

Data-driven Symbolic Regression for Identification of Nonlinear Dynamics in Power Systems

Alex M. Stanković¹, Aleksandar A. Sarić²

Dept. of Electrical and Computer Eng.

Tufts University

Medford, MA, USA

¹ astankov@ece.tufts.edu

² aleksandar.saric.95@gmail.com

Andrija T. Sarić

Dept. of Power, Electr. & Com. Eng.

Faculty of Technical Sciences

Novi Sad, Serbia

asaric@uns.ac.rs

Mark K. Transtrum

Dept. of Physics & Astronomy

Brigham Young University

Provo, UT, USA

mktranstrum@byu.edu

Abstract—The paper describes a data-driven system identification method tailored to power systems and demonstrated on models of synchronous generators (SGs). In this work, we extend the recent sparse identification of nonlinear dynamics (SINDy) modeling procedure to include the effects of exogenous signals and nonlinear trigonometric terms in the library of elements. We show that the resulting framework requires fairly little in terms of data, and is computationally efficient and robust to noise, making it a viable candidate for online identification in response to rapid system changes. The proposed method also shows improved performance over linear data-driven modeling. While the proposed procedure is illustrated on a SG example in a multi-machine benchmark, it is directly applicable to the identification of other system components (e.g., dynamic loads) in large power systems.

Index Terms— Power system, Dynamic model, System identification, Nonlinear dynamics, Sparse model.

I. INTRODUCTION

Dynamic modeling of power systems has received increased attention lately due to changes in components (such as inverter-connected sources and loads) and in data availability (such as Phasor Measurement Units—PMUs). While there exists a plethora of dynamical models for components and interconnections, their practical identifiability from data is a long-standing problem. Our own efforts in this direction have focused on model reduction and global identifiability of models in transient stability studies [1], and have provided some new connections with analytical tools such as the singular perturbation theory. In this paper, we open another line of inquiry—assuming we have a practically identifiable model, and an optimized set of parameters for it, what can be said about the inevitable discrepancies between the model predictions and the measurements? In other words, is there a structure in the error signals that can be utilized to possibly amend the model, while maintaining identifiability? This is tantamount to appending the original model with new parts that are driven by the actual data. The tool that we explore for this purpose belongs to the family of symbolic regression methods.

Symbolic regression parallels the idea of standard regression in searching a space for the best fit to data but operates in a space of mathematical expressions to derive a

model that best matches a given set of data in terms of accuracy and simplicity (parsimony). The candidate expressions are formed by combining building blocks that comprise functions, constants, and other mathematical expressions. Challenges come from the large space of candidate models and the (likely) non-uniqueness of solutions.

We focus on transient stability, which focuses on system-level behavior, and de-emphasizes microscopic (component-level) accuracy. This means that we do not have to explore all possible microscopic dynamics, enabling us to focus on a much smaller, nested hierarchy of models. We use physical intuition of the phenomena to restrict the set of building blocks to include in the algorithm. This decision could potentially limit performance since there is no theoretical guarantee of completeness [2]. However, we show in our application section that our observed performance is encouraging.

Our point of departure is the equation-based modeling framework [3], where authors developed an algorithm for the sparse identification of nonlinear dynamics (SINDy) from data relying on the fact that most dynamical systems of interest have relatively few nonlinear terms, each belonging to a family (i.e., polynomial nonlinearities, etc.). This method uses sparsity promoting techniques to find models that automatically balance sparsity in the number of terms with model accuracy. This algorithm is extended to include controls in [4].

In this paper, we use a modified SINDy algorithm to identify system-wide aspects of dynamic components (in our case a SG example). The model identification is performed from locally recorded measurements only. Having in mind the structure of the standard power system equations, the library of dynamical elements in SINDy is extended with state multiplication and trigonometric functions of states and external measurements.

