Electric Power Systems Research 189 (2020) 106824

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Interleaving physics- and data-driven models for power system transient dynamics

Aleksandar M. Stanković[a,*], Aleksandar A. Sarić[a], Andrija T. Sarić[b], Mark K. Transtrum[c]

[a] Dept. of Electrical and Computer Engineering, Tufts University, Medford, MA, USA
[b] Dept. of Power, Electronics and Com. Eng., Faculty of Technical Sciences, Novi Sad, Serbia
[c] Dept. of Physics and Astronomy, Brigham Young University, Provo, UT, USA

ARTICLE INFO

Keywords:
Power system dynamics
Modeling
Physics-based models
Data-driven models
Compressed sensing
Koopman modes

ABSTRACT

The paper explores interleaved and coordinated refinement of physics- and data-driven models in describing transient phenomena in large-scale power systems. We develop and study an integrated analytical and computational data-driven gray box environment needed to achieve this aim. Main ingredients include computational differential geometry-based model reduction, optimization-based compressed sensing, and a finite approximation of the Koopman operator. The proposed two-step procedure (the model reduction by differential geometric (information geometry) tools, and data refinement by the compressed sensing and Koopman theory based dynamics prediction) is illustrated on the multi-machine benchmark example of IEEE 14-bus system with renewable sources, where the results are shown for doubly-fed induction generator (DFIG) with local measurements in the connection point. The algorithm is directly applicable to identification of other dynamic components (for example, dynamic loads).

1. Introduction

Development of mathematical techniques that operate directly on observations or measurements (i.e., data-driven methods) is of increasing relevance for large-scale power systems. Reasons for unavailability of verified system-level equation-based models needed for studies of stability, dynamic performance and restoration vary, and include: 1. component (particularly load) variations, 2. obsoleteness/undocumented modifications, and 3. models in the form of computer code or tabulated experimental data.

We envision development of a two-step procedure, in which nominal dynamical models are tested for practical identifiability in the first step using differential geometric tools, and then are appended by the data-driven refinement on the second step, using tools from compressed sensing and Koopman operator theory.

Equation-free approaches are effective for many power electronic and power systems [1,2], where this method assumes full access to the lower level models that can be precisely initialized and simulated at will. We are interested in approaches that would be effective in more restricted situations in which only some data (experimental or simulated) are available without access to the detailed simulation model or to the system itself.

Diffusion maps for manifold learning [3] are an example of methods that by-pass the need to precisely select variables and parameters as well as the need to derive accurate, closed-form equations (i.e., white or gray-box approaches).

Available SCADA and especially Phasor Measurement Unit (PMU) based measurements provide a very large volume of data (typically the transient responses) in modern power systems. It is a common expectation that only a few parameters are needed to characterize the modes that are active at a given operating point, thus suggesting a possibility of significant data size reduction. Data compression typically relies on sparsity of the signal of interest in a transformed basis (for example, in Fourier transform basis).

In this paper we explore the interplay among several concepts needed to achieve coordinated physics- and data-driven modeling of power system dynamics.

The outline of the paper is as follows: in Section 2 is provides the problem formulation; Section 3 describes the basis for dynamic model reduction by differential geometric (information geometry) tools, while Section 4 describes the compressed sensing algorithm; Section 5 provides the details of the data-driven Koopman operator based prediction; the overall procedure is applied to the benchmark 14-bus system in Section 6, and Section 7 presents conclusions.

* Corresponding author.
E-mail addresses: astankov@ece.tufts.edu (A.M. Stanković), asaric@uns.ac.rs (A.T. Sarić), mktranstrum@byu.edu (M.K. Transtrum).

https://doi.org/10.1016/j.epsr.2020.106824
Received 30 September 2019; Received in revised form 30 March 2020; Accepted 2 August 2020

2. Problem formulation

The equation-driven framework for dynamical power system models is based on (nonlinear) differential and algebraic equations (DAEs), respectively [4]

$$d\boldsymbol{x}/dt = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{p}, t) \tag{1}$$

$$\boldsymbol{0} = \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{p}, t) \tag{2}$$

where $\boldsymbol{x}$ is the vector of state variables, $\boldsymbol{z}$ are the algebraic variables, $\boldsymbol{p}$ are parameters, and t is the (scalar) time variable.

