Data-Driven Modeling of Microgrid Transient Dynamics through Modularized Sparse Identification

Apoorva Nandakumar, Yan Li, *Senior Member, IEEE*, Honghao Zheng, *Senior Member, IEEE*, Junhui Zhao, Dongbo Zhao, *Senior Member, IEEE*, Yichen Zhang, *Senior Member, IEEE*, Tianqi Hong and Bo Chen

***Abstract*— Modularized sparse identification (M-SINDy) is developed in this paper for effective data-driven modeling of the nonlinear transient dynamics of microgrid systems. The high penetration of power-electronic interfaces makes microgrids highly susceptible to disturbances, causing severe transients, especially in the islanded mode. The M-SINDy method realizes distributed discovery of nonlinear dynamics by partitioning a higher-order microgrid system into multiple subsystems and introducing pseudo-states to represent the impact of neighboring subsystems. This specific property of the proposed algorithm is found to be very useful while working with re-configurable and scalable microgrids. The governing equations of the subsystems are identified through regression by mapping the nonlinear system's data to a linear system in a large functional space. The effectiveness of the M-SINDy method is tested and validated through numerical examples and comparisons with other existing identification models on a typical islanded microgrid. The method can simultaneously compute the governing equations of different subsystems and is examined to be robust to measurement noises and partial observations. This paper highlights the advances of data science in providing a potent tool for modeling and analyzing higher-order nonlinear microgrid systems. Dynamic discovery of system transients from measurements can be beneficial for designing control strategies that improve the overall microgrid stability and reliability.**

***Index Terms*— Sparse identification, Modularized design, Data-driven modeling, Distributed energy resources (DERs), Microgrids, Transient dynamics, Psedo-states.**

I. Introduction

MICROGRIDS are self-sufficient energy systems that can provide a reliable power supply and enable better interconnection between multiple renewable energy resources [1], such as wind generation, photovoltaic, energy storage systems, etc [2]. It can be defined as a group of interconnected loads and distributed energy resources (DERs) that act as a single controllable entity [3]. They can operate in parallel with the main grid or as an autonomous island by controlling the DERs with power-electronics interfaces. Thus, microgrids have the flexibility to operate in different modes based on the distribution network and/ or economic requirements. The transient dynamics of microgrid needs to be investigated for the planning and operation of the system control structure as the operation modes of a microgrid are significantly different from that of a traditional power system.

The transient dynamics of microgrids are typically studied by modeling the system as a set of differential algebraic equations (DAEs) [4], [5]. While the system topology to develop the steady state algebraic constraint is easily obtainable, it is rather difficult to develop a high-fidelity model that represents the transient dynamics of the different DERs, and loads connected in the system. Also, online calibration of the control parameters in a large-scale system with high dimensions poses additional challenges. New DERs are integrated into existing microgrids to improve the overall system performance and support increased load requirements. The addition of new components to an already convoluted microgrid requires modification of the existing DAEs. These modifications in the physical system necessitate the development of new model equations.

With the latest advancements in the power industry, operational time series data can be easily obtained. The abundance of data over time can be exploited to develop a data-driven model which has the flexibility to be modified as needed. The time series data can be corrupted by external noise. Thus, it is essential for the identification method to be robust to external noises.

The characteristics of a system that is being identified can be categorized based on multiple factors. The dynamic behavior of a system can be distinguished as linear and non-linear system based on its differential state equations. Certain identification techniques can be employed if the various functions that define the differential state equations are sparse in the set of all possible functions. Some identification techniques may depend on both the state variables data and the output variables data. The dimensionality of the data obtained from the system is also very crucial to develop a data driven identification technique. The ability of the identification model to adapt to system changes in real time is also crucial. Based on this discussion, typical examples of data-based identification techniques from various literature are summarized below.

Identification of linear systems: The subspace method is a

This work is supported by National Science Foundation under the award DMS-2229435 and Office of Naval Research under the award N00014-22-1-2504.

A. Nandakumar and Y. Li are with the Department of Electrical Engineering, The Pennsylvania State University, University Park, PA, 16802 U.S. (e-mail: aqn5335@psu.edu, yql5925@psu.edu).

H. Zheng and B. Chen are with Commonwealth Edison Company, Chicago, IL, 60197 U.S.

J. Zhao is with Eversource Energy, Berlin, CT, 06037 U.S.

D. Zhao is with Global Technology, Eaton Corporation, Golden, CO, 80401 U.S.

Y. Zhang is with the Department of Electrical Engineering, The University of Texas at Arlington, Arlington, TX, 76010 U.S.

T. Hong is with Argonne National Laboratory, Lemont, IL, 60439 U.S.

classic example that uses input and output data for identifying linear time-invariant systems [6]. Its attractive features can be listed as the ability to model multiple-input multiple-output systems with a simple parametrization and present robust non-iterative numerical solutions [7]. However, this method is limited to linear dynamical systems and cannot be extended to nonlinear systems, such as a complex microgrid. Dynamic mode decomposition with control (DMDc) [8] is also capable of using high dimensional data for system identification. This method also works with the assumption of linear dynamics and cannot provide a good fit for non-linear identification.

Identification of non-linear systems: A typical method using Koopman operator is employed for non-linear system identification. This method converts the nonlinear system so that it can be represented in terms of linear functions. Koopman theory in combination with DMDc [9] has seen a lot of developments in recent years. A model-free method using Koopman operator approximation has been proposed in [10] for power system models. A linear approximation of the nonlinear system independent of state variables is obtained using this method. This method has various advantages relating to stability studies [11] and trajectory prediction [12]. But, it is designed to be employed for offline applications due to its computational complexity. Non-linear models based on machine learning concepts, i.e. neural network based system identification [13], [14], are also becoming increasingly popular with the advances in computing power [15], [16]. Model predictive control along with deep reinforcement techniques [17], [18] yield precise results in the large data limit. A combination of a binary tree algorithm with nonlinear autoregressive with exogenous input identification has been discussed in [19] for modeling nonlinear loads in power systems. However, machine learning and deep learning based algorithms often do not result in closed-form expressions that are easy to interpret. Another common challenge with these algorithms arise from the overfitting issue [20].

Considering the sparse nature of the equations describing the power system models, sparse sampling and dimensionality reduction based method has been studied to characterize and model the nonlinear dynamical systems in [21]. This method has the capability to correctly identify the dynamical parameter regime and reconstruct the full state dynamics. Sparse identification of nonlinear dynamics (SINDy) has been proposed in [22]–[25]. This algorithm can utilize high-dimensional data without having to bear the ill effects of overfitting. SINDy with control was also developed as a sustainable alternative for the online identification of non-linear dynamics in response to rapid system changes [26]. It has been found that this method can adapt to new test cases in real time and identify a model which is robust to external noises. Despite the promising approach of these methods to identify the nonlinear transient dynamics, scaling to higher dimensional models increases the computational cost and effort significantly. Inspired by the aforementioned studies, this paper focuses on extending the concept of SINDy to develop a modularized SINDy (M-SINDy), which can identify the nonlinear dynamics of microgrids with reduced computational effort. **The novelty of this paper are listed below:**

(1) Modularized SINDy (M-SINDy) algorithm has been developed which can significantly improve the computational efficiency while studying higher order microgrid models. The entire system is broken into multiple subsystems which can run parallelly with a smaller computational effort. This method also introduces the concept of pseudo-state variables that has the ability to represent multiple state and output variables and largely reduce the dimensionality of the system. This method proves to be an extremely useful tool while modeling adaptive microgrids with re-configurable structures.

(2) A data-driven model is developed to identify the nonlinear transient dynamics of microgrids, originally described by DAEs. The developed algorithm can model the microgrid dynamics in the form of Ordinary Differential Equations (ODEs). The effectiveness of the proposed algorithm has been numerically verified in the paper.

(3) The paper also investigates the robustness of M-SINDy algorithm with noisy measurements and partial system observations. It discusses and demonstrates that the effective selection of candidate functions for M-SINDy is vital to extend the algorithm for online applications and development of system controls.

The remainder of this paper is organized as follows. Section II introduces the proposed M-SINDy method and explains the concept of pseudo-state variables. It also includes derivations corresponding to the theoretical modeling of the microgrid DAEs. Section III details the steps required to implement the proposed algorithm. Numerical examples are provided in Section IV to verify the effectiveness of the proposed method with multiple test cases. Comparison results with a commonly used identification technique are also described in this section. Conclusions and future work are discussed in Section V.

II. Modularized Sparse Identification (M-SINDy) for Modeling Nonlinear Dynamics

The essential idea of M-SINDy is to partition a large-scale nonlinear dynamical system into small-scale subsystems without sacrificing the model accuracy by accounting for the unused states with the help of pseudo-state variables. The pseudo-state variables are introduced to represent the impact of the neighboring subsystems. The algorithm is used to identify the governing equations of the subsystems through the measurement data obtained from the physical system [27], [28]. The distribution system's operating data is available in abundance as a result of the widely installed advanced metering infrastructure. M-SINDy can be parallelly implemented on these subsystems to discover their non-linear transient dynamics. The method is developed based on the fact that the governing equations are sparse in a high dimensional functional space, i.e., only a few relevant terms are used to determine each subsystem.

A. Transient Modeling of Nonlinear Microgrids

Transient stability refers to the ability of power generation units in a microgrid to remain synchronized under credible disturbances. Transient dynamics of microgrids are mainly governed by the variations in Distribution Energy Resources (DERs) and loads. The averaged model is used to the represent the corresponding power-electronic interfaces of the DERs. The transient dynamics of the microgrid model can be mathematically described by a set of differential-algebraic equations (DAEs) [1],

$$\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \mathbf{y}(t), \mathbf{u}(t)\big), \tag{1a}$$
$$\mathbf{0} = \mathbf{g}\big(\mathbf{x}(t), \mathbf{y}(t), \mathbf{u}(t)\big), \tag{1b}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state variable vector, e.g., state variables in the controller of DER power-electronic interfaces, $\mathbf{y} \in \mathbb{R}^m$ is the algebraic variable vector, e.g., bus voltage amplitude and angle, and $\mathbf{u} \in \mathbb{R}^p$ represents the input variations/ disturbances, e.g., power output fluctuation of PV and power load changes. The specifics of the microgrid model used in this paper is provided in the appendix.

