alternative is to build dynamic reduced-order models [25]; unfortunately, this approach will increase the existing model uncertainties. Further, the KNM assumes that the electric loads are modeled as constant admittances, thereby not allowing for capturing the rich dynamics of composite electric loads [26]. Another important point that has been overlooked is that a topology processor is required to determine the KNM in real time; the existing topology processors embedded on the static state estimators cannot supply this demand. Nevertheless, although of reduced dimension, the graph of the KNM is full [24] rather than sparse, hence sparsity techniques cannot be exploited. The computational burden is another important barrier to the adoption of model-based centralized DSEs.

As an attempt to overcome the aforementioned problems, model-based decentralized DSEs have been investigated; see, for instance, [29], [30]. The idea is to individually estimate the states of each generator at their own location, as in Fig. 1b. Given that each decentralized DSE is designed for a particular generation location, only a PMU at the generator terminal is assumed. The decentralized DSE requires neither the model of the network nor the model of the electric loads, and it is computationally inexpensive because of a reduced number of involved state variables. Further, it offers the additional benefit of performing model calibration [31] without taking the generator offline, which is an appealing attribute from the system’s reliability and economic standpoints. However, although a decentralized DSE allows for local stability assessment, control, and protection, for wide-area stability assessment and control, the DSE states will still need to be communicated to centralized processing, like any other architecture shown in Fig. 1; and, although the data transmissions in this case are not raw measurements but actual system states, it is still necessary to ensure data integrity, cybersecurity, and robustness against bad or missing information at the receiving end, i.e., at the control center. Additionally, in the event of loss of communication with a decentralized DSE, the states of the associated generator must be forecast. To circumvent some of these issues, Paul et al. [32] proposed a computationally distributed but physically centralized DSE, wherein the conceptual idea is to send all the measurements to a centralized location instead of performing the estimation separately for each generator; see Fig. 1c. Although ingenious, the approach in [32] relies on the reference signals of voltage and power that are accessible only to ISOs and RTOs, thereby precluding its adoption by transmission companies and alike utilities. The state forecasting capability is also not available.

Given the promise that decentralized DSEs have in terms of reducing the computational burden on increasing proliferations of sensors and data, and given its favorable trade-off between local and wide-area assessments, we delve deeper into the decentralized DSE. Presuming that model-based decentralized DSEs are available, we pose the following questions:

- Suppose that the state estimates are sent to a centralized location. How could one verify the data integrity of the received state estimates without relying on any model?
- How could one forecast the dynamic states in case of a communication problem with a given decentralized DSE?

To address these questions, we suggest the use of the robust, data-driven DSE, which is presented next.

III. ROBUST KOOPMAN KALMAN FILTER

This section introduces the data-driven method, namely, the Koopman operator-based method that enables studying nonlinear dynamical systems without relying on any model. Further, the developed robust Koopman Kalman filter (KKF) [10], [11] lays the foundation for ensuring that the state estimates received from the decentralized DSEs will be robustly filtered.

A. The Koopman Operator

Consider a discrete-time autonomous dynamical system:

\[ x_k = f(x_{k-1}), \]  

where the state \( x \) is an element of the state space \( S \subset \mathbb{R}^n \), \( f : S \rightarrow S \) is the discrete map, and \( k \in \mathbb{Z} \) is the time index. Define \( g : S \rightarrow \mathbb{R} \), a vector-valued observable in (1). The Koopman operator, \( \mathcal{K} \), is a linear transformation on this vector space:

\[ \mathcal{K} g(x_k) = g \circ f(x_k) = g(f(x_k)). \]

The Koopman eigenvalues, \( \mu_i \), and the Koopman eigenfunctions (KEFs), \( \varphi_i \), of \( \mathcal{K} \) are defined as follows:

\[ \mathcal{K} \varphi_i(x_k) = \mu_i \varphi_i(x_k), \quad i = 1, 2, \ldots \]
If all the elements of $\mathbf{g}$ lie within the span of the KEFs, then we have:

$$
\mathbf{g}(\mathbf{x}_k) = \sum_{i=1}^{\infty} \varphi_i(\mathbf{x}_k) \phi_i = \sum_{i=1}^{\infty} \varphi_i(\mathbf{x}_0) \mu_i \phi_i^k, \quad (4)
$$

where $\phi_i$ are the Koopman modes [13], and $\{\mu_i, \varphi_i, \phi_i\}, i = 1, 2, ..., \infty$ are referred to as the Koopman tuples. The interested reader is referred to [13] for more details on the derivation of (4). For power systems, because of the existence of multiple attractors, we estimate a subset of the Koopman tuples, such that $\mathbf{g}(\mathbf{x}_k) \approx \sum_{i=1}^q \varphi_i(\mathbf{x}_k) \phi_i = \sum_{i=1}^q \varphi_i(\mathbf{x}_0) \mu_i \phi_i^k$.

