Preprints.org

Article

On SINDy Approach to Measure-based Detection of Nonlinear Energy Flows in Power Grids with High Penetration IBR-based Renewables

Reza Saeed Kandezy [*] , John Jiang , Di Wu

Article

On SINDy Approach to Measure-based Detection of Nonlinear Energy Flows in Power Grids with High Penetration IBR-based Renewables

Reza Saeed Kandezy [1]*, John Jiang [2] and Di Wu [3]

[1] Department of Electrical and Computer Engineering, University of Oklahoma, Norman, OK, USA;
    reza.kandezy@ou.edu
[2] Department of Electrical and Computer Engineering, University of Oklahoma, Norman, OK, USA; jnjiang@ou.edu
[3] Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND, USA;
    di.wu.3@ndsu.edu
* Correspondence: reza.kandezy@ou.edu

Abstract: The complexity of modern power grids, exacerbated by integrating diverse energy sources, especially inverter-based resources (IBRs), presents a significant challenge to grid operation and planning since linear models fail to capture the intricate IBR dynamics. This study employs the Sparse Identification of Nonlinear Dynamics (SINDy) method to bridge the gap between theoretical understanding and practical implementation in power system analysis. It introduces the novel Volterra-based Nonlinearity Index (VNI) to examine system-level nonlinearity comprehensively. The distinction of dynamics into first-order linearizable terms, second-order nonlinear dynamics, and third-order noise elucidates the intricacy of power systems. The findings demonstrate a fundamental shift in system dynamics as power sources transit to IBRs, revealing system-level nonlinearity compared to module-level nonlinearity in conventional synchronous generators. The VNI quantifies nonlinear-to-linear relationships, enriching our comprehension of power system behavior and offering a versatile tool for distinguishing between different nonlinearities and visualizing their distinct patterns through the proposed VIN profile.

Keywords: Inverter-based resources; Measure-based method; Model identification; Non-linear dynamics; Power system; SINDy; Synchronous generators; System-level nonlinearity; Volterra-based nonlinearity index

1. Introduction

The systemic nonlinearity conundrum is a central focus in power system operation due to the intricate dynamics within modern power grids [1]. As these grids increasingly incorporate diverse energy sources, including inverter-based resources (IBRs), conventional linear models fail to capture the complex web of nonlinear interactions, feedback loops, and emergent behaviors [2]. The recognition and comprehensive examination of system-level nonlinearity are paramount for safeguarding the resilience, stability, and efficiency of contemporary electrical grids. The development and application of advanced modeling techniques, such as the Sparse Identification of Nonlinear Dynamics (SINDy), are indispensable in understanding the higher-order dynamics and nuanced interdependencies underpinning modern power systems [3].

A concise synthesis of studies emphasizes the transformative impact of IBR integration on power systems. A seminal work by Mishra et al. (2022) uncovers the intrinsic nonlinearities within conventional power grids, highlighting the challenge of maintaining stability amid dynamic interactions among synchronous generators and intricate load profiles [4]. Ekomwenrenren et al. (2021) empirically demonstrate deviations from conventional linearized frequency control in IBR grids, revealing nuanced nonlinear frequency responses unique to these systems [5]. Orihara et al. (2021) delve into the pivotal dynamics of virtual inertia in IBR grids, offering insights into nonlinear control mechanisms [6]. Keyon et al. (2020) examine the impact of varying IBR penetration levels on power system dynamics, illustrating the transition from first-order to high-order behavior [7]. Stankovic et al. (2020) provide a perspective on employing sparse identification techniques to capture

dynamics in power grids [8]. These studies collectively unveil the intricate contours of this nascent paradigm and the exigent hurdles it presents to the domain of power systems engineering.

Real-time support is instrumental in tackling emerging challenges and ensuring the adaptability of power systems as they evolve to accommodate various energy sources. Understanding the distinctions between synchronous generators (SGs) and the interactions within systems that combine SGs and IBRs or rely exclusively on IBRs is vital for system operators [9]. The most significant impacts in power system analysis and management materialize in real-time operations, where precise mathematical system dynamics models are often unattainable [10].

Measurement-based methodologies, rooted in empirical data and supported by advanced monitoring technologies and data analytics, provide a pragmatic solution for real-time decision support [11]. These techniques empower grid operators to make informed, agile decisions that ensure the continuity, stability, and reliability of electrical power systems amidst evolving energy landscapes characterized by nonlinear dynamics and dynamic grid architectures [11].

The emergence of advanced approaches has become a focal point in various fields due to their effectiveness as foundational frameworks. The transformative potential of innovative approaches that extend and refine the SINDy paradigm, making them applicable and relevant across diverse domains [12,13]. The development and adaptation of SINDy techniques

- provide a novel avenue for understanding complex systems,
- enrich our understanding of fundamental principles and
- pave the way for groundbreaking applications and insights in various fields of study [12,13].

This research presents a substantial contribution to power systems analysis and modeling. One of our primary contributions is applying the SINDy algorithm to the power system (IEEE 15-bus), unraveling the intricate dynamics that govern this complex system under various operational scenarios. Our investigation is a foundational step toward leveraging the Volterra-based Nonlinearity Index (VNI) as a novel proposed analytical instrument for quantifying nonlinearity within dynamic systems, providing a quantitative measure that enhances our understanding of the balance between linear and nonlinear behaviors in power grids.

Additionally, our study contributes to the ongoing discourse surrounding power grid dynamics, particularly in integrating inverter-based resources. By examining scenarios with synchronous generators and IBRs at varying levels of integration, the results offer insights into the nonlinear behaviors, at both module and system levels, that emerge when transitioning from traditional power generation to the integration of renewable energy sources. This contribution not only deepens the understanding of fundamental differences in system dynamics but also has practical implications for grid operators and planners aiming to optimize grid performance as renewable energy penetration continues to grow.

The introduction of higher-order polynomial function libraries to model IBR integration represents a significant departure from traditional modeling approaches, reflecting the evolving needs of power grid analysis as renewable energy takes center stage. The findings pave the way for a better understanding of the intricacies of power systems and offer practical solutions for building more resilient and efficient grids.

The subsequent sections of this manuscript are organized as follows: Section 2 provides a detailed explanation of the SINDy algorithm. Section 3 demonstrates the conducted study and the respective results followed by the introduction of the proposed index for system nonlinearity. The final segment comprises the concluding remarks, emphasizing the results' significance and elucidating potential research directions for future studies.

2. SINDy Algorithm

In this section, we will delve into the mechanics of the SINDy framework, exploring its fundamental principles and methodologies. We will begin by understanding how SINDy leverages sparsity methods, compressed sensing, and sparse regression techniques to identify concise and accurate models for complex dynamical systems. We will dissect the critical steps of the SINDy approach, from data collection and constructing a library of candidate functions to sparse regression

and model construction. Additionally, we will highlight the significance of sparsity in simplifying system dynamics and improving interpretability. Furthermore, we will discuss SINDy's applications, limitations, and the importance of adapting this methodology to the intricacies of electrical power systems. Finally, we will introduce a novel understanding and view of the SINDy method and its result by introducing a three-section data analysis structure, which extends SINDy's capabilities for enhanced data-driven research across diverse domains.