In power systems, the network power flow serves as a constraint of the overall system's transient dynamics. Thus, the overall system can be represented as a set of complicated ordinary differential equations (ODE) which includes the algebraic constraints i.e. the algebraic part shown in Eq. (1(b)) can be visualized as a constraint of Eq. (1(a)) and can be merged to form a set of complicated ODEs. In order to derive this comprehensive ODEs from the original DAEs, we obtain the time series Taylor expansion of Eq.1(b) as,

$$\mathbf{0} = \frac{\partial \mathbf{g}}{dt}\Delta t + \frac{1}{2}\cdot\frac{\partial^2 \mathbf{g}}{dt^2}\Delta t^2 + \mathcal{O}(\Delta t^3), \tag{2}$$

To theoretically obtain a continuous time differential equation for the overall system, we can consider a high sampling rate with a very small value for the Δt. Thus, the higher order Δt terms can be ignored from the Taylor series expansion to obtain Eq. (3) which can be further expanded to form Eq. (4).

$$\frac{\partial \mathbf{g}(\mathbf{x}(t), \mathbf{y}(t), \mathbf{u}(t))}{dt} = \mathbf{0} \tag{3}$$

$$\frac{\partial \mathbf{g}}{\partial \mathbf{y}(t)}\dot{\mathbf{y}}(t) + \frac{\partial \mathbf{g}}{\partial \mathbf{x}(t)}\dot{\mathbf{x}}(t) + \frac{\partial \mathbf{g}}{\partial \mathbf{u}(t)}\dot{\mathbf{u}}(t) = \mathbf{0} \tag{4}$$

Eq. (4) can be rewritten to compute the differential form of the algebraic variables $\boldsymbol{y}(t)$ as,

$$\dot{\mathbf{y}}(t) = -\Big(\frac{\partial \mathbf{g}}{\partial \mathbf{y}(t)}\Big)^{-1}\Big(\frac{\partial \mathbf{g}}{\partial \mathbf{x}(t)}\dot{\mathbf{x}}(t) + \frac{\partial \mathbf{g}}{\partial \mathbf{u}(t)}\dot{\mathbf{u}}(t)\Big) \tag{5}$$

From the perspective of microgrid modeling, $\frac{\partial \mathbf{g}}{\partial \mathbf{y}}$ in Eq. (5) represents the network admittance matrix, which is typically a non-singular matrix. The work in [29] provides additional details to justify this statement.

$$\dot{\mathbf{y}}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{y}(t), \mathbf{u}(t)) \tag{6}$$

$$\dot{\mathbf{z}}(t) = [\mathbf{f}(t), \mathbf{h}(t)] \tag{7}$$

$$\dot{\mathbf{z}}(t) = \mathbf{F}(\mathbf{z}(t), \mathbf{u}(t)) \tag{8}$$

The $\dot{\mathbf{z}}(t)$ in Eq. 8 represents the overall differential equations of the microgrid system where $\mathbf{z}(t) = [\mathbf{x}(t) \quad \mathbf{y}(t)]$. Thus, we can conclude that the transient dynamics of the microgrid can be modeled as an ODE without loss of generality, since the algebraic constraint in Eq. (1(b)) has been absorbed into the differential part.

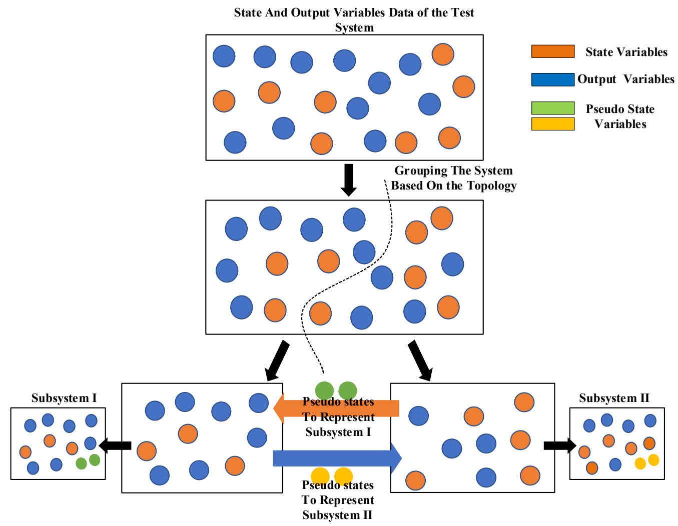


Fig. 1. Pictorial representation of M-SINDy algorithm for system identification

B. Preparing Data for Identifying the Governing Equations using M-SINDy

The basic idea behind the proposed modularized data-driven modeling algorithm is illustrated by Fig. 1, in which the identification of two subsystems is given as an example. The first step towards modularized identification is to prepare the time series data of the state variables, algebraic variables and input disturbance variables. $\mathbf{z_j}(t) \in \mathbb{R}^q$ represents the measurement of the state and algebraic variables ($\mathbf{z}(t) = [\mathbf{x}(t), \mathbf{y}(t)]$) pertaining to the subsystem $\mathbf{j}$. This data is collected using a high resolution measurement system and can be used to compute the vector derivatives, $\dot{\mathbf{z}}_j(t)$ analytically or numerically. In this work, $\dot{\mathbf{z}}_j(t) \approx \frac{\Delta \mathbf{z}_j}{\Delta t}$ is used to approximate the numerical values of the derivatives that represent the overall modified ODE of the original microgrid DAE. Here, Δt is the sampling time. Additionally, the data of the induced disturbances, i.e $\mathbf{u_j}(t)$, is collected in order to define the pool data which is built based on the candidate functions defined specifically for the microgrid system.

The measurement data of the module sampled over the given time period $t = [t_1, t_2, \cdots, t_k]$ can be organized as shown in Eq. (9). Here, $\mathbf{X_j}$ represents the time series values of each variable in the subsystem at individual time instances $t = t_1, t_2, \cdots, t_k$ and $\mathbf{z}_\mathbf{j}^\mathsf{T}(t_1) = [\mathbf{z_1}(t_1) \quad \mathbf{z_2}(t_1) \quad \cdots \quad \mathbf{z_q}(t_1)]$. The time series data of the pseudo-state variables, $\mathbf{v_j}(t) \in \mathbb{R}^l$ is shown by $\mathbf{v}_\mathbf{j}^\mathsf{T}(t_1) = [\mathbf{v}_1(t_1) \quad \mathbf{v}_2(t_1) \quad \cdots \quad \mathbf{v}_l(t_1)]$. Correspondingly, the derivatives $\dot{\mathbf{X}}_\mathbf{j}$ can be computed to perform the modularized sparse regression for system identification.

The $\mathbf{X}$ represents the time series measurements collected for performing the system identification. It includes $\mathbf{x}(t), \mathbf{y}(t)$ and $\mathbf{v}(t)$. The $\mathbf{x}(t)$ represents the state variables in the model.

$$\mathbf{X_j} = \begin{bmatrix} \mathbf{z}_\mathbf{j}^\mathsf{T}(t_1) & \mathbf{z}_\mathbf{j}^\mathsf{T}(t_2) & \cdots & \mathbf{z}_\mathbf{j}^\mathsf{T}(t_k) \\ \mathbf{v}_\mathbf{j}^\mathsf{T}(t_1) & \mathbf{v}_\mathbf{j}^\mathsf{T}(t_2) & \cdots & \mathbf{v}_\mathbf{j}^\mathsf{T}(t_k) \end{bmatrix}^\mathsf{T} \tag{9}$$

In microgrids, the system's state variables ($\mathbf{x}(t)$), bus volt-

This article has been accepted for publication in IEEE Transactions on Sustainable Energy. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2023.3273127

age magnitudes and phase angles ($\mathbf{y}(t)$), and line currents ($\mathbf{v}(t)$) can be measured to build the data matrix for system identification. The authors recommend using the line currents as pseudo-state variables since the current over the distribution feeders reflect the impact of neighboring subsystems.

The transient dynamics are driven by the input disturbances and the data regarding these disturbances, applied at different time intervals, can be prepared as

$$\mathbf{U}_j = [\mathbf{u_j}(t_1) \quad \mathbf{u_j}(t_2) \quad \cdots \quad \mathbf{u_j}(t_k)]^\mathsf{T} \tag{10}$$

The proposed modular design of the system identification enables focus on the variables corresponding to the subsystem of interest. The other subsystems can be reduced to only hold the value of the pseudo-states, e.g., the current flowing through the bus connecting the subsystems. This pseudo-state variable can be used in addition to the subsystem's states to represent the overall system. The method is advantageous as it can perform subsystem specific study and minimize the overall computational effort.

C. Establishing Non-linear Candidate Functions

The objective of sparse identification is to map the non-linear dynamics of the microgrid system to a set of candidate functions with linear coefficients using regression. The non-linear state equations corresponding to the transient dynamics of the original system can be converted into a high dimensional linear functional space. The library of candidate functions to realize this can be formulated as,

$$\begin{array}{lccccc} \Theta(\mathbf{X_j}, \mathbf{U_j}) = [& \mathbf{1} & \mathbf{X_j} & \mathbf{X_j}^{P^2} & \cdots & \\ & \sin(\mathbf{X_j}) & \cos(\mathbf{X_j}) & \mathbf{U_j} & \mathbf{X_j}\mathbf{U_j} & \cdots \\ & \sin(\mathbf{U_j}) & \cos(\mathbf{U_j}) & \cdots & \mathbf{X_j}\sin(\mathbf{X_j}) & \cdots] \end{array} \tag{11}$$

The candidate functions to identify the transient dynamics of a microgrid model is chosen according to Eq. (11) based on,

- The microgrid transient dynamics is dependent on the variations in the loads connected to the system. The complex ZIP [30] power load models are non-linear in nature due to the quadratic terms of bus voltages and line currents.
- The inverter control equations of the DER are defined in the dq framework and the conversion of frames from abc to dq model involves sinusoidal functions.

The pool data matrix is defined such that the number of candidate functions (columns of $\Theta(\mathbf{X_j}, \mathbf{U_j})$) is significantly lower than the number of data samples (rows of $\Theta(\mathbf{X_j}, \mathbf{U_j})$). The sparcity property of the microgrid transient dynamics equations enables proper tracking of these equations with fewer functions. Thus, the restricted basis of non-linear functions shown in Eq. (11) can be used for system identification.