**B. Koopman Canonical Coordinates [33]**

Consider the discrete-time autonomous dynamical system:

$$
\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}), \quad \mathbf{y}_k = \mathbf{h}(\mathbf{x}_k), 
\quad (5)
$$

where $\mathbf{x}$ and $\mathbf{f}$ are defined as in (1), $\mathbf{y} \in \mathbb{R}^m$ is the system observation vector, and $\mathbf{h} : S \rightarrow \mathbb{R}^m$. Let $\mathcal{F}^q = \text{span}\{\varphi_i\}_{i=1}^q$ be a subset of the KEFs, such that $\mathbf{x}, \mathbf{y} \in \mathcal{F}^q$. Then, we have:

$$
\mathbf{x}_k = \sum_{i=1}^q \varphi_i(\mathbf{x}_{k-1}) \phi_i^k \mu_i, \quad \mathbf{y}_k = \sum_{i=1}^q \varphi_i(\mathbf{x}_{k-1}) \phi_i(\mathbf{y}) \mu_i, 
\quad (6)
$$

Now, suppose that the Koopman tuples are ordered such that complex conjugate pairs appear adjacent to each other. It can be shown [11] that a nonlinear change of coordinates is given by $\mathcal{T}_K : \mathbb{R}^n \rightarrow \mathbb{R}^q$, expressed as:

$$
\mathcal{T}_K(\mathbf{x}) = \mathbf{x} = [\mathbf{x}_1, ..., \mathbf{x}_q]^T, 
\quad (7)
$$

where $\mathbf{T}$ denotes the transpose of a vector, and $\mathbf{x}_i = \varphi_i$, if $\varphi_i$ is real-valued,

$$
\mathbf{x}_{i+1} = 2\Re\{\varphi_i\} \text{ and } \mathbf{x}_{i+1} = -2\Im\{\varphi_i\}, \text{ if } (\varphi_i, \varphi_{i+1}) \text{ is a complex conjugate pair},
\quad \mathbb{R}\{\varphi_i\} \text{ and } \mathbb{R}\{\varphi_i\} \text{ are the real and imaginary parts of } \varphi_i, 
\text{ respectively}. 
$$

Using (7), and after some algebraic manipulation [11], we have:

$$
\mathbf{x}_k = \Omega \mathbf{x}_{k-1}, \quad \mathbf{x}_k = \Phi(\mathbf{x}) \mathbf{x}_k, \quad \mathbf{y}_k = \Phi(y) \mathbf{x}_k, \quad (8)
$$

and, remarkably, the nonlinear dynamical system in (5) can be mapped to the linear dynamical system (8). The KEFs define the Koopman canonical coordinates (KCC). If $\varphi_i$ is real-valued, $\Omega_{i,i} = \mu_i$. If $(\varphi_i, \varphi_{i+1})$ form a complex conjugate pair, then we have:

$$
\begin{bmatrix}
\Omega_{i,i} & \Omega_{i,i+1} \\
\Omega_{i+1,i} & \Omega_{i+1,i+1}
\end{bmatrix} =
\begin{bmatrix}
\Re\{\mu_i\} & \Im\{\mu_i\} \\
-\Im\{\mu_i\} & \Re\{\mu_i\}
\end{bmatrix},
$$

and, thus, $\Omega$ is a block diagonal matrix.

The matrix $\Phi(\mathbf{x}) \in \mathbb{R}^{n \times q}$ is a mapping between the states in the KCC, $\mathbf{x}$, and the original state space coordinates, $\mathbf{x}$. If $\varphi_i$ is real-valued, $\Phi_{i,i} = \mathbf{x}_i$. Conversely, if $(\varphi_i, \varphi_{i+1})$ form a complex conjugate pair, $\Phi_{i,i} = \mathbb{R}\{\varphi_i\}$, $\Phi_{i+1,i} = \mathbb{R}\{\varphi_{i+1}\}$.

The matrix $\Phi(y) \in \mathbb{R}^{m \times q}$ is a mapping between the states in the KCC, $\mathbf{x}$, and the observation vector in the original state space coordinates, $\mathbf{y}$. If $\varphi_i$ is real-valued, $\Phi_{i,i} = \mathbf{y}_i$. If $(\varphi_i, \varphi_{i+1})$ form a complex conjugate pair, $\Phi_{i,i} = \mathbb{R}\{\varphi_i\}$, $\Phi_{i+1,i+1} = \mathbb{R}\{\varphi_{i+1}\}$.