The outline of the paper is as follows: in *Section II* we give the problem formulation; *Section III* describes the SINDy algorithm with extended the library of elements, while *Section IV* describes the SINDy's extension with external measurements; *Section V* shows the steps and convergence details of SINDy algorithm; the proposed algorithm is applied to a SG example in *Section VI*; *Section VII* presents conclusions. The *Appendix* contains basic SG's parameters adapted for SINDy-based model identification.

This work has been supported by NSF under grant ECCS-1710944, by CURENT Engineering Research Center of the National Science Foundation and the Department of Energy under NSF Award Number EEC-1041877, by ONR under grant N00014-16-1-3028, and in part by the Ministry of Education and Science of the Republic of Serbia, under project III-42004.

II. PROBLEM FORMULATION

The Differential Algebraic Equation (DAE) based form of power system dynamical models used in transient stability is [1]

$$\dot{\mathbf{x}} = d\mathbf{x}/dt = \mathbf{f}(\mathbf{x}, \mathbf{z}, t) \quad (1)$$

$$\mathbf{0} = \mathbf{g}(\mathbf{x}, \mathbf{z}, t) \quad (2)$$

where $\mathbf{x}$ is the vector of (differential) state variables, $\mathbf{z}$ are the algebraic variables, and t is the (scalar) time variable. The system measurement vector is assumed to be of the form

$$\mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{z}, t). \quad (3)$$

To apply the SINDy framework, we formally transform into ODE form by solving for $\mathbf{z}$ in (2) and substituting to (1) and (3), we have

$$\dot{\mathbf{x}} = \mathbf{f}'(\mathbf{x}, t) \quad (1')$$

$$\mathbf{y} = \mathbf{h}'(\mathbf{x}, t). \quad (3')$$

For power system-based equations, for derivation of (1') and (3'), see, for example [9].

III. SPARSE IDENTIFICATION OF NONLINEAR DYNAMICS (SINDy)

The SINDy algorithm seeks to identify a fully nonlinear dynamical system, $\dot{\mathbf{x}}$ (1') from measurement data set, $\mathbf{y}$ (3'). It relies on the fact that many dynamical systems have relatively few terms in the right-hand side of the governing equations (1') and (3'). Based on standard equations in [9], an appropriate library of candidate nonlinear functions is as follows

$$\boldsymbol{\Theta}^T(\mathbf{X}) = \begin{bmatrix} \text{---} 1 \text{---} \\ \text{---} \mathbf{X} \text{---} \\ \text{---} \sin(\mathbf{X}) \text{---} \\ \text{---} \cos(\mathbf{X}) \text{---} \end{bmatrix} \quad (4)$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \dots \\ \mathbf{x}_n \end{bmatrix} \quad (5)$$

Our dynamical model may be written in the following form

$$\dot{\mathbf{X}} = \boldsymbol{\Xi} \boldsymbol{\Theta}^T(\mathbf{X}) \quad (6)$$

The matrix $\boldsymbol{\Xi}$ in the above library is sparse for most dynamical systems (including power system dynamics). For identification of $\boldsymbol{\Xi}$, the sparse (symbolic) regression may be applied [4-6]. The selection of the library of candidate dynamics is a crucial choice in the SINDy algorithm. Note that in a power system case mixed nonlinear (trigonometric) multiplications of variables are possible (for example, $V\sin(\delta - \theta)$ and $V\cos(\delta - \theta)$ —see [9]). The library may be simply extended to support other nonlinearities (for example a saturation).

Each row of (6) represents a row in (1') and a sparse vector of coefficients ξ_k corresponding to the k -th row of $\boldsymbol{\Xi}$ can be found using a sparse regression algorithm [4], as

$$\xi_k = \underset{\xi_k}{\operatorname{argmin}} \|\dot{\mathbf{x}}_k - \xi_k \boldsymbol{\Theta}^T(\mathbf{X})\|_2^2 + \lambda^2 \|\xi_k\|_1 \quad (7)$$

where $\mathbf{x}_k$ represents k th row of $\mathbf{X}$ (5). The penalizing $\|\cdot\|_1$ term promotes sparsity in the vector ξ_k . The parameter λ is selected to identify the optimal model that best balances low model complexity with accuracy. A coarse sweep of λ is performed to identify the rough order of magnitude, where terms are eliminated and where error begins to increase [7].