The polynomial terms in the candidate functions library are given by $\mathbf{X_j}, \mathbf{X_j^{P^2}}, \cdots, \mathbf{X_j^{P^n}}$. The trigonometric terms in the candidate library are represented as $\sin(\mathbf{X_j}), \cos(\mathbf{X_j}), \cdots$. Examples for the expanded time series candidate functions are given by Eq. (12) and Eq. (13), where $\mathbf{X_j^{P^2}}$ represents quadratic polynomial function and $\sin(\mathbf{X_j})$ represents first order trigonometric function.

$$\mathrm{X}_j^{P^2} = \begin{bmatrix} \mathbf{z_{j_1}}^2(t_1) & \mathbf{z_{j_1}}(t_1)\mathbf{z_{j_2}}(t_1) & \cdots & \mathbf{z_{j_2}}^2(t_1) & \cdots & \mathbf{z_{j_n}}^2(t_1) \\ \mathbf{z_{j_1}}^2(t_2) & \mathbf{z_{j_1}}(t_2)\mathbf{z_{j_2}}(t_2) & \cdots & \mathbf{z_{j_2}}^2(t_2) & \cdots & \mathbf{z_{j_n}}^2(t_2) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{z_{j_1}}^2(t_k) & \mathbf{z_{j_1}}(t_k)\mathbf{z_{j_2}}(t_k) & \cdots & \mathbf{z_{j_2}}^2(t_k) & \cdots & \mathbf{z_{j_n}}^2(t_k) \end{bmatrix} \tag{12}$$

$$\sin(\mathbf{X_j}) = \begin{bmatrix} \sin(\mathbf{z_{j_1}}(t_1)) & \sin(\mathbf{z_{j_2}}(t_1)) & \cdots & \sin(\mathbf{z_{j_n}}(t_1)) \\ \sin(\mathbf{z_{j_1}}(t_2)) & \sin(\mathbf{z_{j_2}}(t_2)) & \cdots & \sin(\mathbf{z_{j_n}}(t_2)) \\ \vdots & \vdots & \ddots & \vdots \\ \sin(\mathbf{z_{j_1}}(t_k)) & \sin(\mathbf{z_{j_2}}(t_k)) & \cdots & \sin(\mathbf{z_{j_n}}(t_k)) \end{bmatrix} \tag{13}$$

An example of a second order candidate function representing the impact of input disturbance is given in Eq. (14).

$$\mathrm{X}_jU = \begin{bmatrix} \mathbf{u}_1(t_1)\mathbf{z_{j_1}}(t_1) & \mathbf{u}_1(t_1)\mathbf{z_{j_2}}(t_1) & \cdots & \mathbf{u}_1(t_1)\mathbf{z_{j_n}}(t_1) \\ \mathbf{u}_1(t_2)\mathbf{z_{j_1}}(t_2) & \mathbf{u}_1(t_2)\mathbf{z_{j_2}}(t_2) & \cdots & \mathbf{u}_1(t_2)\mathbf{z_{j_n}}(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_1(t_k)\mathbf{z_{j_1}}(t_k) & \mathbf{u}_1(t_k)\mathbf{z_{j_2}}(t_k) & \cdots & \mathbf{u}_1(t_k)\mathbf{z_{j_n}}(t_k) \end{bmatrix} \tag{14}$$

D. Sparse Regression for Modular Identification

The final step of system identification is the sparse regression of the pool data to identify the governing equations of the subsystems. This step determines the function given in Eq. (8) which pertains to the subsystem in study. The microgrid dynamics can be modeled using only a few terms from the set of all non-linear candidate functions. This sparse nature of the microgrid system equations can be leveraged to obtain the linear combination of non-linear terms that would accurately describe the system's governing equations using the regression method.

If the sparse vector coefficients identified using regression is given by $\Xi = \begin{bmatrix}\xi_1 & \xi_2 & \cdots & \xi_s\end{bmatrix}$, then $\dot{\mathbf{X}}$ can be represented as,

$$\dot{\mathbf{X}}_\mathbf{j} = \Xi\Theta^\mathsf{T}(\mathbf{X_j}, \mathbf{U_j}), \tag{15}$$

Here, $\mathbf{U_j}$ represents the external disturbances that drives the subsystem dynamics. $\mathbf{X_j} = [\mathbf{Z_j} \quad \mathbf{V_j}]^T$ where $\mathbf{Z_j}$ is the time series data of the state and algebraic variables and $\mathbf{V_j}$ is the time series data of the pseudo-state variables. To identify the vector Ξ, multiple regression methods can be used. Ordinary Least Squares (OLS) regression, Ridge regression, Least Operator Shrinkage and Selection Operator (LASSO) are some commonly used regression techniques. The LASSO-type optimization problems can be solved by using various proximal Newton methods which are also used for solving convex composite optimization problems [31]. This paper utilizes the LASSO regression technique which is outlined by Eq. (16) in which Λ gives the sparsity constraint.

This article has been accepted for publication in IEEE Transactions on Sustainable Energy. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2023.3273127

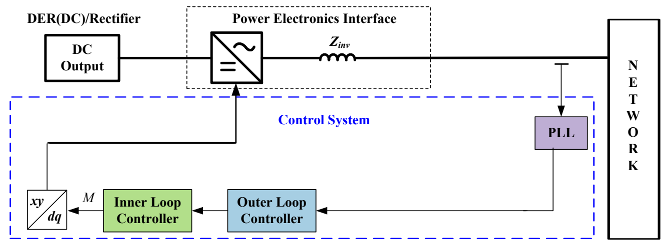


Fig. 2. A simplified one DER based microgrid model with inverter controls

$$\xi = \arg\min_{\xi'} ||\boldsymbol{\Theta}\xi' - \boldsymbol{x}||_2 + \lambda||\xi'||_1, \tag{16}$$

Empirical studies shows that not all state variables would impact the entire microgrid dynamics. Hence, this sparse nature of the system can be leveraged to develop a sparse identification based data-driven model for microgrids. This has been verified with multiple test cases in Section IV.

III. Algorithm to Implement M-SINDy

The data set for the original SINDy based identification algorithm requires the information regarding all the state variables, output variables and input disturbances to model the entire system dynamics. The proposed M-SINDy method can be used to identify only a part of the overall dynamics which is of interest to the system operator. This property is specifically useful in the power systems domain while working with the re-configurable and scalable microgrids. The algorithm utilizes the state variables, output variables (algebraic variables) and input disturbance variables corresponding only to a section of the entire system. The variables in the unused section of the system can be combined to create pseudo-state variables which reduces the dynamical order of the system without compromising on the features of the overall system. The step-by-step implementation of the proposed M-SINDy is given by Algorithm 1.

Algorithm 1: Algorithm to implement M-SINDy

Data:
$\mathbf{z} \leftarrow$ state variables & algebraic variables
$\mathbf{P}^0 - \mathbf{P}^n \leftarrow$ polyorder
$\mathbf{n} \leftarrow$ order of the system
$\mathbf{u} \leftarrow$ input disturbances
usesine $\leftarrow$ 1 or 0 (defines use of trigonometric candidate functions)
Result:
$\Xi \times \Theta \leftarrow$ identify the set of differential equations that define the transient system dynamics

```
Step 1: Select the time series state and algebraic data
  required based on the method of type of identification
  method used.
if Centralized SINDy then
    z = set of all the state and algebraic variables
    X = [Z  U]^T
else
    if Modular SINDy then
        z = set of the state and algebraic variables
          corresponding to the chosen module
        define v as the set of pseudo-state variables to
          represent the impact of the unused state and
          algebraic variables
        X = [Z  V  U]^T
    end if
end if
Step 2: Compute the derivatives of X since F(X) = Ẋ
while i ← 1 : t do
    dx = (x(i+1) − x(i)) / dt
end while
Step 3: Build the pool data based on the candidate
  functions defined for the system identification method
  Θ(X) ← library of all the non-linear functions based
  on the inputs given (x, u, n, usesine, polyorder).
Step 4: Sparse Regression
Ξ = arg min_Ξ' ||ΘΞ' − X||_2 + λ||Ξ'||_1
Step 5: Compute the identified dynamics as
  Ẋ_id = Ξ × Θ
Integrate X_id to obtain the identified data X
Step 6: Obtain the root mean square to compare the
  identified system with ground truth
```

IV. Numerical Examples

A typical microgrid system shown in Fig. 3. This is used to verify the effectiveness of the proposed method for modeling the transient dynamics of microgrids using measurement data. The test system includes 35 buses, 5 DERs, and 7 constant power loads. The microgrid is modeled to operate in the islanded mode with one grid forming DER and four grid following DERs. The grid forming DER uses Vf control strategy and the grid following DERs uses PQ control strategy. The detailed block diagrams denoting the control strategies are discussed in the appendix.

The proposed method can be applied to constant power load, constant current load and constant impedance load models. Constant power loads are adopted in this paper since it can contribute more significantly to the system's transient dynamics. Transients are also introduced to the test system by inducing an approximated step change in the power references of the PQ-controlled DERs.

The training data is prepared by modeling and simulating the test system in MATLAB as a set of DAEs with added white noise. In practice, the training data set can be obtained from metering devices. The following two test cases have been studied in detail.

- A simple system with only one grid forming DER has been simulated to explain the nuances of the M-SINDy algorithm developed in this paper. A simplified one DER model is shown in Fig. 2.
- A more complicated system with five DERs has been simulated to verify the effectiveness of scaling the M-SINDy algorithm to higher order systems.

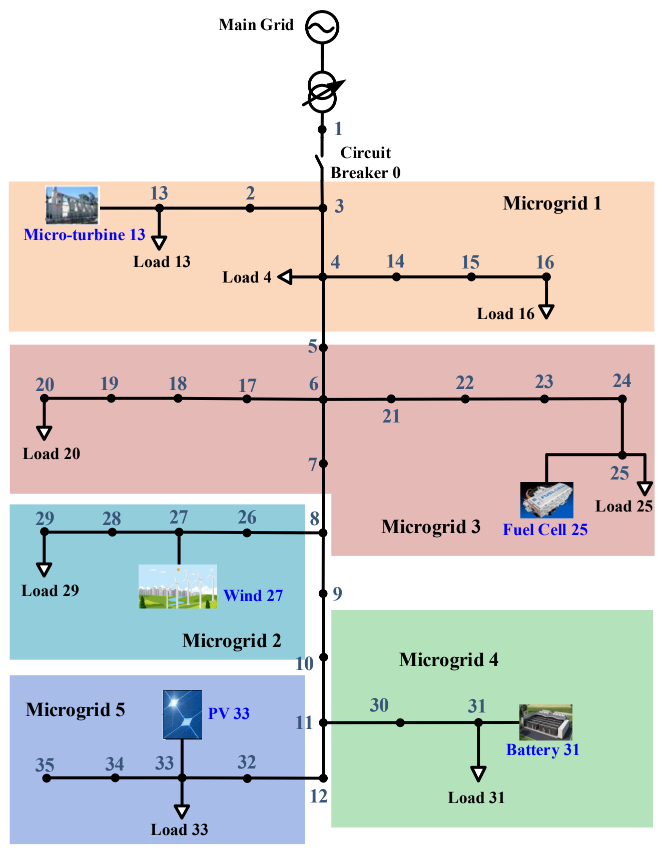


Fig. 3. A typical microgrid test system with multiple distributed energy resources and loads in islanded mode.