System measurements are assumed to be of the form

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{p}, t) \tag{3}$$

In realistic conditions, measured ($\boldsymbol{y}^m$) and calculated ($\boldsymbol{y}^c$) values of variables in measurement points may be different, due to: uncertain parameters (vector $\boldsymbol{p}$), un-modeled dynamics, different initial conditions, etc.

System identification is the standard approach to solve the above problem via parameter optimization as

$$\boldsymbol{p} = \arg\min_{\boldsymbol{p}} \|\boldsymbol{y}^m - \boldsymbol{y}^c\|_2 \tag{4}$$

subject to (1–3). An alternative is based on the equation-free models [1,2,5]. However, both approaches have a drawback in practical applications, due to the sensitivity of the optimized solution to model parametrization and to initial conditions.

In this paper we advance a hybrid approach, in which a portion of the equation-based model is retained. This portion is certified via differential-geometric tools to be robustly identifiable in a given measurement structure. We propose a gray box modeling approach, consisting of equation-based modeling for the retained part in the reduction procedure

$$\boldsymbol{y}^{c,red} = \boldsymbol{\phi}(\boldsymbol{x}^{red}) \tag{5}$$

and data-driven modeling for the difference between measured ($\boldsymbol{y}^m$) and calculated ($\boldsymbol{y}^{c,\ red}$) values, or

$$\Delta\boldsymbol{y} = \boldsymbol{y}^m - \boldsymbol{y}^{c,red} \tag{6}$$

where the transformation $\boldsymbol{y}^{c,red} = \boldsymbol{\zeta}(\boldsymbol{y}^m)$ is possibly nonlinear, and unknown in advance.

Flow-chart of the proposed gray box environment with three main ingredients (Information Geometry-based the dynamic model reduction, compressed sensing, and the Koopman theory-based dynamic prediction) is shown in Fig. 1.

3. Reduction of the dynamic model

Information geometry refers to a geometric interpretation of statistics in which a parametric model is interpreted as a manifold [6]. Parameters of the model act as coordinates, which effectively shift emphasis from model parameters to model behaviours, i.e., model properties that are invariant to re-parameterization [4,6,7]. Combining this theory with computational differential geometry leads to a rich toolbox for exploring global properties of parametric models.

The fundamental quantity in Information Geometry is the Jacobian matrix $\boldsymbol{J_p}(t) = \partial\boldsymbol{h}(t)/\partial\boldsymbol{p}$. The Jacobian is calculated by solving the first order sensitivities of (1–3).

The Fisher Information Matrix (FIM) is the Riemannian metric on the manifold and is constructed from the Jacobian as $\boldsymbol{I} = \boldsymbol{J}_p^T\boldsymbol{J}_p$. The metric quantifies the local, linear properties of the model manifold by defining an inner product between vectors in parameter space that is induced by properties in the model's behaviour space.

Local quantities are extended to a nonlocal analysis through a *geodesic*. Geodesics are analogues of straight lines on curved surfaces. They are constructed numerically by solving a second order ordinary differential equation (ODE) in parameter space

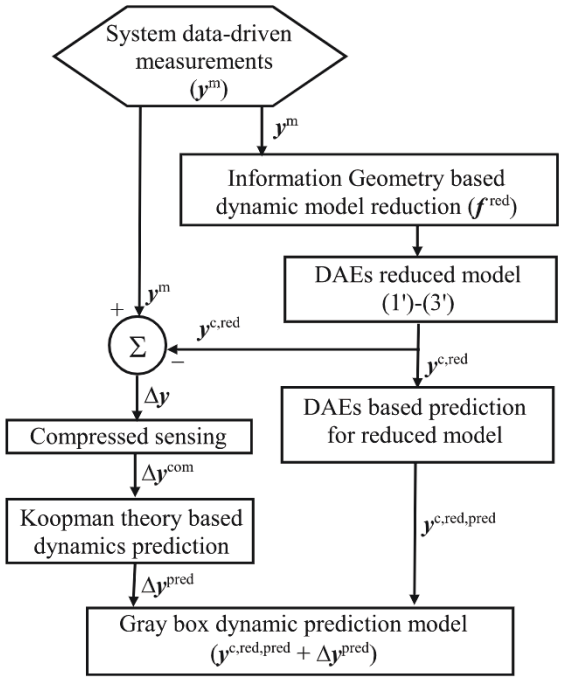

Fig. 1. Flow-chart of the proposed gray box environment.