IV. THE SINDy WITH MEASUREMENTS

The SINDy model can be extended to include input "measurements" as a part of the DAEs prediction model. These, in turn, can also come from a fixed (say physics-based) portion of the model. Then, the nonlinear dynamical system with input measurements ($\mathbf{y}$) from (1') is

$$\dot{\mathbf{x}} = \mathbf{f}''(\mathbf{x}, \mathbf{y}, t) \quad (1'')$$

$$\mathbf{y} = \mathbf{h}''(\mathbf{x}, t) \quad (3'')$$

The SINDy algorithm is simply generalized to include actuation, as this merely requires building a larger library $\boldsymbol{\Theta}^T(\mathbf{X}, \mathbf{Y})$ of candidate functions that include the measurement set ($\mathbf{Y}$)

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \mathbf{y} \end{bmatrix} = \boldsymbol{\Xi} \boldsymbol{\Theta}^T(\mathbf{X}, \mathbf{Y}) \quad (6')$$

$$\xi_k = \underset{\xi_k}{\operatorname{argmin}} \left\| \begin{bmatrix} \dot{\mathbf{x}} \\ \mathbf{y} \end{bmatrix}_k - \xi_k \boldsymbol{\Theta}^T(\mathbf{X}, \mathbf{Y}) \right\|_2^2 + \lambda^2 \|\xi_k\|_1 \quad (7')$$

V. THE SINDy ALGORITHM AND CONVERGENCE

Eq. (7') can be rewritten as

$$\mathbf{x} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda^2 \|\mathbf{x}\|_1 \quad (7'')$$

where $\mathbf{x} \in \mathbb{R}^n$ is unknown ($\mathbf{x} \equiv \xi_k$) associated with vector of observations ($\mathbf{b} \equiv \begin{bmatrix} \dot{\mathbf{x}} \\ \mathbf{y} \end{bmatrix}_k$) in $\mathbb{R}^m$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a matrix with $m \geq n$ [$\mathbf{A} \equiv \boldsymbol{\Theta}^T(\mathbf{X}, \mathbf{Y})$].

The SINDy algorithm is [8]

- Calculate

$$\mathbf{x}^0 = \mathbf{A}^\dagger \mathbf{b} \quad (8)$$

where $\mathbf{A}^\dagger$ is a matrix $\mathbf{A}$ pseudo-inverse.

- Solve

$$\mathbf{S}^k = \{j \in [n]: |\mathbf{x}_j^k| \geq \lambda\}, \quad k \geq 0 \quad (9)$$

$$\mathbf{x}^{k+1} = \underset{\mathbf{x} \in \mathbb{R}^n: \operatorname{supp}(\mathbf{x}) \subseteq \mathbf{S}^k}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2, \quad k \geq 0 \quad (10)$$

The convergence results of the SINDy algorithm may be summarized in [8, Theorem 2.10]:

- $\mathbf{x}^k$ converges to a fixed point of the iterative scheme (8)-(10) in at most n steps;
- a fixed point of the iterative scheme is a local minimizer of the objective function (7').
- a global minimizer of objective function (7') is a fixed point of the iterative scheme;
- $\mathbf{x}^k$ strictly decreases the objective function (7') unless the iterates are stationary.

VI. APPLICATION

We consider transient stability-related models in the DAE form in a Matlab-derived simulation environment. Our Matlab code is fully general in the sense that it allows for a variety of recorded measurements:

- rotor angle, $\delta(t)$ and speed, $\omega(t)$;
- active, $P_g(t)$ and reactive, $Q_g(t)$ power generations;
- nodal voltage magnitude, $V(t)$ and angle, $\theta(t)$.

These measurements are obtained by simulations on the IEEE 14-bus test example [9, Fig. 2.4]. We start transients following the three-phase short circuit in bus 4, which cleared after 250 ms. Note that the size of the test system is mostly irrelevant for this study because the proposed methodology is

based only on local measurements at the connection point with the rest of the power system. We deal with the model of a synchronous generator (SG)—for its mathematical description and used parameters, see [9].