A. Case I: Data-driven modeling of system with one DER

1) Time domain trajectories and error analysis: The microgrid system modeled has the initial voltage reference set to 1.0 p.u. and the initial power load set to 0.27399 p.u. at $t = 0.0s$. This system is perturbed to change the voltage reference value from 1.0 p.u. to 0.9 p.u. at $t = 0.5s$. The second perturbation decreases the active power load by 30% of its initial value at $t = 1.0s$. The third disturbance increases the active power load by 30% of its initial value at $t = 1.5s$. The final disturbance brings back the power load to its initial value, but changes the voltage reference to 1.1 p.u. These changes in the control inputs of the system contributes to significant variations in the system's transient dynamics. A sampling time of 10 ms has been recorded for this test case. Typically, smart meters have a sampling rate of 1 Hz -1000 Hz [32] and the data for identification in this work had a small sampling time for improved accuracy (10 ms).

The inclusion of the disturbances in power loads ($\mathbf{u_1}(t)$) and the variations in the voltage references ($\mathbf{u_2}(t)$), can be leveraged to track the transient system dynamics. Fig. 4 shows the voltage magnitude and phase angle corresponding to the bus to which the micro-turbine based DER (grid-forming DER) is connected in the one DER test system.

The V-f controller is designed to synchronize and stabilize the frequency of the entire microgrid at 60 Hz. The induced power load disturbances can cause oscillations in the frequency before reaching the steady state value. The figure showing

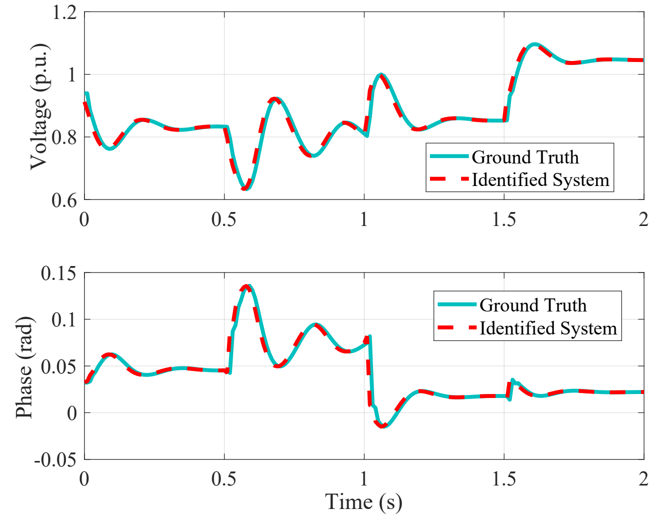


Fig. 4. Comparison of ground truth and identified bus voltage magnitude and phase angle

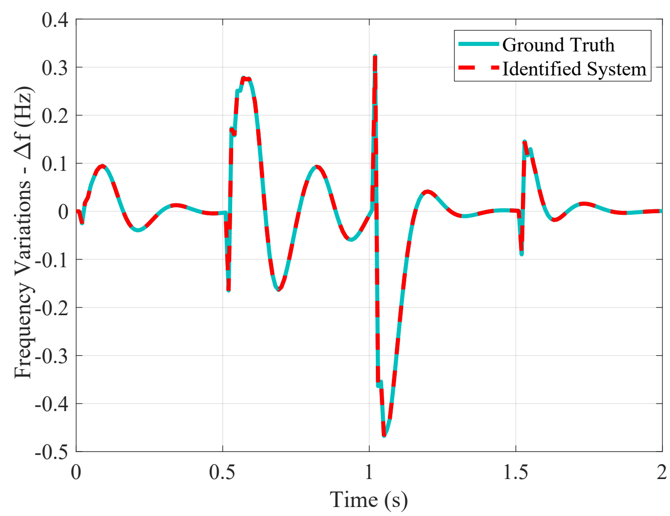


Fig. 5. Comparison of ground truth and identified frequency variations in the grid forming DER

these variations in the frequency of the grid forming DER is shown in Fig. 5.

- The comparisons elucidates the exact tracking performance of the M-SINDy algorithm by comparing the voltage and frequency data obtained from the true microgrid model with the identified model.
- The root mean square error between the identified model and the true microgrid model was computed to be in the range of $[7.95e^{-03}\% \quad 0.0835\%]$ with an average error value of $\pm 0.0125\%$.

2) Effect of polyorders on system identification: The candidate functions used to identify the system comprises of constant values, polynomials of different orders, and trigonometric terms. In this test, the dependence of the quality of identification on the different orders of the polynomial terms has been discussed. Fig. 6 shows the transient dynamics of true and identified states 5 and 8 of the microgrid model

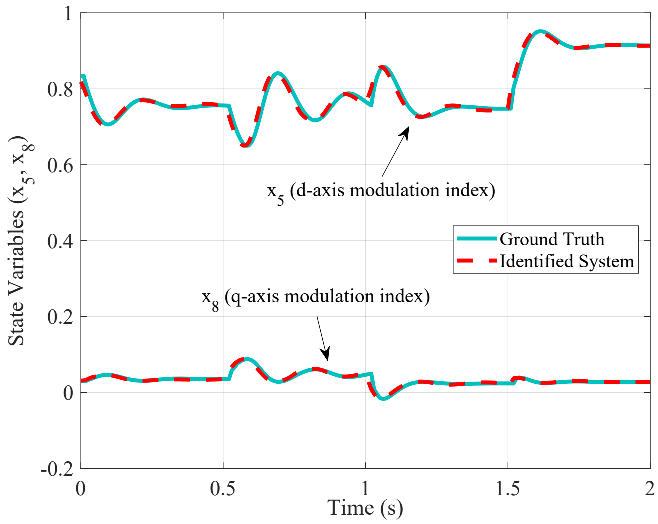


Fig. 6. Comparison of ground truth and identified state variables using polyorder 1 in the M-SINDy algorithm

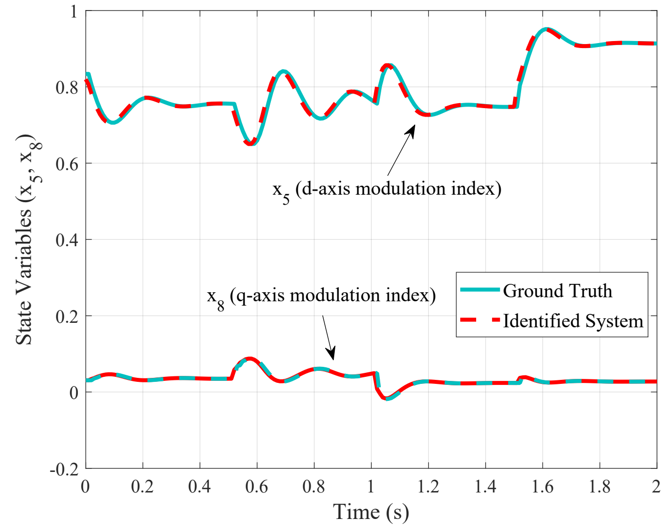


Fig. 7. Comparison of ground truth and identified state variables using polyorder 2 in the M-SINDy algorithm

using the M-SINDy algorithm with polyorder 1. These states represent the inverter modulation indices in the d and q frame respectively. Fig. 7 is the representation of the same state variables which has been identified using polyorder 2. We can see that,

- By comparing the errors between the identified states and the true states in both cases, it has been verified that the nonlinearity of the microgrid model is represented better by increasing the polyorder of the candidate function.
- While the average root mean square error for the case with polyorder 1 was $\pm 0.0125\%$, the average root mean square error for the case with polyorder 2 was found to be $\pm 0.0053\%$.
- The mean square error significantly reduced as the polyorder, used to track the nonlinearity in the system, was increased. This indicates that higher polyorders are desired for practical applications.

3) Trigonometric functions based Identification: In this test, the candidate functions chosen to represent the nonlinearity of the microgrid system consists only of trigonometric terms, i.e., sine and cosine functions of higher orders to represent the entire system. Any periodic function can be represented as the sum of sine and cosine terms. Fig. 8 validates this statement as the identified model consists only of trigonometric terms and closely follows the actual system dynamics which comprises of both polynomials and trigonometric functions.

4) Noisy data test: In practice, the sensor data obtained in real-time is susceptible to be corrupted with some noise. In order to understand the robustness of the M-SINDy algorithm and its effectiveness to identify the original system with such sensor noises, a small amount of white Gaussian noise with a maximum of 10% signal to noise ratio (SNR) was introduced to the true system data. This corrupted data was propagated through the identification algorithm and the results obtained verifies teh robustness of the proposed method. Fig. 9 shows the voltage magnitude and the phase data identified from the true system corrupted with noisy measurements (20 dB SNR).