$$\frac{\partial^2 p^i}{\partial\tau^2} = \sum_{j,k}\Gamma^i_{jk}\frac{\partial p^j}{\partial\tau}\cdot\frac{\partial p^k}{\partial\tau};\ \Gamma^i_{jk} = \sum_{\ell,m}(\boldsymbol{I}^{-1})^{i\ell}\frac{\partial y_m}{\partial p^\ell}\cdot\frac{\partial^2 y_m}{\partial p^j\partial p^k} \tag{7}$$

The geodesic is parameterized by its arc length as measured on the model manifold, denoted by τ. Γ are the Christoffel symbols [6], which are formally expressed in terms of first and second order parametric sensitivities of (1–3). However, the unique structure of the geodesic Eq. (7) enables for some numerical shortcuts. Because of the sums over the indices j and k in (7), it is not necessary to calculate all the Christoffel symbols. Instead, the relevant combinations can be calculated directly and efficiently following methods described in [4,6,7 and references therein].

To solve (7), it is also necessary to specify initial conditions. A unique solution to the geodesic equation is determined by specifying a starting point and direction in parameter space. Here, we take the initial values to be the "true" parameter values and the initial direction to be the eigenvector of the FIM with smallest eigenvalue. Our analysis then proceeds by first solving (1–3) along with their first order sensitivities, and constructing the Jacobian and FIM. Next the relevant combinations of the Christoffel symbols are evaluated and combined into the right-hand side of (7). These steps become the right-hand-side routine of a numerical ODE integrator that repeats these steps as it marches through parameter space along the geodesic.

Often, the geodesic will exhibit a "finite-time" singularity, where "time" refers to the parameter τ, not the time parameter, t in (1–3). When this occurs, the geodesic extends parameters to extreme values, such infinity or zero, for finite values of τ. We then say the corresponding parameters are practically unidentifiable and that the model manifold has a boundary. Furthermore, the complete boundary of a model manifold is typically made up of several "faces" (like the faces of a high-dimensional polyhedron). Each face corresponds to a different parameter being pushed to an extreme limit. A knowledge of these "faces" on the model manifold is useful for identify-ability analysis, model reduction, and experimental design.

For the case of a doubly-fed induction generator ((DFIG) that we consider in detail later (see Appendix for detailed equations), the following modeling recommendations were obtained from above analysis (for more details, see [7]):

- Inertia (H_m) is well-conditioned and can be estimated from the motion equation [7, Table I, **Case 1**] .
- Time constants (T_p and T_e) are ill-conditioned (two zero eigenvalues dominantly influence their participation factors) and cannot be estimated simultaneously from electrical equations and available measurements [7,Table II, **Cases 2 − 4**].
- Reactances (x_s and x_μ) are well-conditioned and may be estimated simultaneously from electrical equations and available measurements [7,Table 1, **Case 5**].
- Reactance x_s and time constant T_e are both non-identifiable due to a correlation in which x_s becomes zero and T_e becomes infinite. Similarly, there is also a correlation in which x_μ and x_s become zero and T_e becomes infinite.

The reduced DAEs model may be represented as

$$d\boldsymbol{x}^{red}/dt = \boldsymbol{f}^{red}(\boldsymbol{x}^{red}, \boldsymbol{z}, \boldsymbol{p}, t) \tag{1′}$$

$$\boldsymbol{0} = \boldsymbol{g}(\boldsymbol{x}^{red}, \boldsymbol{z}, \boldsymbol{p}, t) \tag{2′}$$

$$\boldsymbol{y}^{c,red} = \boldsymbol{h}(\boldsymbol{x}^{red}, \boldsymbol{z}, \boldsymbol{p}, t) \tag{3′}$$

Note that vector of algebraic variables ($\boldsymbol{z}$) is not reduced in this model. This conclusion is not general, and it depends on the type of the model that is a candidate for reduction.