The proposed model identification approach may be applied for the dynamic model reduction, where the time variation of state variables, $\mathbf{x}(t)$ is obtained from simulations by the detailed model. Derivatives $\dot{\mathbf{x}}_k(t)$ used in (7) are obtained numerically (backward differencing) from data.

A. Linear Terms in Right-side of Differential Equations

A first group of tests is performed with an aim for identification of four single (δ , $\dot{\omega}$, e'_q , and e'_d) and two paired (δ & $\dot{\omega}$ and e'_q & e'_d) differential equations in the fourth-order SG model from [9], using only linear terms in the differential

equations. The "fictive current measurements" i_d and i_q are then calculated, assuming that $e'_q(t)$ & $e'_d(t)$ are known.

Optimization problem (7) is solved by:

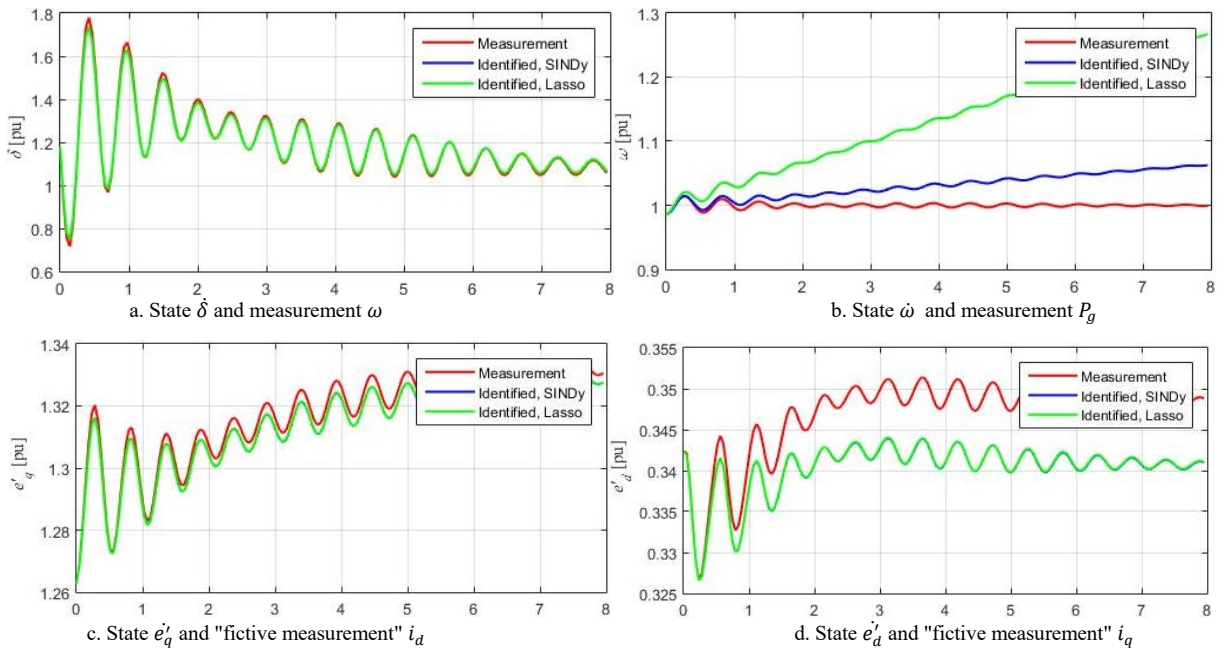
- 1) SINDy algorithm [4, 8],
- 2) Lasso software [10].

Note that output from Lasso for the constant terms is always equal to zero, as Lasso tries to remove those predictors (input variables) which no have additional information about the output (important for the state prediction). Thus, constant terms in the mechanical equations (Ω_b , $P_m/2H + D/2H$, and v_f/T'_{d0}) are excluded from Lasso optimization.