Integrating sparsity methods in analyzing dynamical systems has emerged as a significant advancement, employing compressed sensing and sparse regression techniques to identify concise and accurate models representing the underlying nonlinear dynamics [13]. SINDy is a measure-based method specifically designed to discover governing equations or mathematical models from observed data. The SINDy approach focuses on dynamical systems described by the equation

$$\frac{dx(t)}{dt} = f(x(t)),$$

(1)

where $x(t) \in R^n$ represents the system's state at time t and $f(x(t))$ encompasses the dynamic constraints governing the system's equations of motion, including parameters, time dependence, and external forcing [3].

Leveraging recent progress in compressed sensing and sparse regression, the sparsity perspective enables the identification of the nonzero terms in f without computationally demanding brute-force searches. Convex methods that scale well with Moore's law allow for identifying sparse solutions with high probability, striking a balance between model complexity and accuracy, thereby avoiding overfitting the model to the available data [14]. An example case is illustrated in Figure 1 to demonstrate SINDy's algorithm [3].
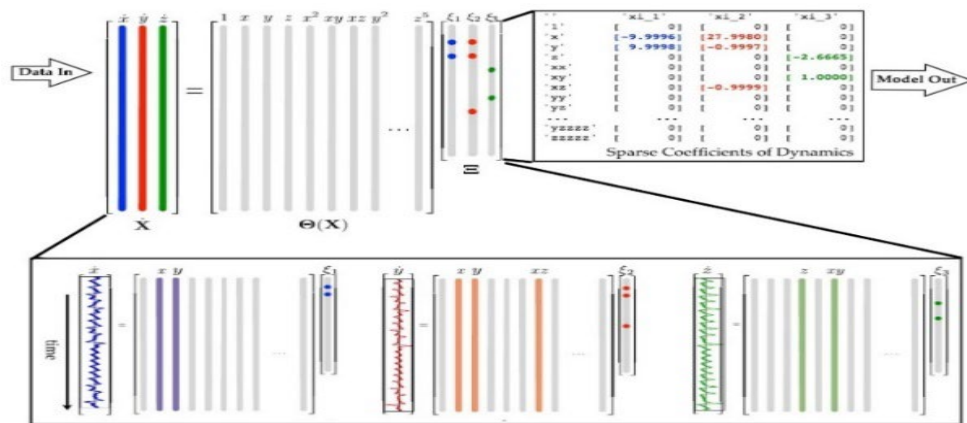

Figure 1. Example illustration of SINDy algorithm [3].

To determine the function f from available data, a time history of the system's state, denoted as $x(t)$, is collected. The derivative of $x(t)$, denoted as $\dot{x}(t)$, is directly or numerically approximated. The data is sampled at various time instances, $\{t_1, t_2, \dots, t_m\}$ and organized into X and $\dot{X}$ matrices. The matrix X is constructed as:

$$X = \begin{bmatrix} | & | & & | \\ x(t_1), x(t_2), \dots, x(t_m) \\ | & | & & | \end{bmatrix}^T$$

(2)

and the matrix $\dot{X}$ is constructed as:

$$\dot{X} = \begin{bmatrix} | & | & & | \\ \dot{x}(t_1), \dot{x}(t_2), \dots, \dot{x}(t_m) \\ | & | & & | \end{bmatrix}^T$$

(3)

The next step in the SINDy approach involves defining a library of candidate nonlinear functions, denoted as $\Theta(X)$, where X is the data matrix that contains observed data points of the system variables [3] The library is constructed by carefully selecting relevant nonlinear functions

based on prior knowledge and theoretical considerations. These functions can include polynomials, trigonometric functions, exponentials, logarithmic functions, and other suitable nonlinear expressions [3].

$$\Theta(X) = [1, X, X^2, X^3, \ldots, sin(X), cos(X), \ldots] \tag{4}$$

Higher-order polynomials are denoted as X^2, X^3, and so on. Each column in the $\Theta(X)$ matrix represents a candidate function for the right-hand side of the dynamical equation [3].

Assuming that only a few of these nonlinearities are active in each row of f, a sparse regression problem is formulated to determine the sparse vectors of coefficients,

$$\Xi = \begin{bmatrix} | & | & & | \\ \xi_1, \xi_2, \ldots, \xi_n \\ | & | & & | \end{bmatrix} \tag{5}$$

which indicate the active nonlinearities. Mathematically, this can be expressed as:

$$\dot{X} = \Theta(X)\Xi. \tag{6}$$

Each column, ξ_k, of the Ξ matrix corresponds to a sparse vector of coefficients that determines the active terms in the right-hand side of one of the row equations, $x_k = f(x)$ [3], [12].

Given the data matrix X and the library of candidate functions $\Theta(X)$, SINDy formulates the sparse regression problem as follows:

$$minimize \left\| \dot{X} - \Theta(X)\Xi \right\|_2 + \lambda \left\| \Xi \right\|_1 \tag{7}$$

where Ξ is the sparse vector of coefficients representing the importance or relevance of each term in the library, $\left\|.\right\|_2$ denotes the L2 norm, $\left\|.\right\|_1$ represents the L1 norm, and λ is a regularization parameter that controls the trade-off between data fidelity and sparsity. The first term ensures that the model predictions, obtained by multiplying $\Theta(X)$ with Ξ, are close to the observed data, while the second term encourages a sparse solution by promoting a minimal number of nonzero coefficients [3].

Sparse regression is a crucial step in the process, where a library of candidate functions is subjected to analysis to identify the most concise model that accurately captures the underlying nonlinear dynamics. The sparsity principle is central to this approach, as it seeks to select a subset of functions from the candidate library that is most relevant to the system's dynamics. By incorporating regularization techniques, such as L1 regularization (or the Lasso), the model achieves sparsity by encouraging the coefficients of irrelevant terms to be zero, thereby emphasizing the significant functions while minimizing the overall number of terms [18]. This strategy simplifies the representation of system dynamics, leading to improved interpretability and a more concise model.

Once the Ξ matrix is determined, a model for each row equation can be constructed using the library of candidate functions and the corresponding sparse coefficients. Specifically, the k_{th} row equation, $x_k = f(x)$, can be expressed as [3]:

$$x_k = \Theta(x)\xi_k, \tag{8}$$

where $\Theta(x)$ is a vector of symbolic functions of the elements of x. It is important to note that $\Theta(x)$ differs from $\Theta(X)$ in that it represents symbolic functions of the state variables, unlike $\Theta(X)$, which represents a data matrix. Consequently, the overall representation of the system dynamics can be written as follows:

$$\dot{x} = f(x) = \Xi^T \Theta(x) \tag{9}$$

Each column requires a separate optimization procedure to determine the sparse vector of coefficients, ξ_k, for the corresponding row equation. It is also possible to normalize the columns of $\Theta(X)$, particularly when the entries of X are small [3].