**C. Robust Koopman Kalman Filter**

In the KKF form, (8) becomes:

$$
\mathbf{z}_k = \Omega \mathbf{z}_{k-1} + \mathbf{w}_{k-1}, \\
\mathbf{y}_k = \Phi(y) \mathbf{z}_k + \mathbf{v}_k, 
\quad (9)
$$

where $\mathbf{w}$ stands for the system process error, and $\mathbf{v}$ denotes the measurement noise. In what follows, we rely on [34] to solve (9). The interested reader is referred to [11] for details on the derivation of the robust KKF.

Now, the practical experience suggests that the synchronous generators oscillate coherently; see Fig. 2. This is equivalent of saying that the dynamics of electrical power systems evolve on a low-dimensional attractor. In the next section, this property is leveraged to perform a sparse selection of the state estimates used to identify the KKF, which is applicable to larger power systems.

**IV. SPARSE SELECTION OF STATE ESTIMATES**

Following Manohar et al. [17], suppose that a state $\mathbf{x}$ evolving according to the nonlinear dynamics (1) has a compact representation in a transform basis $\Psi$. In a universal basis $\Psi \in \mathbb{R}^{n \times n}$, $\mathbf{x}$ might have a sparse representation:

$$
\mathbf{x} = \mathbf{\Psi} \mathbf{s}, 
\quad (10)
$$

$s \in \mathbb{R}^n$ is a sparse vector. In a tailored basis $\mathbf{\Psi}_r \in \mathbb{R}^{n \times r}$, such as a basis defined by a proper orthogonal decomposition, $\mathbf{x}$ might have a low-rank representation:

$$
\mathbf{x} = \mathbf{\Psi}_r \ell, 
\quad (11)
$$

$\ell \in \mathbb{R}^r$. We seek to find a matrix $\mathbf{C} \in \mathbb{R}^{p \times n}$ consisting of a small number ($p \ll n$) of optimized measurements:

$$
\mathbf{b} = \mathbf{C} \mathbf{x}, 
\quad (12)
$$

$\mathbf{b} \in \mathbb{R}^p$, which facilitates the accurate reconstruction of $\mathbf{s}$ or $\ell$, and thus $\mathbf{x}$. Note that $\mathbf{C} = [e_{\gamma_1} e_{\gamma_2} ... e_{\gamma_p}]$, where $e_{\gamma_i}$ is the unit vector with a unit entry at index $\gamma_i$ and zeros elsewhere. Combining (10) and (12) yields:

$$
\mathbf{b} = (\mathbf{C}^\dagger \mathbf{s}) \ell = \mathbf{C}^\dagger \mathbf{s} \ell. 
\quad (13)
$$

Eq. (13) is referred to as the compressed sensing problem. Conversely, combining (11) and (12) yields:

$$
\mathbf{b} = (\mathbf{C}^\dagger \mathbf{\Psi}_r) \ell = \mathbf{C}^\dagger \mathbf{\Psi}_r \ell. 
\quad (14)
$$

If $\mathbf{C}$ is properly structured such that $\mathbf{\ell}$ is well conditioned, it is possible to solve for the low-rank coefficients $\ell$ given the measurements $\mathbf{b}$ in (14) as follows:

$$
\hat{\ell} = \begin{cases}
\theta^\dagger \mathbf{b} = (\mathbf{C} \mathbf{\Psi}_r)^{-1} \mathbf{b}, & p = r, \\
\theta^\dagger \mathbf{b} = (\mathbf{C} \mathbf{\Psi}_r^\dagger) \mathbf{b}, & p > r,
\end{cases} 
\quad (15)
$$

$\theta^\dagger$ denotes the Moore-Penrose pseudoinverse of $\theta$. Thus, $\mathbf{x}$ can be estimated as $\mathbf{x} = \mathbf{\Psi}_r \hat{\ell}$. From (15), one seeks columns of $\mathbf{\Psi}_r$ corresponding to point sensor locations in the state space, $e_{\gamma_i}$, that optimally condition the inversion of the matrix $\mathbf{\theta}$. The structure of the elements $e_{\gamma_i}$ affect the condition number of $\mathbf{C}$ and consequently of $\mathbf{M}_r = \theta^\dagger \mathbf{\ell}$. The condition number of the system might be indirectly bounded by optimizing the
spectral content of $M_\gamma$ using its determinant, trace, or spectral radius. For example, we have:

$$\gamma_* = \arg \max_{\gamma, |\gamma| = p} \det M_\gamma = \arg \max_{\gamma} \prod_i |\lambda_i(M_\gamma)|$$