Results obtained for six different cases are presented in Table I, while the transient responses for SG in bus 4 are shown in Figure 1. For numerical integration, the Matlab's **ode45** function is extended with additional external inputs (in our case, with the measurements).

TABLE I: EXPECTED MODEL IDENTIFICATION FOR LINEAR TERMS IN RIGHT-SIDE OF DIFFERENTIAL EQUATIONS

States	Used Measur.	Expected coefficients			Identified, SINDy			Identified, Lasso		
		Constants (C)	Variables (V)	Measurements (M)	C	V	M	C	V	M
δ	ω	$-\Omega_b = -2\pi 60 = -376.99$	0	$\Omega_b = 2\pi 60 = 376.99$	-376.99	0	376.99	-376.99	0	376.99
$\dot{\omega}$	P_g	$\frac{P_m}{2H} + \frac{D}{2H} = 0.474$	$-\frac{D}{2H} = 0.198$	$-\frac{1}{2H} = 0.395$	0.474	-0.197	0.384	0.474	-0.177	0.376
e'_q	i_d	$\frac{v_f}{T'_{d0}} = 0.873$	$-\frac{1}{T'_{d0}} = -0.211$	$-\frac{x_d - x'_d}{T'_{d0}} = -0.857$	0.873	-0.210	-0.857	0.873	-0.210	-0.857
e'_d	i_q	0	$-\frac{1}{T'_{q0}} = -0.667$	$\frac{x_q - x'_q}{T'_{q0}} = 1.347$	0	-0.682	1.347	0	-0.682	1.347
δ	P_g	$-\Omega_b = -2\pi 60 = -376.99$	0	$\Omega_b = 2\pi 60 = 376.99$	-376.99	0	376.99	-376.99	0	376.99
$\dot{\omega}$	P_g	$\frac{P_m}{2H} + \frac{D}{2H} = 0.474$	$-\frac{D}{2H} = -0.198$	$-\frac{1}{2H} = -0.395$	0.433	-0.164	-0.384	0.474	0	-0.374
e'_q	i_d	$\frac{v_f}{T'_{d0}} = 0.873$	$-\frac{1}{T'_{d0}} = -0.211$	$-\frac{x_d - x'_d}{T'_{d0}} = -0.857$	0.873	-0.210	-0.857	0.873	-0.190	-0.857
	i_q			0			1.861			1.845
e'_d	i_d	0	$-\frac{1}{T'_{q0}} = -0.667$	0	0	-0.653	-0.783	0	-0.653	-0.523
	i_q			$\frac{x_q - x'_q}{T'_{q0}} = 1.347$			1.347			1.347