The SINDy has been extensively studied and validated in various scientific domains, showcasing its effectiveness in uncovering governing equations from data. However, the method has limitations. SINDy is sensitive to noise and requires careful model selection to balance sparsity and

accuracy [15]. Its application to densely coupled dynamics poses challenges as disentangling individual contributions becomes difficult [13]. Furthermore, while SINDy captures dynamics from data, it does not explicitly incorporate physical constraints, necessitating additional techniques or prior knowledge incorporation to ensure compliance with fundamental principles [3], [12]. Awareness of these limitations is essential for effective and informed utilization of the SINDy method.

A fundamental and critical undertaking within power system analysis involves the adaptation and contextualization of the general SINDy methodology tailored to the intricacies of electrical power systems. This essential initiative involves the development of a power-specific algorithm that aligns the SINDy principles with the dynamics inherent to power systems. Such an adaptation holds paramount significance as it effectively bridges the gap between the versatile SINDy framework and the unique complexities of power system dynamics.

Algorithm 1:

Input:

Time history of the system's state, denoted as x(t), where x(t)∈Rn.

Regularization parameter λ.

Library of candidate functions, $\Theta(X)$, where X is the data matrix that contains observed data points of the system variables.

Step 1: Data Collection

Initialize empty matrices X and $\dot{X}$, where $\dot{X}$ represents the time derivatives of X.

For each time instance t in the set of time instances:

Collect data at time t and store it in x(t).

Compute the derivative of x(t) at time t, denoted as $\dot{x}$ (t).

Append x(t) to the X matrix.

Append x˙(t) to the $\dot{X}$ matrix.

Step 2: Construct Library of Candidate Functions

Initialize an empty list for the library of candidate functions.

For each candidate function in the set of candidate functions:

Compute the values of the candidate function using data matrix X.

Append the function values to the library.

Step 3: Sparse Regression

Initialize an empty list Ξ to store the sparse coefficient matrices for each variable.

For each system variable k:

Perform sparse regression using data matrices X, $\dot{X}$, the library $\Theta(X)$, and regularization parameter λ to obtain Ξ_k.

Sparse Regression Formulation:

The sparse regression problem for variable k can be expressed as:

$$minimize \left\| \dot{X} - \Theta(X)\Xi \right\|_2 + \lambda \left\| \Xi \right\|_1$$

where Ξ_k represents the sparse coefficient matrix for system variable k.

$\left\| . \right\|_2$ denotes the L2 norm.

$\left\| . \right\|_1$ represents the L1 norm.

λ is a regularization parameter that controls the trade-off between data fidelity and sparsity.

Step 4: Model Construction

Initialize an empty list for the models representing the system dynamics.

For each system variable k:

Construct the model for variable k using the library of candidate functions $\Theta(x)$ and the corresponding sparse coefficient Ξ_k.

Model Construction:

The model for system variable k can be expressed as:

$$x_k = \Theta(x)\xi_k$$

where $\Theta(x)$ is a vector of symbolic functions of the elements of x.

ξ_k is the sparse vector of coefficients that determines the active terms in the system variable $x_k = f(x)$.

Output:

The list of models represents the system dynamics for each system variable, providing concise and accurate descriptions of the underlying nonlinear dynamics.

The proposed three-section data analysis structure represents a notable evolution in data analysis, particularly within the SINDy framework. This innovative structure introduces a novel third section that extends the conventional SINDy methodology, providing new avenues for enhanced data analysis. The initial two sections focus on identifying linearizable and nonlinear dynamics within the system. The addition of the third section substantially enhances the overall data analysis process by categorizing and managing negligible data, often regarded as noise, which significantly improves the precision and accuracy of system modeling. This comprehensive approach enables a more profound understanding of intricate nonlinear behaviors, benefiting applications across diverse domains, from power systems to the natural sciences. The inclusion of the third section underscores the adaptability and versatility of the SINDy methodology, allowing for a more nuanced examination of complex system dynamics, a critical component of contemporary data-driven research.

The first section of the proposed data analysis framework plays a fundamental role in identifying and characterizing first-order impacts within the system. Its unique focus lies in evaluating the nonlinearities across nodes, with a particular emphasis on those that are trivial or readily linearizable. This systematic approach dissects the system dynamics to isolate elements that exhibit straightforward and manageable nonlinearities amenable to linear approximations. This categorization enhances the precision and tractability of the data analysis process, providing valuable insights into complex system behaviors encompassing both linear and nonlinear components, particularly in applications spanning diverse domains, including power systems.

The second section within the outlined data analysis structure takes a central role in the comprehensive examination of system dynamics. It is dedicated to discerning and categorizing true nonlinearity, which differs significantly from the more straightforward and readily linearizable elements identified in the first section. True nonlinearity represents intricate and non-trivial characteristics that defy simple linearization, delving deep into complex system behaviors. Focusing on these inherently nonlinear dynamics offers profound insights into the intricate interdependencies and feedback loops characterizing real-world systems, transcending linear approximation constraints. This in-depth analysis is pivotal for understanding nonlinearity's nuances across various domains, providing a fundamental foundation for a richer comprehension of system dynamics, particularly in the context of power systems and beyond.

The third section within the data analysis structure serves a crucial role in isolating and addressing components of system dynamics categorized as negligible. These elements include tolerable errors, inherent noise, and other factors exerting minimal influence on the overall system behavior. While individually minor, their cumulative impact can introduce variations and perturbations in the system's dynamics. However, by considering these factors within a dedicated section, they can be managed and refined effectively, enhancing the overall modeling accuracy of the system. This meticulous categorization offers a valuable framework for researchers and system operators, allowing them to discern essential dynamics from negligible ones, ensuring a more accurate representation of system behavior. This process is fundamental for optimizing system models and is highly relevant to applications in various domains, with particular significance in power system analysis.