4. Compressed sensing

The key idea behind data compression is the hope that in a transformed basis (such as the one obtained by Fourier transform) the data will be sparse. Instead of collecting high dimensional measurements and then compressing, it is possible to acquire a few ("compressed") measurements and then solve for the sparsest high-dimensional signal that is consistent with available measurements. This procedure enables a set of measurements to be recovered from what would otherwise be highly incomplete information. Compressibility means that for the recorded signal with an appropriate basis, only a few elements of the basis ("modes") are needed ("active"), reducing the number of numerical values that must be stored for an accurate representation. This means that a compressible signal $\boldsymbol{x} \in \mathbb{R}^n$ may be written as a sparse vector $\boldsymbol{s} \in \mathbb{R}^n$ on a transform basis $\boldsymbol{\Psi} \in \mathbb{R}^{n\times n}$ [8,9]

$$\boldsymbol{x} = \boldsymbol{\Psi s} \tag{8}$$

Therefore, once a signal is compressed, one needs only store sparse vector $\boldsymbol{s}$ rather than the entire $\boldsymbol{x}$. If signal $\boldsymbol{x}$ (with n measurements) is K-sparse in $\boldsymbol{\Psi}$, measurements $\boldsymbol{y} \in \mathbb{R}^p$ with $K < p < < n$ are given as

$$\boldsymbol{y} = \boldsymbol{\Phi x} \tag{9}$$

where the measurement matrix $\boldsymbol{\Phi} \in \mathbb{R}^{p\times n}$ represents a set of p linear measurements on the state $\boldsymbol{x}$. The choice of $\boldsymbol{\Phi}$ is of critical importance in compressed sensing and will be discussed in more details in Section VI.

Substituting (9) to (8) we have

$$\boldsymbol{y} = \boldsymbol{\Phi\Psi s} = \boldsymbol{\Theta s} \tag{10}$$

where this system of equations is underdetermined, since there are potentially infinitely many consistent solutions ($\boldsymbol{s}$).

The sparsest solution ($\hat{\boldsymbol{s}}$) satisfies the following optimization problem:

$$\hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \|\boldsymbol{s}\|_1 \tag{11a}$$

subject to:

$$\boldsymbol{y} = \boldsymbol{\Phi\Psi s} = \boldsymbol{\Theta s} \tag{11b}$$

where $\|\cdot\|_1$ is the ℓ_1 norm, given by $\|\boldsymbol{s}\|_1 = \sum_{k=1}^{n} |s_k|$. Note that the ℓ_1 minimum-norm solution is sparse, while the ℓ_2 minimum-norm solution is not (see Fig. 6). Alternative formulations of (11) (for example, with quadratic constraints, bounded residual correlation and other) may be found in [8].

The transient responses in power systems typically are recorded with a wide range of timestamps (from several seconds for SCADA measurements, over 50-100 ms for calculated values by transient analysis software (for example PSS/E), to microseconds for PMU-based measurements). Intuitively, we expect that detailed transient responses may be identified from a reduced number of saved discrete points. Additionally, initial conditions tend to vary slowly during the daily load/generation profile variations.

We may conclude that there are two promising options for compression of recorded measurements:

- Reduction of the number of elements in vector $\boldsymbol{s}$ (variable in transform basis). In our case, this is a reduction in the number of necessary Fast Fourier Transformation (FFT) coefficients.
- Reduction of sampling points in the time axis and along the initial condition axis (for suitably slowly varying signals).

Note that above algorithm for compressed sensing is applied for the difference between measured and calculated values (6) (see Fig. 1).

5. Data-driven Koopman operator

For real-valued observation (measurement) functions ϕ, $\mathbf{M} \to \mathbb{R}$ ($\mathbf{M}$ denotes the state space, and $\mathbb{R}$ denotes the scalar measurement space), which are elements of an infinite-dimensional Hilbert space, the dynamic model (1–3) may be re-written in compact form as [also see (9)]

$$y = \phi(\boldsymbol{x}) \tag{12}$$

The Koopman operator ($\mathcal{K}$) is an infinite-dimensional *linear* operator, acting on a Hilbert space of measurement functions (ϕ) as [8,10,11]

$$\mathcal{K}_t\phi = \phi o \boldsymbol{F}_t \tag{13}$$

where $\boldsymbol{F}_t$ is a transition function of states ($\boldsymbol{x}$) in state space ($\boldsymbol{M}$), determined by (1–3).

Schematic illustration of the Koopman operator for the nonlinear dynamical system is shown in Fig. 2.