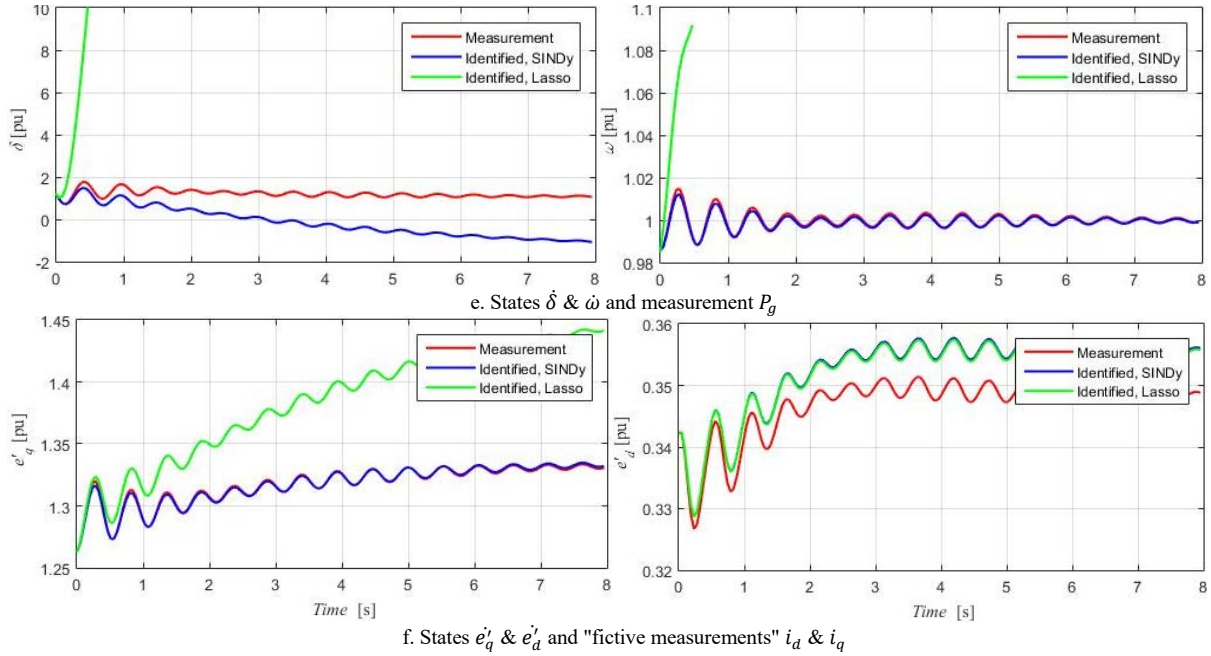


Figure 1. Transient responses for analyzed cases from Table I

B. Nonlinear (Trigonometric) Terms in Right-side of DAEs

The second group of tests is performed with the fourth-order dynamical model [9]; for this model, the measurements $P_g(t)$, $V(t)$ and $\theta(t)$ are simulations with PSAT-based dynamic model [9]. [Note that measurement $P_g(t)$ may be calculated and included directly into differential equations (as a function of measurements $V(t)$, $\theta(t)$, and state variables $e'_q(t)$ and $e'_d(t)$). However, in this case, the quadratic sine and cosine terms need to be included in the model library [9].

In this case, there are 7 variables in vector $\mathbf{X}$ (5) (δ , ω , e'_q , e'_d , V , θ , and P_g) and 10 elements in a library of candidate nonlinear functions in $\Theta^T(\mathbf{X})$ (4) (Const., δ , ω , e'_q , e'_d , P_g , V , θ , $V\cos(\delta - \theta)$, and $V\sin(\delta - \theta)$ —for details, see [9].

Model parameters expected from this analysis are shown in Table II. Results obtained by the previously mentioned two solvers: 1) SINDy algorithm [4, 8], and 2) Lasso software [10]) are compared regarding the Mean Squared Error (MSE) between measured and identified state variables:

- | | |
|--------------------|------------------|
| 1) SINDy algorithm | MSE = 262.496 |
| 2) Lasso software | MSE = 757064.255 |

The results for both solvers are shown in Figure 2. From these results, we may observe that the identification matrix (Ξ) is the full matrix and MSE is still very high (especially for Lasso algorithm); a sparsity in both software is additionally induced through the use of ℓ_1 regularization in (7). The number of non-zero elements in the identification matrix (Ξ) in Lasso software may be controlled by the parameter DFmax (maximum number of non-zero coefficients in the model). The reported MSE is still high in both cases. Additionally, the identified model in terms of elements is different from the one expected in Table II; this is especially important for trigonometric terms.

From the above results, we can conclude that the list of parameters for identification needs to be shortened. While we omit the details here for space reasons, by fixing zero elements in mechanical equations, the calculated MSE (by Matlab

software) is decreased to 0.048. Optimization in SINDy algorithm (7) is least-square based and the measurement noise does not have significant influence on the optimization results.

VII. CONCLUSION

In this paper, we have demonstrated the applicability of the sparse identification of nonlinear dynamics (SINDy) algorithm for data-driven identification of dynamic models in power systems. The SINDy model has been extended to discover the nonlinear model (described by state multiplications and trigonometric terms) with exogenous signals measurements. Since the proposed extension only identifies a few "active" terms in the dynamic responses, it requires less data than other leading machine learning techniques and prevents overfitting. The proposed model identifies strongly nonlinear systems from local measurement data only, and the model identification is fast enough to discover models in real-time, even in response to abrupt changes to the model (for example, for outages or short-circuits). Note that a critical step in the proposed model is the choice of library functions, which is based on the expert knowledge. A poor choice of library functions will generally yield a non-sparse model.