$$= \arg \max_{\gamma} \prod_i \sigma_i(M_\gamma), \quad (16)$$

where $\lambda_i$ and $\sigma_i$ are, respectively, the $i$-th eigenvalue and singular value of $M_\gamma$. The QR factorization with column pivoting decomposes a matrix $M_\gamma \in \mathbb{R}^{m \times n}$ into a unitary matrix $Q$, an upper triangular matrix $R$, and a column permutation matrix $C$, that is:

$$M_\gamma C^\top = QR. \quad (17)$$

The key idea from [17] is, when applied to an appropriate basis, the QR pivoting procedure provides an approximate greedy solution method for the optimization in (16), also known as a submatrix volume maximization because the matrix volume is the absolute value of the determinant.

A. Sparse Selection of State Estimates

In [17], the QR pivoting procedure is proposed as a tool to optimize sensor placement, in particular for the reconstruction of high-dimensional states from point measurements given tailored bases. Instead of measurements, here, we have access to the state estimates received from the robust, model-based, decentralized DSEs. Our objective is to shrink the number of state estimates used by the robust KKF so as to speed up its processing.

Algorithm 1 Sparse selection of state estimates

1: procedure
2: $\Psi_r \leftarrow \text{svd}(\hat{x})$
3: if $p == r$ then
4: $\gamma \leftarrow \text{pivot}(\Psi_r)$ \quad $[Q, R, \text{pivot}] = qr(\Psi_r)$
5: else if $p > r$ then
6: $\gamma \leftarrow \text{pivot}(\Psi_r, \Psi_r^\top)$ \quad $[Q, R, \text{pivot}] = qr(\Psi_r, \Psi_r^\top)$
7: \quad $b \leftarrow \hat{x}_\gamma$ \quad $\hat{x}_\gamma = \hat{x}_{l:1:1:p}$

After $b$ is determined as in Algorithm 1, we make use of (7) such that:

$$T_{\gamma}(b) = \hat{x} = [\hat{x}_1, ..., \hat{x}_p]^\top. \quad (18)$$

V. NUMERICAL RESULTS

We carry out simulations on the 2,000-bus synthetic Texas system [35] comprising 544 synchronous and 87 non-synchronous generators; the latter are modeled as variable-speed wind generators with full converters and do not contribute to the electromechanical modes. A three-phase short-circuit is applied to Bus 1017 and cleared after 10 milliseconds. Fig. 2 shows the rotor angle of the synchronous generators relative to the system average angle, with removed mean. We observe that 456 of 544 synchronous generators present coherent dynamics, whereas 88 of 544 synchronous generators do not contribute at all to the dynamics; see constant line at zero degrees in Fig. 2.

Upon application of the KMD explained in Section III, the 10 most important Koopman eigenvalues—i.e., the ones with the smaller damping coefficient—are shown in Fig. 3. They were estimated using 10 seconds of rotor speed estimates. The black crosses indicate results obtained with all 544 estimates obtained from the decentralized DSE (assuming we have local PMU measurements), whereas the blue circles indicate results obtained with 240 principled selected estimates from the sparse state measurement selection technique explained in Section IV. The remaining 5 seconds of the time domain simulation are used for testing. We consider three different scenarios, as presented next.

1) No outliers: The results are shown in Fig. 4a. True state, received state estimate, and filtered state estimate, respectively, refer to the outcome of the time-domain simulation, the data received from the robust model-based decentralized DSEs, and the outcome of the robust KKF.

2) Impulsive noise: The results are shown in Fig. 4b. We observe that the robust KKF is able to suppress the impulsive noise.

3) Loss of the communication link with a decentralized DSE: The results are shown in Fig. 4c. This case demonstrates the forecasting capability of the robust KKF.

VI. CONCLUSIONS AND FUTURE RESEARCH

The proposed hybrid framework offers a balance between model-based and data-driven methods, and it has several important advantages compared to other methodologies. It is completely independent of the network model, does not attempt to model the dynamics of the loads, and, most importantly, does not depend on the infrastructure that is exclusively available at the control centers.
Although promising, the proposed approach requires further investigations. One important aspect is the system observability at the upper level, i.e., the sparse selection must be conditioned to a certain degree of redundancy. We will rely on the correspondence between the Koopman operator and the Lie derivatives to pursue this effort. In addition, the estimation of the KEFs [36] is an important open problem and must be addressed. Future work will also look into the applicability of using other observables from sensors in the upper level system identification by either partially or completely bypassing the lower level DSE in certain scenarios to gain increased agility.

REFERENCES