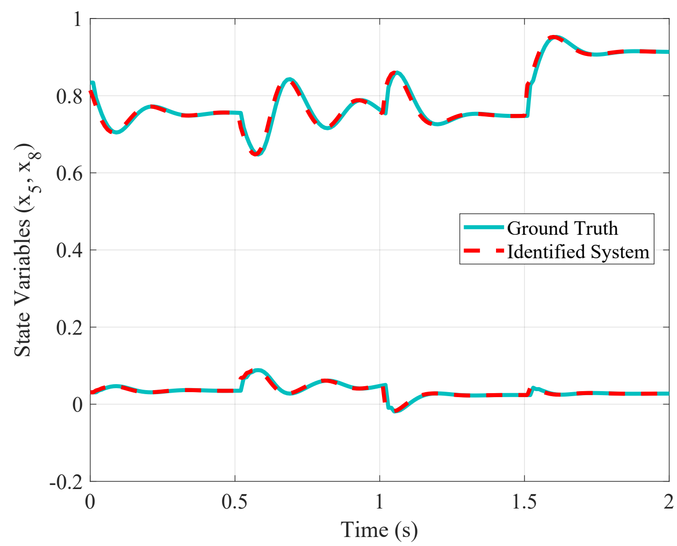


Fig. 8. Comparison of ground truth and identified system using only trigonometric functions

B. Case II: Data-driven modeling of system with five DERs

1) Time domain trajectories and error analysis: In order to verify the effectiveness of the proposed M-SINDy algorithm for identifying higher order microgrid models, a test case system with 5 DERs was simulated. The size of the system drastically increased in comparison to the previously tested microgrid system with one DER. The new system comprises of 40 state variables, which includes 8 state variables corresponding to each one of the 5 DERs, i.e, micro-turbine, fuel cell, photovoltaic system, wind and battery. The system is perturbed with two distinct disturbances at different time intervals. A sampling time of 10 ms has been recorded for this test case. The active power load is increased from its original value by 5% at $t = 0.0s$. Following this, at $t = 1.5s$, the active power load is decreased by 10% from its original value. At $t = 3.0s$, the active power load is set to its original value and

This article has been accepted for publication in IEEE Transactions on Sustainable Energy. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2023.3273127

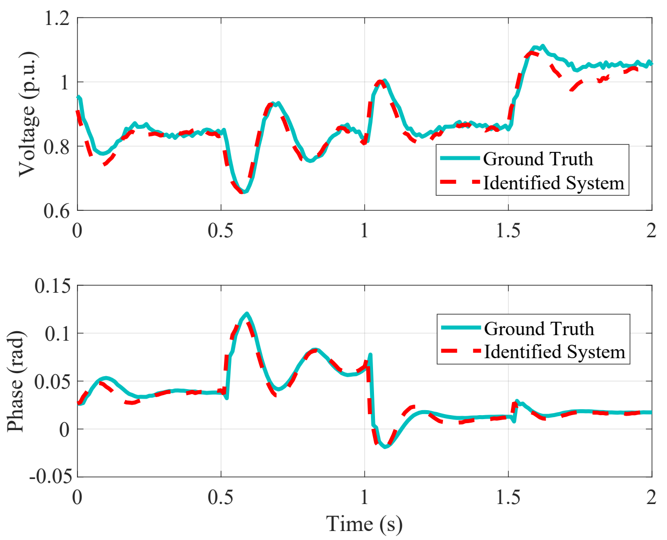


Fig. 9. Comparison of ground truth and identified system when the ground truth is corrupted with gaussian white noise

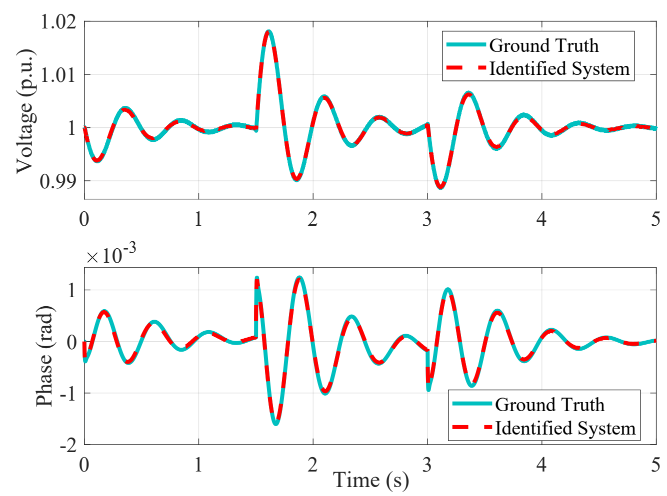


Fig. 10. Comparison of ground truth and identified bus voltage and phase angle corresponding to the grid forming micro-turbine based DER

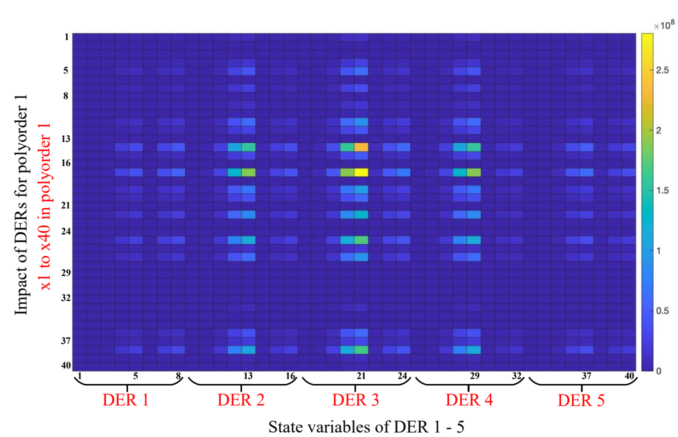


Fig. 11. Color-map detailing the dominant dynamics in the identified system

the power generation of the battery based DER is increased by 10%. The initial per unitized value of the battery power generation is 0.021. These variations in the control inputs are incorporated in the pool data to track the system changes.

- The comparison validates the sparse identification based system modeling in successfully converting the DAEs based model into an ODEs based model.
- To further illustrate the effectiveness of the M-SINDy algorithm in identifying the true system dynamics, the voltage plots have been obtained and are shown in Fig. 10. The magnitude and phase of the bus 31 voltage are tracked very closely by the proposed algorithm and the root mean square error between the original system and the identified system was found to be around $\pm 2.33e^{-03}\%$.
- The detailed colormap representing the dominant dynamics of the 5 DER microgrid test system is shown in Fig. 11. It can be verified that the dominant states correspond to the state variables 4,5,7 and 8 of each DER. These states define the d-axis and q-axis modulation indices of the connected inverter. Theoretically, these are the state variables that experience maximum variations in their dynamics. Thus, it is ideal for identified Ξ matrix to follow a similar pattern.

2) Noise data test: As mentioned in the previous section, all the data collected from the sensor in real-time are susceptible to external noise. In order to ensure that the M-SINDy based identification method can be employed for a higher order system corrupted by noise, a gaussian white noise was manually introduced to the simulated data. The results of the addition of additive white gaussian noise (AWGN) to the simulated voltage data are shown in Fig. 12. The voltage magnitude and phase shown in this figure are well tracked by the ODEs identified using the M-SINDy algorithm. Thus, the proposed identification method would prove to be ideal for cases where the measured data is possibly corrupted by external noise.

3) Identification with partial observation: One of the assumptions in the sparse identification technique is the availability of all the state variables and the output variables. In reality, it would be possible to obtain all the data regarding state variables only if these states are observable and measurable. Hence, it is vital to show that the proposed method can identify the transient dynamics using only the data regarding the output variables which are usually measurable. In this case, internal states data is not accessible and hence, only the bus voltage magnitudes and phases are used as the measurement data to identify the transient dynamics in the bus voltages of the system. The output voltages correspond to the algebraic part of the DAEs. The M-SINDy algorithm can accurately model the algebraic constraint as a differential equation. The dynamics of the bus voltages are a reflection of the overall dynamics of the system. The voltages can be treated as a top layer which is determined by the state variables in the hidden layer.

Fig. 13 corresponds to the bus 31 voltage magnitude and phase data. The system dynamics have been traced using only the voltage data. If the internal state data is not an observable state, it could cause inaccuracies while designing the control

This article has been accepted for publication in IEEE Transactions on Sustainable Energy. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2023.3273127

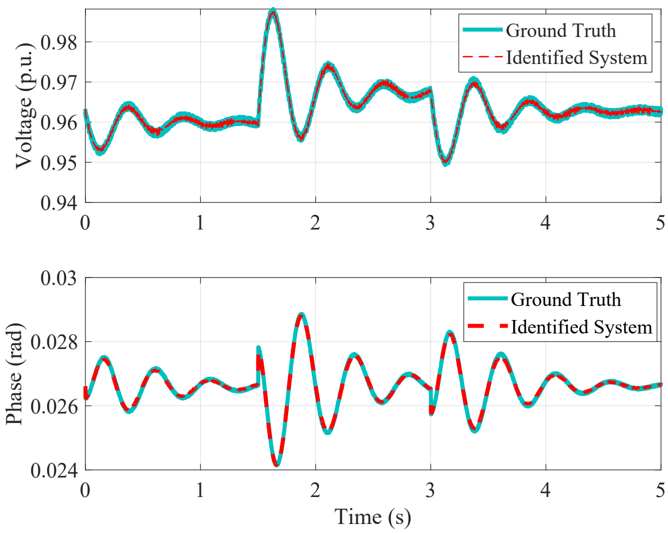


Fig. 12. Comparison of ground truth and identified system when the ground truth is corrupted with gaussian white noise

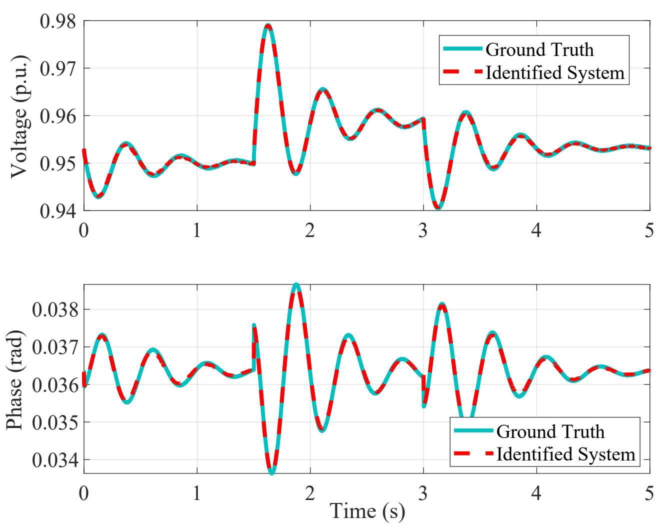


Fig. 13. Comparison of ground truth and identified system with partial observation

algorithm for the exact output tracing. This method overcomes the need to obtain all the internal states data for system identification.