The Koopman operator provides an alternative perspective for the evolution of measurements in discrete-time $y_k = \phi(\boldsymbol{x}_k)$ from (9), which is of the primary interest of this paper.

For the sampled-data system with time step Δt from (13) is

$$\phi(\boldsymbol{x}_{k+1}) = \mathcal{K}_{\Delta t}\phi(\boldsymbol{x}_k) \tag{14}$$

In other words, this operator defines an infinite-dimensional linear dynamical system that uses the observation (in our case the measurements) in the function of the state [$y_k = \phi(\boldsymbol{x}_k)$] to the next time step.

The Koopman operator is linear, a property which is based on the linearity of the addition operation in function space

$$\mathcal{K}_{\Delta t}(\alpha_1\phi_1(\boldsymbol{x}_k) + \alpha_2\phi_2(\boldsymbol{x}_k)) = \alpha_1\mathcal{K}_{\Delta t}\phi_1(\boldsymbol{x}_k) + \alpha_2\mathcal{K}_{\Delta t}\phi_2(\boldsymbol{x}_k) \tag{15}$$

The discrete-time Koopman operator, $\mathcal{K}_{\Delta t}\phi(\boldsymbol{x}_k)$ corresponding to eigenvalue λ from (14) satisfies

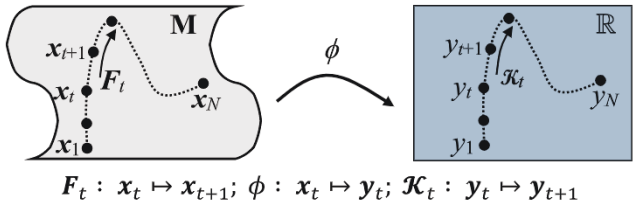

Fig. 2. Schematic illustration of the Koopman operator for the nonlinear dynamical system.

$$\mathcal{K}_{\Delta t}\phi(\boldsymbol{x}_k) = \lambda\phi(\boldsymbol{x}_k) \tag{16}$$

In the case of multiple measurements, these measurements may be arranged in a vector $\boldsymbol{\phi}(\boldsymbol{x})$

$$\boldsymbol{y} = \boldsymbol{\phi}(\boldsymbol{x}) = \begin{bmatrix} \phi_1(\boldsymbol{x}) \\ \phi_2(\boldsymbol{x}) \\ \vdots \\ \phi_P(\boldsymbol{x}) \end{bmatrix} \tag{17}$$

Individual measurements may be expanded in terms of eigenfunctions $\varphi_j(\boldsymbol{x})$, providing the basis for Hilbert space

$$\boldsymbol{\phi}(\boldsymbol{x}) = \sum_{j=1}^{\infty} \varphi_j(\boldsymbol{x})\boldsymbol{v}_j \tag{17}$$

where $\boldsymbol{v}_j$ is the j-th Koopman mode associated with eigenfunction $\varphi_j(\boldsymbol{x})$.

Based on the decomposition from (18), it is possible to represent the dynamics of the measurements $[\boldsymbol{\phi}(\boldsymbol{x})]$ as

$$\begin{aligned} \boldsymbol{y}_k &= \boldsymbol{\phi}(\boldsymbol{x}_k) = \mathcal{K}_{\Delta t}^k\boldsymbol{\phi}(\boldsymbol{x}_0) = \mathcal{K}_{\Delta t}^k\sum_{j=1}^{\infty}\varphi_j(\boldsymbol{x}_0)\boldsymbol{v}_j \\ &= \sum_{j=1}^{\infty}\mathcal{K}_{\Delta t}^k\varphi_j(\boldsymbol{x}_0)\boldsymbol{v}_j = \sum_{j=1}^{\infty}\lambda_j^k\varphi_j(\boldsymbol{x}_0)\boldsymbol{v}_j. \end{aligned} \tag{18}$$

The sequence of triples, $\{\lambda_j, \varphi_j, \boldsymbol{v}_j\}_{j=1}^{\infty}$ is the Koopman mode decomposition [10,11].

The Koopman operator provides a global linear (infinite dimensional) representation of the nonlinear dynamical system. Dynamic mode decomposition approximates the Koopman operator with a best-fit (finite) linear model from time-dependent measurements. There are several algorithms for identifying Koopman embeddings and eigenfunctions from data, such as extended mode decomposition [12], or QR decomposition of the input snapshot matrix [13]. In this paper we apply the basic algorithm from [8].