3. Demonstration Study

In this comprehensive study on an IEEE 15-bus power grid, we employed the SINDy algorithm to analyze voltage waveforms and identify system dynamics under various complex operational

scenarios. The choice of the experimental configuration was deliberate, aligning it with similar studies conducted by other researchers in 2022 [8]. The system architecture consists of 15 buses interconnected through branches representing power transmission lines, each with unique parameters and attributes governing power flow dynamics, as shown in Figure 2. Our discussion also encompasses the test scenarios, including abrupt changes and gradual load variations in the context of conventional SG and IBR at 50% and 100% integration. Our analysis serves as the foundation for introducing the Volterra-based Nonlinearity Index as a novel tool for assessing the level of nonlinearity in dynamic systems, offering significant insights into system dynamics.
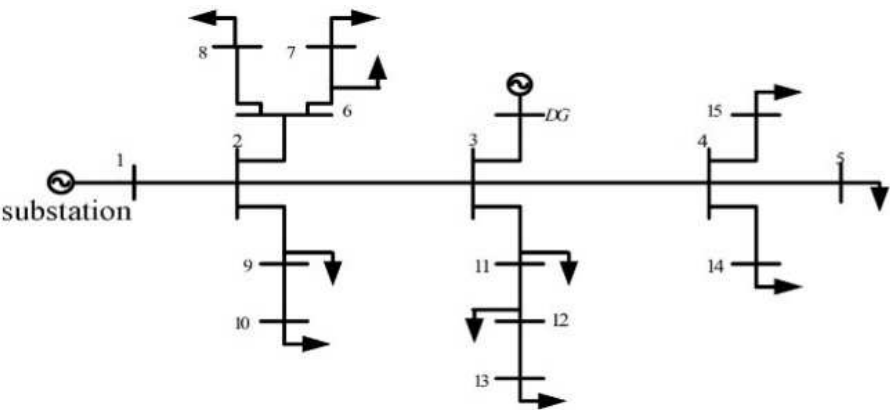

Figure 2. Single-line diagram of the implemented IEEE 15 bus network.

A. Investigation setup:

The system architecture consists of 15 buses, representing distinct nodes within the power system network, and these buses are interconnected through branches that represent the power transmission lines. Each bus in the IEEE 15-bus system has a unique set of parameters and attributes and is connected to neighboring buss via branches characterized by specific impedance, which govern the dynamics of power flow among the interconnected buses. Table 1 provides a comprehensive overview of the network configuration and branch parameters of the IEEE 15-bus system.

Table 1. – The implemented IEEE 15 bus network configuration.

Line index	From bus	To bus	$r+x$i (Ω)	Node index	$P_{load}+jQ_{load}$ (kW+kVAR)
1	1	2	1.53+1.778i	2	100+j60
2	2	3	1.037+1.071i	3	90+j40
3	3	4	1.224+1.428i	4	120+j80
4	4	5	1.262+1.499i	5	60+j30
5	5	9	1.176+1.335i	6	60+j20
6	6	10	1.1+0.6188i	7	200+j100
7	7	6	1.174+0.2351i	8	200+j100
8	8	7	1.174+0.74i	9	60+j20
9	9	8	1.174+ 0.74i	10	60+j20
10	10	11	1.15+ 0.065i	11	45+j30
11	11	12	1.274+1.522i	12	60+j35
12	12	13	1.274+1.522i	13	60+j35
13	13	14	1.075+1.522i	14	120+j80
14	14	15	1.075+ 1.522i	15	60+j10

This investigation explores a comprehensive set of power system conditions, encompassing both abrupt changes (faults) and gradual changes (load variations), in the context of conventional SG and IBRs. The study encompasses three distinct scenarios, representing both single- and multi-dynamic systems. The first scenario examines a system solely supplied by a synchronous generator at Bus 1, with all other generators disconnected from the network. The second scenario incorporates the integration of an IBR at Bus 3, sharing the load demand equally with the synchronous generator at Bus 1. In the third scenario, the network is subjected to a 100% penetration of IBRs, where the demand is supplied by two IBRs located at Bus 1 and Bus 3. Each scenario spans 10 seconds, with the synchronous generators and IBRs initiated at $t = 0s$. At $t = 3.3s$, a three-phase-to-ground fault occurs at Bus 10, cleared after four cycles of the fundamental frequency (60 Hz). Furthermore, at $t = 7s$, a significant load is connected to Bus 14, only to be disconnected at $t = 8s$.

The SINDy algorithm, described in Algorithm 1, analyzed the acquired voltage waveforms, demonstrating promising system identification and modeling capabilities. Employing a refined computational approach, we meticulously consider essential parameters and algorithms to facilitate a comprehensive analysis. Careful consideration is given to the sampling rate (20,000 samples per second) and fundamental frequency (60 Hz) to ensure a high-fidelity representation of electrical phenomena. The simulation duration (10 s) and total sample count (200,000) are determined to capture temporal dynamics faithfully. By leveraging the Hilbert transform, converting voltage waveforms into complex numbers, and subsequent computation of instantaneous phase angles, we gain profound insights into the intricate behavior of the system.

Furthermore, the simulation methodology incorporates the SINDy algorithm, wherein a polynomial library is constructed with a precise second-order polynomial and regularization parameter (0.8). The ensuing coefficients derived from this process are then employed to solve the system's ordinary differential equation, thus elucidating the underlying dynamics. Rigorous evaluation is conducted to assess the accuracy of the predicted slow dynamics and quantify the disparity between the identified fast dynamics and actual data. This meticulous simulation framework, accompanied by its key parameters and algorithms, engenders a robust foundation for comprehensive investigations into the intricacies of power systems.

Applying polynomial function libraries up to the third order in the context of inverter-based resources signifies a notable departure from traditional modeling approaches. In power grid modeling, mainly when dealing with inverter-based resources, these higher-order polynomial functions allow for a more intricate representation of the dynamic behavior within the system. Unlike first-order models that may oversimplify the interactions between various components, including polynomial functions up to the third order enables the capture of nonlinearities and interactions characteristic of inverter-based resources. These functions provide a flexible framework to model the complex interplay between inverter controllers, grid conditions, and the response of renewable energy sources to changing environmental factors.

The application of these polynomial function libraries has theoretical and practical implications. Theoretically, it acknowledges the importance of capturing higher-order dynamics and interactions. It aligns with the principles of complex systems theory, emphasizing the significance of nonlinear dynamics and the emergence of complex behaviors in systems like power grids with significant inverter-based resource penetration. From a practical standpoint, this approach facilitates more accurate modeling, enabling grid operators and planners to understand better and predict the behavior of inverter-based resources. The practical advantages are particularly evident when renewable energy integration is critical. By accommodating higher-order dynamics, these models enhance the ability to simulate, analyze, and optimize the grid's performance, ultimately contributing to a more resilient and efficient power system.

Our investigation introduces a three-section data analysis structure, offering an enhanced approach to analyzing system dynamics. The initial two sections focus on identifying linear and nonlinear dynamics within the system, categorizing elements that are linearizable and those that are inherently nonlinear. The addition of the third section allows for managing negligible data

components, ensuring improved modeling accuracy. This comprehensive approach provides a more profound understanding of intricate nonlinear behaviors across various domains.