4) Subsystem Identification: In this test, the entire microgrid model is split into two subsystems. The details regarding the split subsystems are shown in Fig. 14. The model has been broken into two subsystems at the line connecting buses 7 and 8. Subsystem 1 consists of 2 DERs, 5 loads and 18 buses. Subsystem 2 consists of 3 DERs, 3 loads and 15 buses. If the load/ DER of interest belongs to subsystem 1, the entirety of the dynamics corresponding to subsystem 2 can be reduced to the current flowing between bus 8 and bus 7 (I_{78}). The dynamics of 24 state variables and 15 algebraic variables can be represented by this current flow. This significantly reduces the model complexity. The current flow between the buses at the point of partition can be computed based on the knowledge

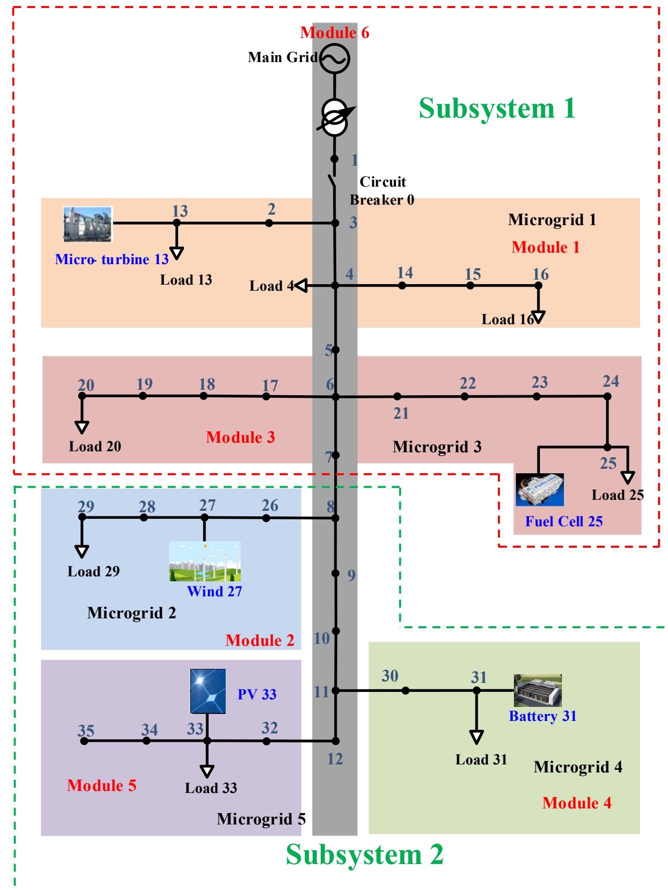


Fig. 14. Representation of a microgrid test system that can be reconfigured into multiple modules for identification

of the bus voltages and the line impedance between these buses. The dynamics of the neighboring subsystems can be understood by the current flow between the two subsystems at the point of partition. Hence, the line current data can be used as the pseudo state variable while modeling a single subsystem using M-SINDy algorithm. Examples of the bus voltages belonging to subsystem 1 and subsystem 2 are shown in Fig. 15 and Fig. 16 respectively.

5) Modular Identification: In this example, the microgrid test system has been broken into multiple modules. The splitting of the test system into 6 modules (5 modules corresponding to each DER, 1 module corresponding to the network backbone connecting the DERs) is shown in Fig. 14. Each of these modules can be parallelly identified using the M-SINDy algorithm which significantly minimizes the computational effort. Table I outlines the computational time required for identification based on the number of modules. Fig. 17 has 6 subplots with each plot representing one of the state variables from the 6 modules.

C. Comparison of data driven modeling techniques

Additional tests were performed to further elucidate the accuracy of the proposed M-SINDy algorithm in comparison to the existing data-driven identification methods. The data provided in the one DER test system was propagated through

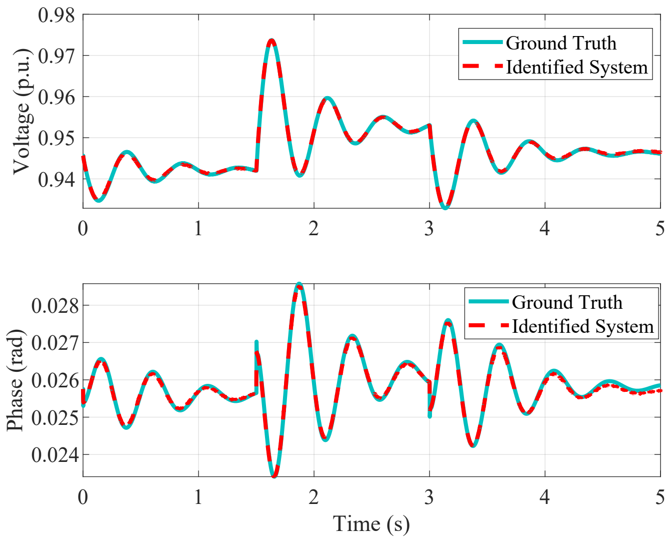


Fig. 15. Bus voltage magnitude and phase comparison - subsystem 1

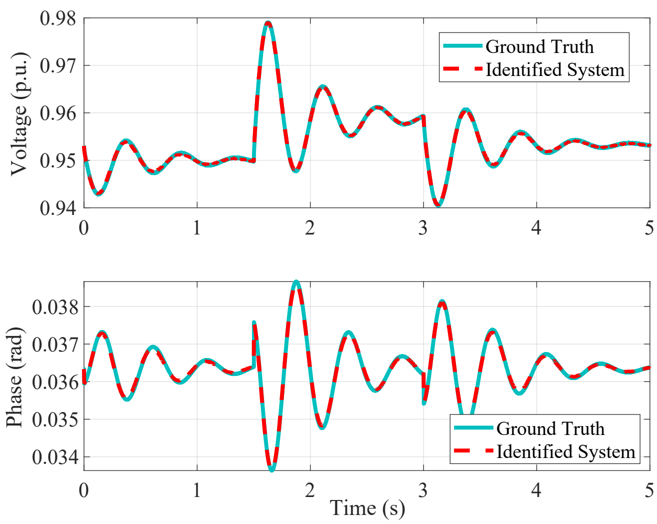


Fig. 16. Bus voltage magnitude and phase comparison - subsystem 2

the layers of the neural network based NARX algorithm and the results are shown in Fig. 18 and Table II. The overall system accuracy, mean square error and computational time comparisons show that the proposed method performs a better system identification with lesser computational effort. The final representation of the identified model as an ODE using the M-SINDy algorithm provides a good correlation with the physical system. NARX algorithm did not provide a similar closed-form expression which is easy to interpret.

V. Conclusions and Future Work

M-SINDy has been developed in this paper to identify the nonlinear transient dynamics of higher order microgrid systems in a data-driven fashion. The higher order system can be decoupled into several small-scale subsystems and the concept of pseudo-states was introduced to identify the governing equations using measurement data. Numerical examples

TABLE I

Comparison of Calculation Time between M-SINDy and SINDy.

Identification method	Polyorder 1	Polyorder 2
Modularized-SINDy (6 modules)	1.355 s	99.166 s
Modularized-SINDy (2 modules)	3.645 s	286.455 s
Centralized-SINDy (1 module)	6.222 s	319.297 s

TABLE II

Comparison of the elapsed time and RMS error calaculation for system identification

Method	Error Percentage	Time (s)
NARX	1.6524 %	2.7524
M-SINDy	0.0265 %	0.2285

validated the effectiveness of M-SINDy through time domain simulations, identification with random white Gaussian noise and partial measurements. It has also been verified that the true dynamical system which is typically modeled as a DAE can be represented as a system comprising only ODEs.

The M-SINDy algorithm can be extended to predict the system dynamics and incorporate new model predictive control algorithms that can significantly improve the transient stability of the system.

VI. Acknowledgement

The authors would like to thank Dr. J. Nathan Kutz at the University of Washington for sharing the base code for Sparse Identification of Non-linear Dynamics with control.

Appendix

The details regarding the modeling of the 5 DER based test system used in the paper is given below. The system topology details are shown in Tables III, IV, V. The one line diagram of the system is shown by Fig. 19 which gives the steady state algebraic values of the buses connected to the inverter based resources. The averaged model of the inverter has been utilized in this work.

TABLE III

DER Generations At Each Bus

Bus	$P_n(kW)$	$Q_n(kVAR)$
13	136.02	74.85
25	24.04	33.62
27	25.68	15.75
31	21.42	21.78
33	26.06	24.48

TABLE IV

Power Loads At Each Bus

Bus	$P_n(kW)$	$Q_n(kVAR)$
4	32.69	15.97
13	54.69	21.26
16	21.56	10.64
20	59.63	38.57
25	40.54	20.63
29	61.35	37.59
31	38.21	16.78
33	65.31	39.45

This article has been accepted for publication in IEEE Transactions on Sustainable Energy. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2023.3273127

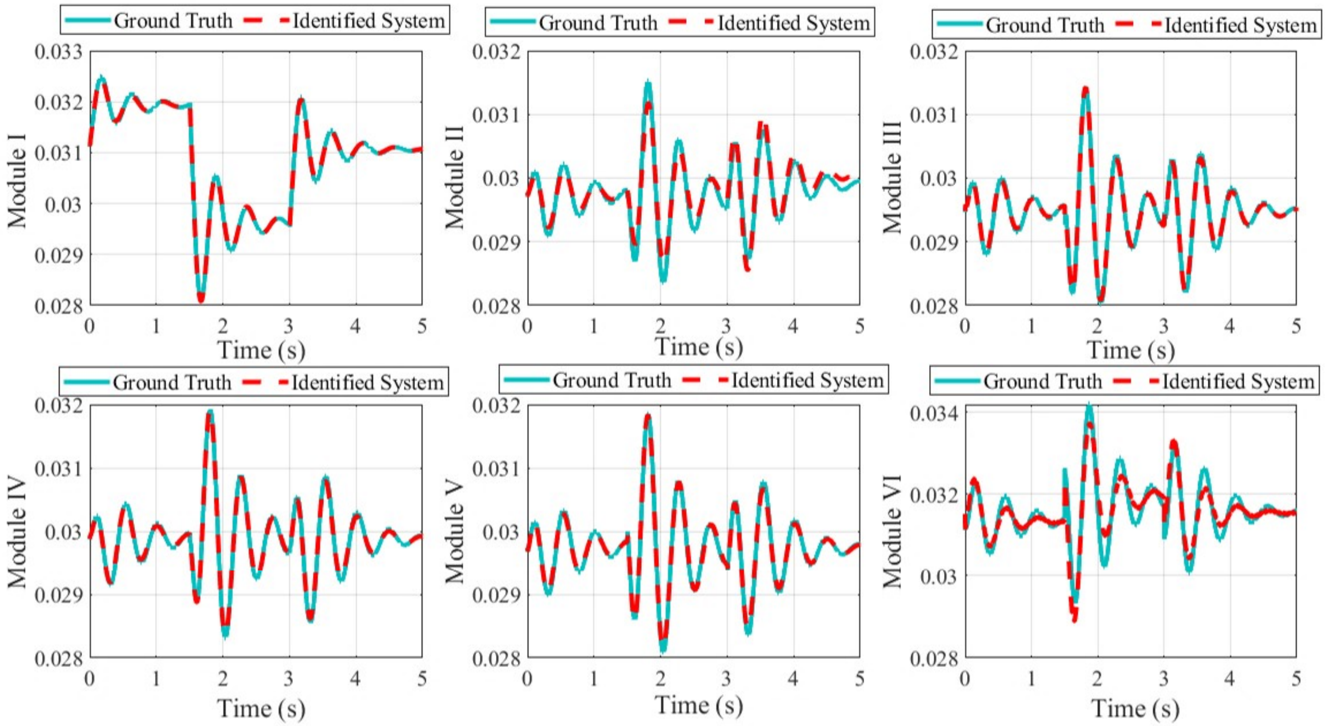


Fig. 17. The comparison between the ground truth and the identified system's states from the 6 modules