Note that above algorithm for data-driven Koopman theory-based dynamic prediction is applied to the compressed difference between measured and values calculated by the reduced model, $\Delta\boldsymbol{y}^{com}$ (see Fig. 1).

6. Application

We have developed a Matlab- and PSAT-based [14] simulation environment, which implements the DAEs-based model, as in (1), (2). The original IEEE 14-bus test system with synchronous generators (SG – describing conventional units and interconnections) [14] is modified to include doubly-fed induction generator (DFIG – capturing prevalent type of wind plants today) in bus 6 and direct-drive synchronous generators (DDSG – used by industry to model solar plants and the new generation of wind) in bus 8, as shown in Fig. 3. Our software environment is general in the sense that it allows for a variety of

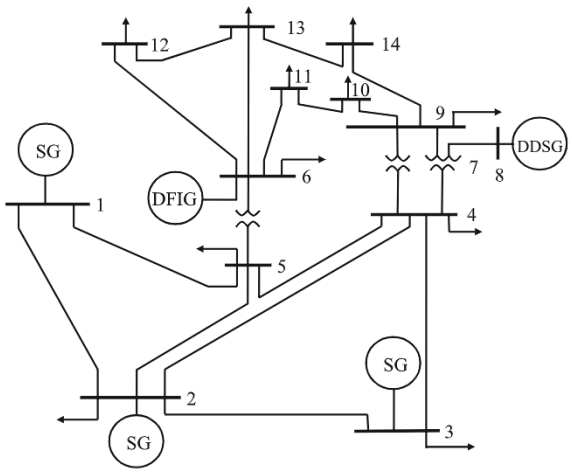

Fig. 3. Modified IEEE 14-bus test system with three types of resources (SG, DFIG, and DDSG).

measurements: rotor speed (ω_m), active (P_g) and reactive (Q_g) power generations, nodal voltage magnitude (V) and angle (θ). For four generating unit types [solar (S), wind (W), thermal (T), and hydro (H)] and four load types the different daily generation/load curves are supposed.

6.1. Dynamic model reduction

For the case of DFIG with three electrical parameters in (A1) (extension to the control parameters may be found in [4]), we identify the following manifold faces, each corresponding to a different unidentifiable parameter in the model (see Appendix for clarification): $x_\mu \to \infty$ and x_s, $T_e \to 0$, forming the reduced three-order model in (1').

Transient responses (recorded for 9 s, with time stamps each 0.05 s) are obtained by three-phase short-circuit on bus 4, following the fault clearing after 250 ms. Total 1440 (every 60 s, or 60 × 24) initial conditions are simulated by full (used as the measured values) and reduced DAEs-based models. Note that for storing only five measurements for the measured and calculated by the reduced model we need about 24 Mb. Additional storage capacity is needed for storing state variables. These results are shown in Fig. 4a, b. Differences and their surface between measured transients and transients calculated by the reduced model are also shown in Fig. 4 (subplots c, d).

From the results presented in Fig. 4 we conclude that the differences between measured ($\boldsymbol{y}^m$) and calculated (in our case by the reduced model) ($\boldsymbol{y}^c$) values [$\Delta\boldsymbol{y} = \boldsymbol{y}^m - \boldsymbol{y}^c$ in (3)] may be significant (in analyzed case 10-20 %, where the higher differences occur in the period right after the fault clearing). Similarly, these differences are challenging for identification, and typically require alternative methods, such as Koopman modes and/or Deep Neural Networks.

6.2. Compressed sensing

The optimization problem (11) is solved by CVX software [15]; several interesting cases (measurement responses) will be described in the sequel.

The basic results for differences (between measured and calculated values by reduced model) of the bus voltage magnitude measurements are omitted from the paper, due to space constraints [similar plot with ones in Fig. 4.c, d (for active powers)].

For initial solution ($\boldsymbol{x}_0$) two alternatives have been explored: (1) minimum energy $\boldsymbol{x}_0 = \boldsymbol{\Theta}\boldsymbol{\Theta}^T\boldsymbol{x}$ [9], and (2) percentage of retained largest FFT coefficients. In the simulations presented below, the second option was used, with the threshold of only 5 % retained the largest FFT coefficients. In the analyzed case, the number of elements in vector $\boldsymbol{x}$ ($\boldsymbol{x}_0$) is 260640. This is the total number of decision variables in the optimization (11), while the maximum number of linear constraints in (11) is also 260640 for the fully determined problem.