B. SG-driven power grid dynamic identification:

In this study we analyze the generic SINDy algorithm on a power grid, IEEE 15-bus, subjected to diverse dynamics and conditions, reaffirming the method's strengths and limitations in identifying system models, addressing the intricate coupling dynamics challenges, and discussing additional limitations that require consideration. The study's findings demonstrate that SINDy, with its utilization of voltage waveforms, successfully captures essential patterns and relationships within the electrical behavior of the power grid. The results highlight the potential of SINDy as a powerful tool for system identification and modeling in power systems. The results related to the SG model exhibit consistencies to foundation theories and findings from other research such as Stankovic's work in 2020 [8]. Our analysis highlights the dominance of first-order terms in the extensive 15-dimensional system. Second-order terms play minor roles and third-order terms are close to zero, affirming the precision of our SINDy-based model with an impressively low error rate, affirming its accuracy in capturing both short-term and long-term dynamics.

In the analysis of a power grid relying exclusively on synchronous generator resources, our investigation aimed to identify the model characterizing the system's dynamics. Figure 3 demonstrates that first-order terms dominate within the all 15-dimensional system. In this case, the SG resources are modeled with the basic model. Second-order terms and third-order terms coefficients are negligible, approximated to zero by the MATLAB calculation.
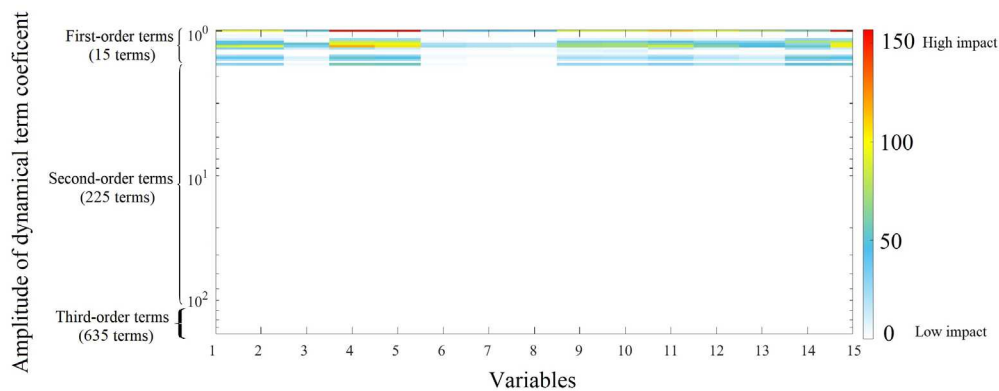

Figure 3. Colormap of the dynamical terms identified by SINDy for the IEEE 15-bus system supplied by basic SG sources.

The approximation of the system dynamics highlights the precision of the model estimated. The error between actual data and the identified model's approximation is impressively less than 0.001 percent, affirming the model's accuracy in capturing both short-term and long-term dynamics.

To reduce the potential bias introduced by our choice of SG model, we replicated the investigation using a 7th-order SG model – the most intricate SG model accessible within the MATLAB framework. This strategic adaptation allowed us to explore the impact of higher-order terms, primarily second-order terms, on the overall dynamics of the system. The result, presented in Figure 4, shows that the dynamic is still dominated by first-order terms where second-order term coefficients are minimal, comprising less than 1 percent of the smallest first-order term, and third-order term coefficients are even more negligible, approximated to zero by MATLAB.

It is noteworthy that the appearance of second-order terms did not yield any significant effects, which aligns with the expectations set by current theoretical frameworks. The conventional understanding of power systems suggests that higher-order terms, particularly second-order, tend to play a relatively minor role in system behavior, especially when compared to the prominence of first-order terms. This observation underscores the consistency of our findings with established theoretical principles, reaffirming the accuracy and reliability of our analysis.
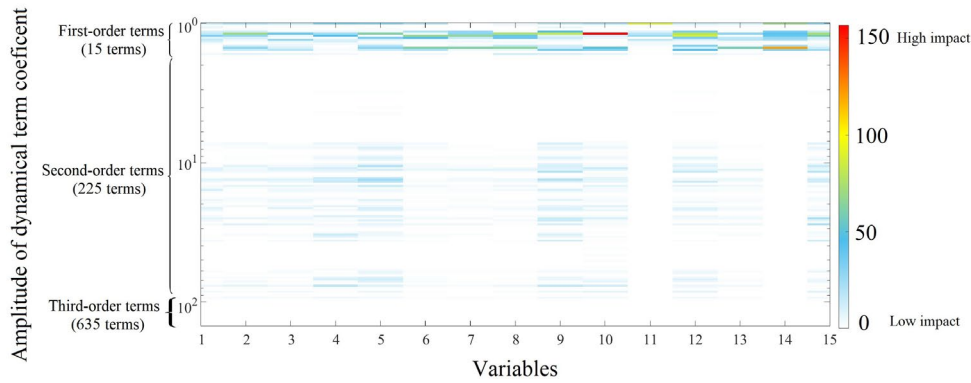
Figure 4. Colormap of the dynamical terms identified by SINDy for the IEEE 15-bus system supplied by 7th-order SG sources.

This study has analyzed the SINDy algorithm in the context of a complex power grid, specifically the IEEE 15-bus system. The obtained findings underscore the potential of SINDy as a robust and effective tool for system identification and modeling within power systems, particularly when utilizing voltage waveforms to capture essential patterns and relationships within the electrical behavior of the grid. Furthermore, our analysis of the SG model reveals consistencies with foundational theories and findings from previous research while exploring the impact of implemented models by adapting a 7th-order SG model to show results further affirm the prevailing influence of first-order terms and the negligible effect of second-order terms.

C. Dynamic identification in IBR integrated power grid

By conducting the investigations on the power system integrated with IBRs this section navigates through these distinct nonlinear behaviors, grounded in the principles of complex systems theory and nonlinear dynamics, to offer valuable insights into the dynamic response of power grids with substantial IBR integration. The results embark on an exploration of the dynamic behavior of IBRs within a power grid, shedding light on the distinctive difference in the nature of nonlinear dynamics caused by IBRs and SGs. The nonlinearity in SGs is characterized as "module-level dynamic nonlinearity," rooted in well-documented electromagnetic principles and iron core saturation effects, primarily influenced by the individual components of SGs. In contrast, the nonlinearity encountered in IBRs reveals a multifaceted character, encompassing both module and system-level nonlinearity. The latter, system-level nonlinearity, is a product of intricate interactions between diverse components, control algorithms, and the inherent variability of input sources, indicative of complex system dynamics.

The investigation will analyze two scenarios. The first scenario will conduct the test over the same IEEE 15-bus system that is supplied by both the basic SG model and IBR where each supplies 50 % of the load demand. The second scenario will investigate the same power system under full penetration of IBRs, i.e., 100 % of the load demands are supplied with IBRs.

Through the first scenario with 50 % integration of inverter-based resources into the power grid, equalizing their role with SGs in supplying load demands, the SINDy algorithm was used to identify the underlying dynamic with the measured data. As Figure 5 illustrates, it was found that the second-order terms become more effective in the dynamic model. The higher impact of the second-order terms will show that the data-based model, i.e., the underlying model within measurements, includes a more non-linearizable nature that shows itself in higher coefficient values for second-order terms. However, the third-order terms, representing the negligible data (noise) are still in the same condition.