TABLE V
LINE IMPEDANCE BETWEEN BUSES

From	To	R(Ω/km)	L(H/km)	Length(m)
3	4	0.0682	0.5422×10^{-3}	45
13	2	0.005	0.1133×10^{-3}	30
2	3	0.005	0.1133×10^{-3}	30
4	14	0.1824	1.474×10^{-3}	50
14	15	0.0422	0.3220×10^{-3}	50
15	16	0.1510	1.684×10^{-3}	50
4	5	0.1678	1.662×10^{-3}	45
5	6	0.4645	3.740×10^{-3}	20
6	17	0.4645	3.740×10^{-3}	20
17	18	0.0952	0.8800×10^{-3}	20
18	19	0.2453	1.6400×10^{-3}	30
19	20	0.1423	1.1600×10^{-3}	30
6	21	0.0640	0.5933×10^{-3}	30
21	22	0.0640	0.5933×10^{-3}	30
22	23	0.0426	1.9622×10^{-3}	45
23	24	0.0480	5.900×10^{-3}	40
24	25	0.0480	2.2075×10^{-3}	40
6	7	0.0476	1.1675×10^{-3}	40
7	8	0.0423	1.0378×10^{-3}	45
8	26	0.0423	2.4133×10^{-3}	45
26	27	0.3040	3.6200×10^{-3}	30
27	28	0.2026	1.3200×10^{-3}	45
28	29	0.6563	3.600×10^{-3}	30
8	9	0.3530	1.8433×10^{-3}	30
9	10	0.3530	1.8433×10^{-3}	30
10	11	1.2613	2.8100×10^{-3}	30
11	30	0.5040	2.8267×10^{-3}	30
30	31	1.1786	2.8267×10^{-3}	30
11	12	1.1786	2.7200×10^{-3}	30
12	32	0.4394	3.0900×10^{-3}	30
32	33	0.5722	3.0900×10^{-3}	30
33	34	0.5722	1.8267×10^{-3}	30
34	35	0.3406	0.9633×10^{-3}	30

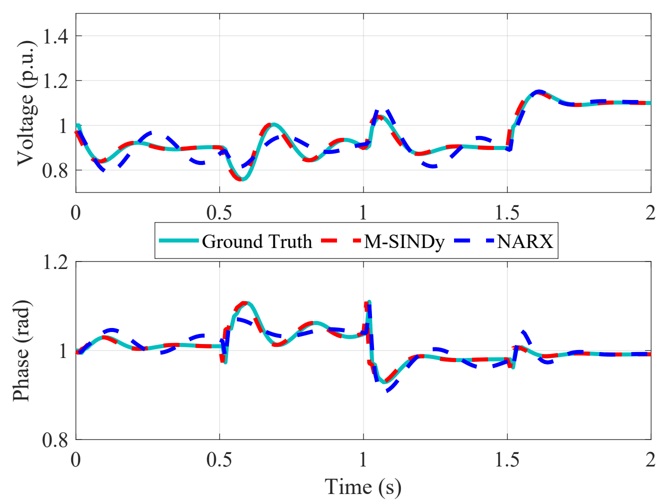


Fig. 18. Comparison plots to explain the performance of the different identification methods

The system has two specific control strategies corresponding to the grid-forming DER (V-f control) and grid-following DERs (P-Q control). Block diagrams and equations explaining the details of the double-loop controller for the grid following and grid forming DERs are shown in Fig. 20.

The differential equations ($\mathbf{f}(\mathbf{x}(t), \mathbf{y}(t), \mathbf{u}(t))$) that describe the dynamics of the control system for the grid-forming DER in Fig. 20 is given by (17)-(24). The differential equations ($\mathbf{f}(\mathbf{x}(t), \mathbf{y}(t), \mathbf{u}(t))$) that describe the dynamics of the control system for the grid-following DER in Fig. 20 is given by (25)-

This article has been accepted for publication in IEEE Transactions on Sustainable Energy. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2023.3273127

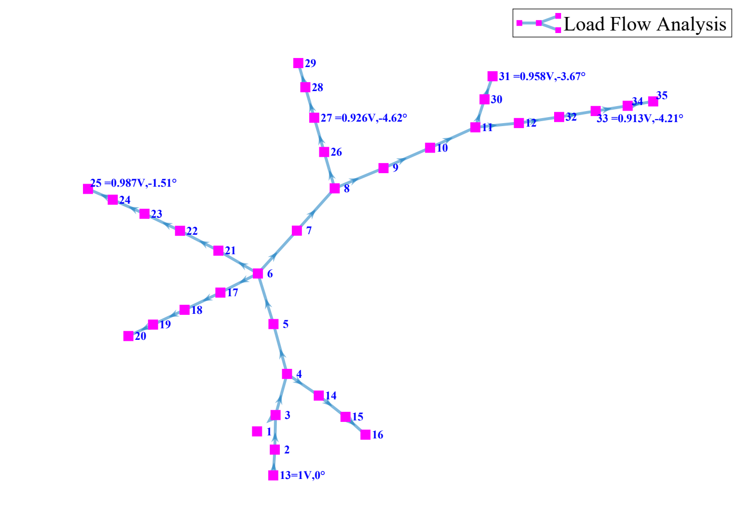


Fig. 19. System Topology

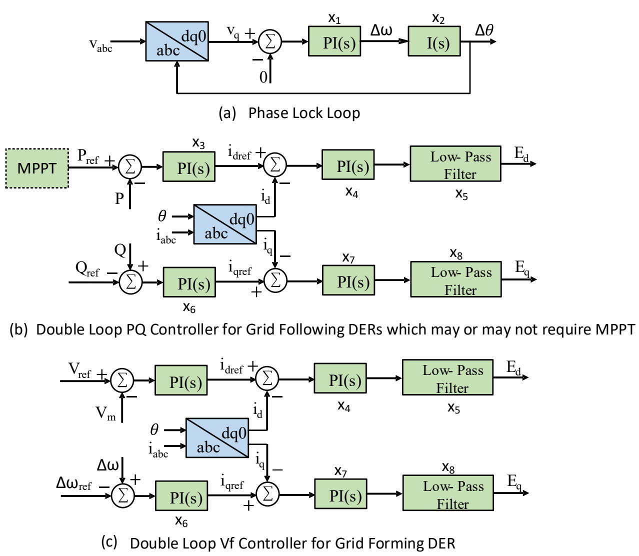


Fig. 20. Double loop controller for grid forming and grid following DERs with/ without MPPT depending on the DER capability to dispatch active power

(32). The algebraic equation, $\mathbf{g}(\mathbf{x}(t), \mathbf{y}(t), \mathbf{u}(t))$, describes the network power flow.

$$\dot{x_1} = V_q - \bar{V}_q \quad (17)$$
$$\dot{x_2} = \Delta\omega \quad (18)$$
$$\dot{x_3} = \bar{V}_m - \sqrt{|V_d|^2 + |V_q|^2} \quad (19)$$
$$\dot{x_4} = \bar{\Delta\omega} - \Delta\omega \quad (20)$$
$$\dot{x_5} = \bar{I}_d - I_d \quad (21)$$
$$\dot{x_6} = \bar{I}_q - I_q \quad (22)$$
$$\dot{x_7} = (\bar{V}_d^{(\mathrm{inv})} - x_7)/T_f \quad (23)$$
$$\dot{x_8} = (\bar{V}_q^{(\mathrm{inv})} - x_8)/T_f \quad (24)$$

$$\dot{x_1} = V_q - \bar{V}_q \quad (25)$$
$$\dot{x_2} = \Delta\omega \quad (26)$$
$$\dot{x_3} = P_{ref} - P \quad (27)$$
$$\dot{x_4} = Q_{ref} - Q \quad (28)$$
$$\dot{x_5} = \bar{I}_d - I_d \quad (29)$$
$$\dot{x_6} = \bar{I}_q - I_q \quad (30)$$
$$\dot{x_7} = (\bar{V}_d^{(\mathrm{inv})} - x_7)/T_f \quad (31)$$
$$\dot{x_8} = (\bar{V}_q^{(\mathrm{inv})} - x_8)/T_f \quad (32)$$

Here, V_d and V_q represent the d-axis and q-axis bus voltages at the point of interconnection between the inverter and the network. $\Delta\omega$ is the change in frequency of the system calculated by the PLL design. T_f is the time constant used in the low pass filter.

REFERENCES

[1] R. Lasseter, "Microgrids," in *2002 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.02CH37309)*, vol. 1, 2002, pp. 305–308 vol.1.

[2] C. Wang, Y. Li, K. Peng, B. Hong, Z. Wu, and C. Sun, "Coordinated optimal design of inverter controllers in a micro-grid with multiple distributed generation units," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2679–2687, 2013.

[3] Y. Li, *Cyber-Physical Microgrids*, 1st ed. Springer, Cham, 2021.

[4] G. Kariniotakis, N. Soultanis, A. Tsouchnikas, S. Papathanasiou, and N. Hatziargyriou, "Dynamic modeling of microgrids," in *2005 International Conference on Future Power Systems*, 2005, pp. 7 pp.–7.

[5] P. N. Papadopoulos, T. A. Papadopoulos, and G. K. Papagiannis, "Dynamic modeling of a microgrid using smart grid technologies," in *2012 47th International Universities Power Engineering Conference (UPEC)*, 2012, pp. 1–5.

[6] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory—Implementation—Applications*. Springer Science & Business Media, 2012.

[7] S. J. Qin, "An overview of subspace identification," *Computers & chemical engineering*, vol. 30, no. 10-12, pp. 1502–1513, 2006.