Although the preliminary application of the proposed model is encouraging, this work motivates several future extensions that are possible within the developed framework:

- Identification of high-dimensional dynamic models with a low-dimensional attractor (the model reduction).
- The model identification with limited measurement set.
- Inclusion of physical constraints in a known model structure, which may benefit real-time identification (constrained sparse identification [11]). It is possible to use a constrained least-squares algorithm, for example, to enforce parameter limits. Enforcing constraints has the potential to further reduce the amount of data required to identify models, and to improve numeric in some cases.
- Regression models for the dynamic response prediction.

The proposed framework has significant potential for real-time identification of nonlinear power system dynamic models, where the transient responses may be generated continuously by the load/generation increments due to the daily load/generation profiles. The ability to identify accurate and efficient models with small amounts of data may be a key factor of recovery in time-critical scenarios, such as the model changes that could lead to the power system instability.

REFERENCES

- [1] Transtrum, M.K., Sarić, A.T., Stanković, A.M.: "Information geometry approach to verification of dynamic models in power systems", IEEE Trans. Power Systems, 2018, **33** (1), pp. 440–450.
- [2] Daniels B.C., Nemenman, I.: "Automated adaptive inference of phenomenological dynamical models", Nature Communications, pp. 1–8.
- [3] Brunton, S.L., Proctor, J.L., Kutz, J.N.: "Discovering governing equations from data by sparse identification of nonlinear dynamical systems", Proc. of the National Academy of Sciences (PNAS) of the United States of America, 2016, **113** (15), pp. 3932–3937.
- [4] Brunton, S.L., Proctor, J.L., Kutz, J.N.: "Sparse identification of nonlinear dynamics with control (SINDy)", IFAC-Papers On Line, 2016, **49** (18), pp. 710–715.
- [5] Kaiser, E., Kutz, J.N., Brunton, S.L.: "Sparse identification of nonlinear dynamics for model predictive control in the low-data limit", Proc. of the

Royal Society A: Mathematical, Physical and Engineering Sciences, 2018, **474** (2219), pp. 1–25.

- [6] Loiseau, J.C., Noack, B.R., Brunton, S.L.: "Sparse reduced-order modelling: sensor-based dynamics to full-state estimation", Journal of Fluid Mechanics, 2018, **844**, pp. 459–490.
- [7] Mangan, N.M., Askham, T., Brunton, J.L., Kutz, J.N., Proctor, S.L.: "Model selection for hybrid dynamical systems via sparse regression", Proc. of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2019, **475** (2223), pp. 1–22.
- [8] Zhang, L., Schaeffer, H.: "On the convergence of the SINDy algorithm", Multiscale Model. Simul., 2019, **17** (3), pp. 948–972.
- [9] Milano, F.: *Power System Modelling and Scripting*. London, UK: Springer, 2010.
- [10] Tibshirani, R.: "Regression shrinkage and selection via the Lasso", J. of the Royal Statistical Society, Series B, 1996, **58**(1), pp. 267–288.
- [11] Loiseau, J.C., Brunton, S.L.: "Constrained sparse Galerkin regression", Journal of Fluid Mechanics, 2018, **838** (10), pp. 42–67.

APPENDIX: SG INPUT DATA

Input SG's parameters for simulations are as follows:

$\Omega_b = 2\pi f = 120\pi$; $\omega_0 = 1$ pu; $P_m = 0.7$ pu; $D = 0.5$ pu; $2H = 2.53$ MWs/MVA; $v_f = 4.1468$; $T'_{d0} = 4.75$ s; $T'_{q0} = 1.5$ s; $x_d = 5$ pu; $x_q = 4.88$ pu; $x'_d = 0.928$ pu; $x'_q = 2.86$ pu.