The results unveil a noteworthy transformation in the power grid dynamics during our exploration of IBR integration. What becomes evident is not only the heightened influence of second-order terms (characterizing nonlinear dynamics) on the system's overall behavior within all the individual buses but also the activation of more terms in the second-order region, thus highlighting

the unmistakable imprint of system-level nonlinearity on the outcomes. This intriguing shift underscores the intricate interplay of components in the network and the inherent variability in each source, collectively contributing to the observed system-level nonlinearity. In comparison to previous scenarios, this outcome signifies a fundamental difference, illuminating how power grids evolve when IBRs are introduced, revealing a dynamic that transcends more module dynamics and delves into the realm of complex, system-wide nonlinear interactions.
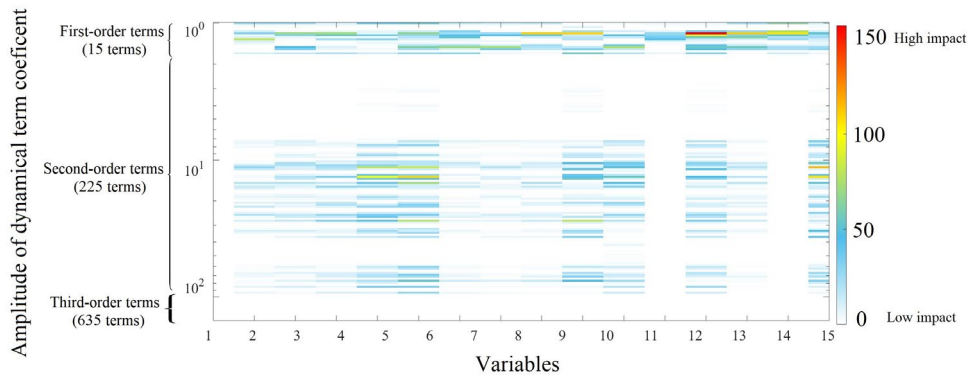

Figure 5. Colormap of the dynamical terms identified by SINDy for the IEEE 15-bus system supplied by basic model SG sources and IBRs.

In the next scenario, where load demands were exclusively supplied by IBRs, our model identification analysis, as depicted in Figure 6, remarkably underscored the dominance of second-order terms. This shift in the composition of dominant terms is a pivotal result that merits in-depth discussion. The prominence of second-order terms in this context carries profound implications for understanding power grid dynamics.
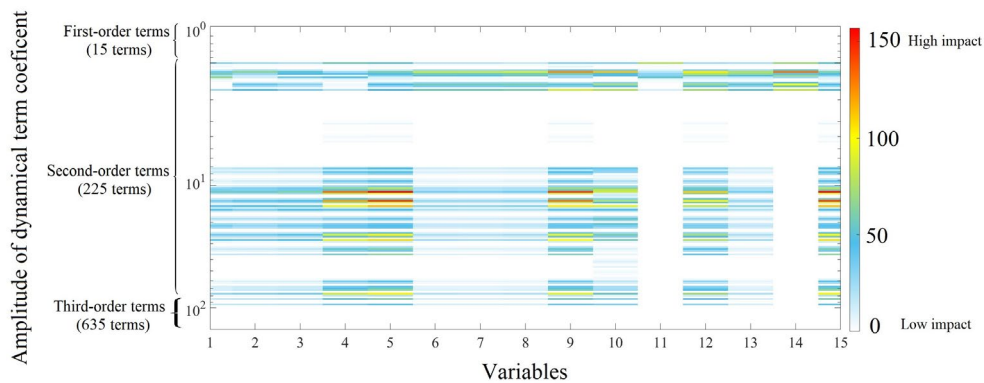

Figure 6. Colormap of the dynamical terms identified by SINDy for the IEEE 15-bus system supplied by IBRs.

This outcome indicates the profound impact of IBR-related nonlinearity on the power grid, highlighting the necessity for more nuanced modeling to represent these complex interactions accurately. The dominance of second-order terms indicates that these nonlinear behaviors have a substantial impact on the system's dynamics. In the context of IBRs, it becomes evident that second-order terms play a key role in capturing and representing the system's response to these nonlinear effects. This underscores the need to consider and model the nonlinearity introduced by IBRs explicitly, as first-order models may need to be revised to represent these intricate interactions.

Moreover, from a theoretical perspective, this result aligns with established principles of nonlinear system dynamics. In complex systems, it is expected to observe higher-order nonlinearities, especially when the interactions among system components are intricate. These second-order terms

can arise due to a variety of reasons, including feedback mechanisms, nonlinear component characteristics, and complex system interactions.

Furthermore, a noteworthy observation emerged when comparing the level of participation of second-order terms in the 100 % penetration of IBRs to the 50 % and zero penetration scenarios. This comparison, presented in Figure 7, accentuated a significant shift, indicative of the substantial impact of variable interactions in contrast to the direct effects of individual variables on the system's dynamics.
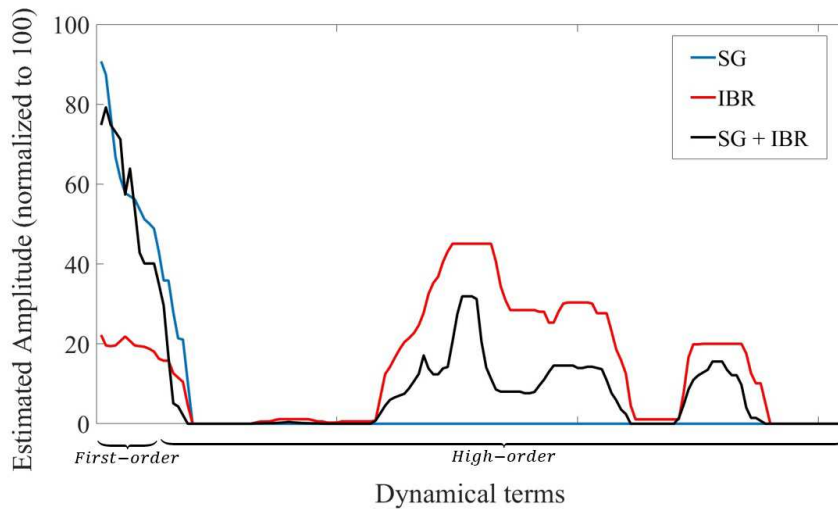

Figure 7. The normalized impact of the activated first-order and high-order terms in the dynamic of the variables for the scenarios with SG, a combination of SG and IBRs, and IBRs.

As observed, the analysis of the 7th-order SG model underscores the prevalence of first-order terms, indicative of linearizable dynamics within the measured data. The limited influence of high-order terms and noise in this model allows for their neglect without substantially affecting the model's accuracy. Furthermore, the nonlinearity in this context is primarily associated with the individual buses connected to SG sources, suggesting a module-level nonlinearity.