[8] J. L. Proctor, S. L. Brunton, and J. N. Kutz, "Dynamic Mode Decomposition with Control," *SIAM Journal on Applied Dynamical Systems*, vol. 15, no. 1, pp. 142–161, 2016. [Online]. Available: https://doi.org/10.1137/15M1013857

[9] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data–driven approximation of the koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, jun 2015.

[10] X. Li, C. Mishra, S. Chen, Y. Wang, and J. De La Ree, "Determination of parameters of time-delayed embedding algorithm using koopman operator-based model predictive frequency control," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 6, pp. 1140–1151, 2021.

[11] Y. Susuki and I. Mezić, "Nonlinear koopman modes and power system stability assessment without models," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 899–907, 2014.

[12] P. Sharma, B. Huang, V. Ajjarapu, and U. Vaidya, "Data-driven identification and prediction of power system dynamics using linear operators," in *2019 IEEE Power Energy Society General Meeting (PESGM)*, 2019, pp. 1–5.

[13] Y. Wang, "A new concept using lstm neural networks for dynamic system identification," in *2017 American Control Conference (ACC)*, 2017, pp. 5324–5329.

[14] H. Bevrani, B. Francois, and T. Ise, "Microgrid Dynamics and Control," 2017.

[15] P. Werbos, "Neural networks for control and system identification," in *Proceedings of the 28th IEEE Conference on Decision and Control,*, 1989, pp. 260–265 vol.1.

[16] R. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*, vol. 4, no. 2, pp. 4–22, 1987.

[17] Y. Tange, S. Kiryu, and T. Matsui, "Model predictive control based on deep reinforcement learning method with discrete-valued input," in *2019 IEEE Conference on Control Technology and Applications (CCTA)*, 2019, pp. 308–313.

[18] Q. Chen, R. Long, S. Quan, and L. Zhang, "Nonlinear Recurrent Neural Network Predictive Control for Energy Distribution of a Fuel Cell Powered Robot," *The Scientific World Journal*, vol. 2014, p. 509729, 2014. [Online]. Available: https://doi.org/10.1155/2014/509729

[19] M. G. Jahromi, S. D. Mitchell, G. Mirzaeva, and D. Gay, "A new method for power system load modeling using a nonlinear system identification estimator," *IEEE Transactions on Industry Applications*, vol. 52, no. 4, pp. 3535–3542, 2016.

[20] U. Fasel, E. Kaiser, J. N. Kutz, B. W. Brunton, and S. L. Brunton, "SINDy with Control: A Tutorial," *arXiv e-prints*, p. arXiv:2108.13404, Aug. 2021.

[21] S. Sargsyan, S. Brunton, and J. Kutz, "Nonlinear model reduction for dynamical systems using sparse sensor locations from learned libraries," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 92, p. 033304, 10 2015.

[22] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016. [Online]. Available: https://www.pnas.org/content/113/15/3932

[23] J. L. Proctor, S. L. Brunton, B. W. Brunton, and J. N. Kutz, "Exploiting sparsity and equation-free architectures in complex systems," *The European Physical Journal Special Topics*, vol. 223, no. 13, pp. 2665–2684, 2014. [Online]. Available: https://doi.org/10.1140/epjst/e2014-02285-8

[24] A. M. Stanković, A. A. Sarić, A. T. Sarić, and M. K. Transtrum, "Data-driven symbolic regression for identification of nonlinear dynamics in power systems," in *2020 IEEE Power Energy Society General Meeting (PESGM)*, 2020, pp. 1–5.

[25] A. Nandakumar, Y. Li, D. Zhao, Y. Zhang, and T. Hong, "Sparse identification-enabled data-driven modeling for nonlinear dynamics of microgrids," in *2022 IEEE Power Energy Society General Meeting (PESGM)*, 2022, pp. 1–5.

[26] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Sparse Identification of Nonlinear Dynamics with Control (SINDYc)**SLB acknowledges support from the U.S. Air Force Center of Excellence on Nature Inspired Flight Technologies and Ideas (FA9550-14-1-0398). JLP thanks Bill and Melinda Gates for their active support of the Institute of Disease Modeling and their sponsorship through the Global Good Fund. JNK acknowledges support from the U.S. Air Force Office of Scientific Research (FA9550-09-0174)." *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 710–715, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896316318298

[27] ——, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016. [Online]. Available: https://www.pnas.org/content/113/15/3932

[28] J. L. Proctor, S. L. Brunton, B. W. Brunton, and J. N. Kutz, "Exploiting sparsity and equation-free architectures in complex systems," *The European Physical Journal Special Topics*, vol. 223, no. 13, pp. 2665–2684, 2014. [Online]. Available: https://doi.org/10.1140/epjst/e2014-02285-8

[29] D. Turizo and D. K. Molzahn, "Invertibility conditions for the admittance matrices of balanced power systems," *IEEE Transactions on Power Systems*, pp. 1–12, 2022.

[30] A. Arif, Z. Wang, J. Wang, B. Mather, H. Bashualdo, and D. Zhao, "Load modeling—a review," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5986–5999, 2017.

[31] X. Li, D. Sun, and K.-C. Toh, "A Highly Efficient Semismooth Newton Augmented Lagrangian Method for Solving Lasso Problems," *SIAM Journal on Optimization*, vol. 28, no. 1, pp. 433–458, 2018. [Online]. Available: https://doi.org/10.1137/16M1097572

[32] J. Frolik, A. L. Lentine, A. Seier, and C. Palombini, "Dynamic communications control for grid agents," in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, 2012, pp. 1–5.

Apoorva Nandakumar received her B.Tech from the Department of Instrumentation and Control Engineering at National Institute of Technology, Tiruchirappalli, TamilNadu, India in 2019. She received her MS degree with specific focus on power electronics from the Department of Electrical Engineering at Penn State University in 2021. She is currently pursuing her Ph.D. degree in the Department of Electrical Engineering at Penn State University. Her research interests includes data driven modeling, stability analysis and control of networked microgrids and software defined networking.

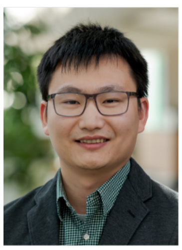
Yan Li earned her Ph.D. degree in Electrical Engineering from the University of Connecticut, Storrs, CT, U.S., in 2019. She also holds a Ph.D. degree in Electrical Engineering from Tianjin University, Tianjin, China, in 2013. She is currently an assistant professor at the School of Electrical Engineering and Computer Science at the Pennsylvania State University, University Park, PA, U.S. Her research interests include cyber-physical microgrids, quantum computing, data-driven modeling and control, stability and resilience analysis, cybersecurity, reachability, software-defined networking, etc.

Honghao Zheng (Senior Member, IEEE) received the B.S. degree from the Huazhong University of Science and Technology in 2009, and the M.S. and Ph.D. degrees from UW-Madison in 2012 and 2015, respectively. He worked for Siemens Industry Inc., in smart grid Research and Development, digital grid. He is currently the manager of grid strategy and analytics at Commonwealth Edison, an Exelon Company. He is also a registered Professional Engineer.

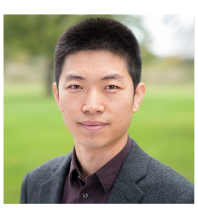
Junhui Zhao is the Manager of Advanced Engineering Analytics with Eversource Energy, Berlin, CT, USA, where he leads efforts to explore and demonstrate the applications of big data and AI technologies in power systems. Before joining Eversource, he was a Principal Quantitative Engineer at ComEd in 2022 and was an Associate Professor with tenure at the University of New Haven, CT, 2014-2021. He received the Ph.D. degree in electrical engineering from Wayne State University, Detroit, USA, in 2014.

Dongbo Zhao (SM' 16) is currently the Global Technology Manager with Eaton in its Eaton Research Lab. He was a Team Lead of DER Integration and a Principal Energy System Scientist with Argonne National Laboratory, Lemont, IL and also an Institute Fellow of Northwestern Argonne Institute of Science and Engineering of Northwestern University, before joining Eaton. His research interests include power system control, protection, reliability analysis, transmission and distribution automation, and electric market optimization. Dr. Zhao is a Senior Member of IEEE since 2016, and a member of IEEE PES, IAS and IES Societies. He has been the editor of IEEE Transactions on Power Delivery, IEEE Transactions on Sustainable Energy, and IEEE Power Engineering Letters.

Yichen Zhang (Senior Member, IEEE) received the B.S. degree in electrical engineering from Northwestern Polytechnical University, China, in 2010, the M.S. degree in electrical engineering from Xi'an Jiaotong University, China, in 2012, and the Ph.D. degree in electrical engineering from The University of Tennessee, Knoxville, in 2018. He was also a research assistant with the Oak Ridge National Laboratory from 2016 to 2018. From 2018 to 2022, he was with the Energy Systems Division, Argonne National Laboratory. He is now an assistant professor at The University of Texas at Arlington. His research interests include power system dynamics, grid-interactive converters, renewable integration, control and decision-making for cyber-physical power systems. He is an associate editor of IEEE Transactions on Power Systems and IEEE Power Engineering Letters.



Tianqi Hong (S'13–M'16) received the B.Sc. degree in electrical engineering from Hohai University, China, in 2011, and the M.Sc. degree in electrical engineering from the Southeast University, China and the Engineering school of New York University in 2013. He received a Ph.D. degree from New York University in 2016. His main research interests are power system analysis, power electronics systems, microgrid, and electromagnetic design. Currently, he is a Principal Energy System Scientist at Argonne National Laboratory. Prior to this, he was a Postdoc Fellow in the Engineering school of New York University and a Senior Research Scientist at Unique Technical Services, LLC, responsible for transportation electrification, battery energy storage integration, and medium capacity microgrid. Dr. Hong is an active reviewer in the power engineering area, and he serves as an Editorial Board Member of International Transactions on Electrical Energy Systems, IEEE Transactions on Power Delivery, IEEE Transactions on Industry Applications, and IEEE Power Engineering Letters. He also serves as Special Activity Co-Chair of the IEEE IAS Industrial Power Converters Committee (IPCC).



Bo Chen (S'12-M'17) is a Principal Engineer at Commonwealth Edison (ComEd), Oak Brook Terrace, IL, USA. He received the Ph.D. degree in electrical engineering from Texas AM University, College Station, USA, in 2017. He worked as a Principal Energy Systems Scientist at the Energy Systems Division, Argonne National Laboratory, IL, USA. His research interests include modeling, control, and optimization of power systems, cybersecurity, and cyber-physical systems. Dr. Chen is the Associate Editor of IEEE Transactions on Smart Grid.