TABLE II: EXPECTED MODEL IDENTIFICATION FOR NONLINEAR (TRIGONOMETRIC) TERMS IN RIGHT-SIDE OF DIFFERENTIAL EQUATIONS

Const.	δ	ω	$\dot{e}'_q$	$\dot{e}'_d$	P_g	V	θ
	$-\Omega_b = -2\pi 60 = -376.99$	$\frac{P_m}{2H} + \frac{D}{H} = 0.4743$	$\frac{v_f}{T'_{d0}} = 0.8730$	0	0	0	0
δ	0	0	0	0	0	0	0
ω	$\Omega_b = 2\pi 60 = 376.99$	$-\frac{D}{2H} = -0.1976$	0	0	0	0	0
e'_q	0	0	$-1 - \frac{(x_d - x'_d)}{x'_d T'_{d0}} = -1.134$	$-1 - \frac{(x_q - x'_q)}{x'_q T'_{q0}} = -1.1375$	0	0	0
e'_d	0	0	0	0	0	0	0
P_g	0	$-\frac{1}{2H} = -0.3953$	0	0	1	0	0
V	0	0	0	0	0	1	0
θ	0	0	0	0	0	0	1
$V\cos(\delta - \theta)$	0	0	$\frac{(x_d - x'_d)}{x'_d T'_{d0}} = 0.9238$	0	0	0	0
$V\sin(\delta - \theta)$	0	0	0	$\frac{(x_q - x'_q)}{x'_q T'_{q0}} = -0.4709$	0	0	0

	'delta_dot'	'omega_dot'	'Elq_dot'	'Eld_dot'	'Pg'	'V'	'theta'
'Const'	[-376.5125]	[0]	[0.8549]	[0]	[0]	[0]	[0]
'delta'	[0]	[0]	[0]	[0]	[0]	[0]	[0]
'omega'	[383.4824]	[-0.2474]	[0]	[0]	[0]	[0]	[0]
'Elq'	[0]	[0]	[0]	[0.3763]	[0]	[0]	[0]
'Eld'	[-19.5531]	[0.7758]	[-1.2372]	[-0.6623]	[0]	[0]	[0]
'Pg'	[0]	[-0.3839]	[-0.5695]	[0.2155]	[1.0000]	[0]	[0]
'V'	[0]	[0]	[0.7672]	[0.0741]	[0]	[1.0000]	[0]
'theta'	[0]	[0]	[0.2847]	[0]	[0]	[0]	[1.0000]
'V*cos(delta-theta)'	[0]	[0]	[0]	[0]	[0]	[0]	[0]
'V*sin(delta-theta)'	[0]	[0]	[0]	[0]	[0]	[0]	[0]

a. SINDy algorithm

	'delta_dot'	'omega_dot'	'Elq_dot'	'Eld_dot'	'Pg'	'V'	'theta'
'Const'	[0]	[0]	[0]	[0]	[0]	[0]	[0]
'delta'	[0]	[0]	[0]	[0]	[0]	[0]	[0]
'omega'	[378.4266]	[0]	[0]	[0.1157]	[0]	[0]	[0]
'Elq'	[0]	[0]	[-0.5558]	[0]	[0]	[0]	[0]
'Eld'	[-15.9290]	[0]	[-0.1767]	[-0.0689]	[0]	[0]	[0]
'Pg'	[0]	[-0.3744]	[-0.5790]	[0.1250]	[0.9708]	[0]	[0]
'V'	[0]	[0]	[0.6595]	[0]	[0]	[0.9708]	[0]
'theta'	[0]	[0]	[0]	[0]	[0]	[0]	[0.9708]
'V*cos(delta-theta)'	[0]	[0]	[0]	[0.0598]	[0]	[0]	[0]
'V*sin(delta-theta)'	[0]	[0]	[0]	[0]	[0]	[0]	[0]

b. Lasso software

Figure 2. Model identification for nonlinear (trigonometric) terms in right-side of differential equations.