In scenarios featuring 50% and 100% integration of IBRs, a notable increase in nonlinearity is evident, manifested by higher coefficients for second-order terms. It is worth noting that the integration of IBRs activates a greater number of terms, and these activated terms are not directly linked to the buses connected to IBR sources, indicating the emergence of system-level nonlinearity resulting from network interactions. The graphical representation vividly portrays the escalating nonlinearity and the increasing influence of second-order terms as IBR penetration rises from 50% to 100%, ultimately leading to the dominance of nonlinear dynamics in the overall system behavior.

D. Volterra-based Nonlinearity Index

In dynamic systems, the interplay between linear and nonlinear behaviors is a common phenomenon, and quantifying this nonlinearity holds paramount importance for comprehending system performance, facilitating effective control, and optimizing signal processing. In this study, the Volterra-based Nonlinearity Index (VNI) is introduced as a novel analytical instrument with the capacity to evaluate nonlinearity in dynamic systems quantitatively. This section not only introduces the fundamental concept of VNI but also explores its mathematical underpinnings. VNI's significance transcends as it enables the quantification of the nonlinear-to-linear relationship within dynamic systems, offering profound insights into the intricate dynamics at play. Moreover, VNI's versatility allows for the recognition of different types of nonlinearities and the quantification of the relative influences of system-level and module-level nonlinearity, further enhancing its utility in system analysis and modeling. Using the case studies conducted in this research, we showcase practical experiments employing VNI, revealing new discoveries regarding nonlinearity in the 7th-order SG

model compared to IBR. These discoveries highlight different patterns of nonlinearity and emphasize the importance of structural analysis in identifying their sources.

Since dynamic systems often exhibit a combination of linear and nonlinear behaviors, characterizing the extent of nonlinearity is pivotal in understanding system performance, control, and signal processing. To elucidate the impact of high-order terms and highlight the effectiveness of the SINDy in capturing system dynamics, we introduce a novel index that assesses the influence of high-order dynamics. The VNI is an analytical tool designed to assess the level of nonlinearity in dynamic systems quantitatively. VNI draws its foundation from the Volterra series [16], a powerful mathematical construction that dissects system responses into linear and higher-order nonlinear components, providing a systematic approach for nonlinear modeling and analysis [16].

VNI is expressed as the ratio of the energy (or magnitude) associated with the nonlinear components to the energy of the linear response within the Volterra series expansion. This formulation encapsulates the inherent nonlinearity of the system and the interplay between linear and nonlinear phenomena. Mathematically, VNI is defined as:

$$VNI = \frac{\sum_{k=1}^{N} \iint |H_k(\tau_1, \tau_2)|^2 d\tau_1 d\tau_2}{|H_0|^2} \tag{10}$$

Here, N represents the selected order of the Volterra series, accommodating the analysis of a range of higher-order nonlinear terms $H_k(\tau_1, \tau_2)$ signifies the Volterra series coefficients on the kth order nonlinear terms, and $|H_0|^2$ represents the squared magnitude of the linear response.

A higher VNI value implies a greater prevalence of nonlinearity in the system. Consequently, VNI serves as a comprehensive gauge of the nonlinear-to-linear relationship within a dynamic system, contributing to a more profound understanding of the system's dynamics and its suitability for specific applications.

It has to be noted that the application of machine learning approaches to the VNI introduces an exciting dimension in the realm of system analysis and modeling. VNI, when coupled with machine learning techniques, can unlock the potential to discern and differentiate various types of nonlinearities inherent within complex dynamic systems. Machine learning algorithms recognize patterns, relationships, and hidden structures within data, and when applied to VNI data, they can extract nuanced distinctions in the system's behavior. These distinctions manifest as different types of nonlinearities that are challenging to identify through conventional methods. This capability has significant implications for characterizing the complex behavior of systems with mixed linear and nonlinear components relying on measurement-based and real-time methods, as it can provide insights into how different nonlinear phenomena manifest and interact in the overall system response.

The VNI framework could find utility in diverse applications across science and engineering disciplines. In control systems, it aids in assessing the stability and robustness of nonlinear control strategies, informing the choice of appropriate controllers. In communication systems, it provides insights into signal quality, especially in scenarios where nonlinear effects can degrade signal integrity. In physical and biological systems, VNI enables researchers to quantify and understand the nonlinear interactions underlying complex behaviors.

Applying the VIN and VIN profile to the case studies in our investigations, the VIN values for three scenarios in the system were supplied with 7th-order SGs, a combination of basic model SG and IBRs, and the IBRs were calculated equal to o.78, 0.54, 3.54, respectively.

To have a visual indicator of the extent of nonlinearity within the dynamic system the VIN profile, as a novel concept used to assess the linearity and nonlinearity within dynamic systems, particularly in system-level interactions is introduced. It is based on the proposed VIN, which quantitatively measures the level of nonlinearity within a dynamic system. The VIN profile associates a profile with the identified dynamics of the system, explicitly relating the slope of the profile, r, to the calculated VIN value through the following relation:

$$r = VIN - 1 \tag{11}$$

This relationship provides a means to classify the prevalence of either linearity or nonlinearity within the system. The calculation of the VIN profile for each case study serves to distinctly illustrate and visualize the distinctions in both the magnitude and character of nonlinearity between them. When the VIN profile exhibits a negative slope, it suggests that linearity dominates the system dynamics. In this scenario, the higher-order terms have less influence on the overall dynamics, indicating that the system's behavior is primarily linear or that nonlinearity is confined to a module level. This means that linear relationships can predominantly explain the system's response and the impact of higher-order terms is limited.

Conversely, when the VIN profile shows a positive slope, it signifies the domination of nonlinearity within the system and suggests that the higher-order terms have a more significant impact on the overall dynamics. In such cases, nonlinearity is not confined to module-level interactions but extends to system-level interactions. By analyzing this profile, researchers and engineers can understand whether linearity or nonlinearity predominates in a given system and whether the nonlinearity is confined to the module level or extends to system-level interactions. The estimated VIN profiles of the scenarios with 7th-order SGs, a combination of basic model SG and IBRs, and the IBRs are presented in Figure 8.
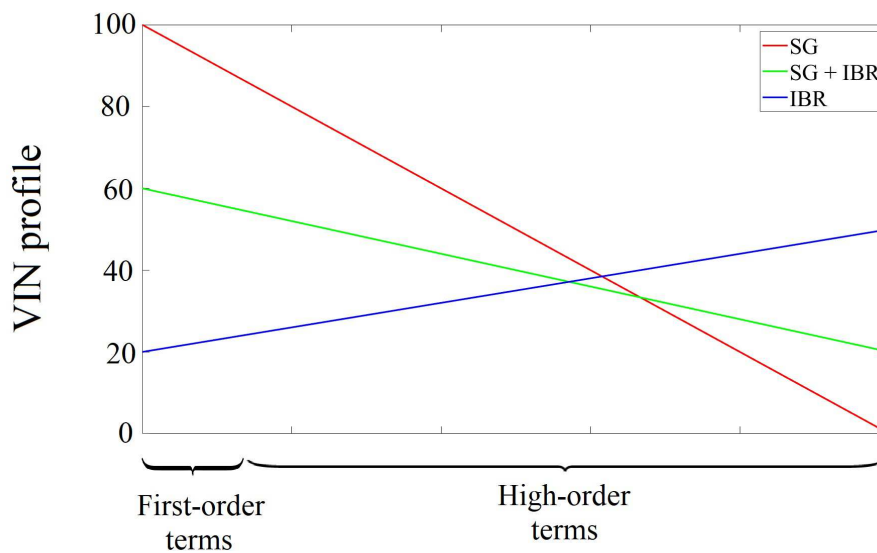

Figure 8. The VIN profile of the dynamic for the scenarios with SG, a combination of SG and IBRs, and IBRs.

4. Conclusions

This study has been dedicated to the development and practical application of SINDy methods tailored to the complex domain of power systems, as exemplified by the IEEE 15-bus network. The primary objective has been to bridge the gap between theoretical understanding and real-world implementation, with a particular focus on the analysis of system dynamics. We harnessed the state-of-the-art SINDy algorithm and the proposed novel Volterra-based Nonlinearity Index to navigate the intricate landscape of power systems.

An essential contribution of this study is the introduction of a clear and concise distinction among different classes of dynamical terms within power systems. We have categorized these dynamics into three distinct groups, offering a systematic dissection of the complexity inherent in power systems. The first-order terms represent system elements characterized by linear behavior, readily amenable to traditional modeling approaches. Second-order terms point to significant nonlinear dynamics, exerting a notable influence on the system's response. Lastly, the third-order terms correspond to noise or negligible data components, underscoring the importance of noise management in our analysis.

Our findings also underscore a fundamental shift in system dynamics as the power source transitions to inverter-based resources, highlighting the presence of system-level nonlinearity in contrast to the module-level nonlinearity observed in conventional synchronous generator resources. This distinction holds significant implications for the modeling and analysis of modern power systems, emphasizing the need for nuanced approaches to understanding their complex behavior.

Furthermore, introducing the Volterra-based Nonlinearity Index adds a new dimension to our understanding of power system performance. VNI enables quantifying the nonlinear-to-linear relationship within dynamic systems, offering profound insights into the intricacies of power system behavior. Its versatility could allow for recognizing various nonlinearities, facilitating the differentiation between module-level and system-level nonlinearity. The VIN profile, derived from VNI, offers a visual representation of the extent and character of nonlinearity within a system, providing additional layers of insight into the behavior of intricate power systems.

The complexities of modern power systems demand ongoing research and development to refine our tools and approaches. Future investigations could further enhance our ability to detect and identify sources of oscillations in real-time, a crucial step towards ensuring the stability and reliability of our power grids.

Author Contributions: The contributions of the authors for this research article are as follows: Conceptualization, Reza Saeed Kandezy and John Jiang; methodology, Reza Saeed Kandezy; software, Reza Saeed Kandezy; validation, Reza Saeed Kandezy, John Jiang, and Di Wu; formal analysis, Reza Saeed Kandezy; investigation, Reza Saeed Kandezy; resources, John Jiang; data curation, Reza Saeed Kandezy; writing—original draft preparation, Reza Saeed Kandezy; writing—review and editing, Di Wu; visualization, Reza Saeed Kandezy; supervision, John Jiang. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: There are no external data for this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Corning, Peter A. "Synergy and self-organization in the evolution of complex systems." Systems Research 12.2 (1995): 89-121.
2. Vu, Thanh Long, and Konstantin Turitsyn. "A framework for robust assessment of power grid stability and resiliency." IEEE Transactions on Automatic Control 62.3 (2016): 1165-1177.
3. Brunton, Steven L., Joshua L. Proctor, and J. Nathan Kutz. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems." Proceedings of the national academy of sciences 113, no. 15 (2016): 3932-3937
4. Mishra, Akhilesh Kumar, et al. "Maiden Application of Integral-Tilt Integral Derivative with Filter (I-TDN) Control Structure for Load Frequency Control." IFAC-PapersOnLine 55.34 (2022): 72-77.
5. Ekomwenrenren, Etinosa, et al. "Hierarchical coordinated fast frequency control using inverter-based resources." IEEE Transactions on Power Systems 36.6 (2021): 4992-5005.
6. Orihara, Dai, et al. "Contribution of voltage support function to virtual inertia control performance of inverter-based resource in frequency stability." Energies 14.14 (2021): 4220.
7. Kenyon, Rick Wallace, et al. "Stability and control of power systems with high penetrations of inverter-based resources: An accessible review of current knowledge and open questions." Solar Energy 210 (2020): 149-168.
8. Stanković, Alex M., et al. "Data-driven symbolic regression for identification of nonlinear dynamics in power systems." 2020 IEEE Power & Energy Society General Meeting (PESGM). IEEE, 2020.
9. Impram, Semich, Secil Varbak Nese, and Bülent Oral. "Challenges of renewable energy penetration on power system flexibility: A survey." Energy Strategy Reviews 31 (2020): 100539.
10. Karangelos, Efthymios, Patrick Panciatici, and Louis Wehenkel. "Whither probabilistic security management for real-time operation of power systems?." 2013 IREP Symposium Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid. IEEE, 2013.
11. Teti, Roberto, et al. "Advanced monitoring of machining operations." CIRP annals 59.2 (2010): 717-739.

12. Fasel, Urban, Eurika Kaiser, J. Nathan Kutz, Bingni W. Brunton, and Steven L. Brunton. "SINDy with control: A tutorial." In 2021 60th IEEE Conference on Decision and Control (CDC), pp. 16-21. IEEE, 2021.

13. Kaheman, Kadierdan, J. Nathan Kutz, and Steven L. Brunton. "SINDy-PI: a robust algorithm for parallel implicit sparse identification of nonlinear dynamics." Proceedings of the Royal Society A 476, no. 2242 (2020): 20200279.

14. Brunton, Steven L., Bernd R. Noack, and Petros Koumoutsakos. "Machine learning for fluid mechanics." Annual review of fluid mechanics 52 (2020): 477-508.

15. Fasel, Urban, J. Nathan Kutz, Bingni W. Brunton, and Steven L. Brunton. "Ensemble-SINDy: Robust sparse model discovery in the low-data, high-noise limit, with active learning and control." Proceedings of the Royal Society A 478, no. 2260 (2022): 20210904.

16. Boyd, Stephen, Leon O. Chua, and Charles A. Desoer. "Analytical foundations of Volterra series." IMA Journal of Mathematical Control and Information 1.3 (1984): 243-282.