The differences (between measured and calculated values by reduced model) are transformed into a grayscale figure (normalized to the range 0-1), which is fully adapted for the compressed sensing algorithm. Grayscale transformation of the initial point ($\boldsymbol{x}_0$) is shown in Fig. 5, where only initial conditions from 500 to 1000 are shown (reduced initial condition's axis is shown, due to the large axes ratio − 1440/181 ≈ 7.95).

Solutions and their histogram of the ℓ_1-minimum norm for variables in the Fourier domain (s_j) for the fully determined problem are shown in Fig. 6. Reconstructed differences of the bus voltage magnitudes are shown in Fig. 7 (values are scaled to 0–1 range).

Based on the results presented in Fig. 6, only a small number of FFT coefficients is necessary for signal reconstruction, suggesting the possibility of extreme compression. One characteristic truncation case is shown in Fig. 8. Based on these results, we conclude that with only 5 % of FFT coefficients the differences of bus voltage magnitudes may be reconstructed properly.

Optimal truncation (optimal hard threshold) for (un)known noise and a rectangular matrix $\boldsymbol{X} \in \mathbb{R}^{n\times m}$ may be found in [16].

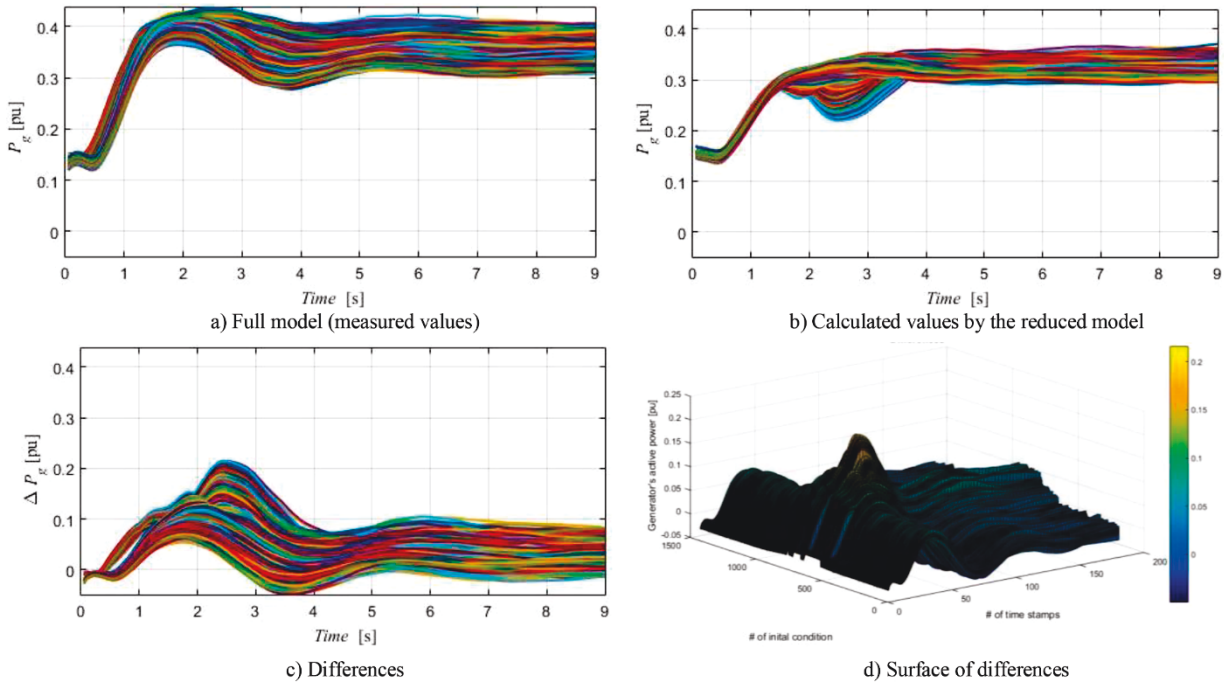

Fig. 4. Measured transient resonses and transient responses calculated by the reduced model, as well as their differences.

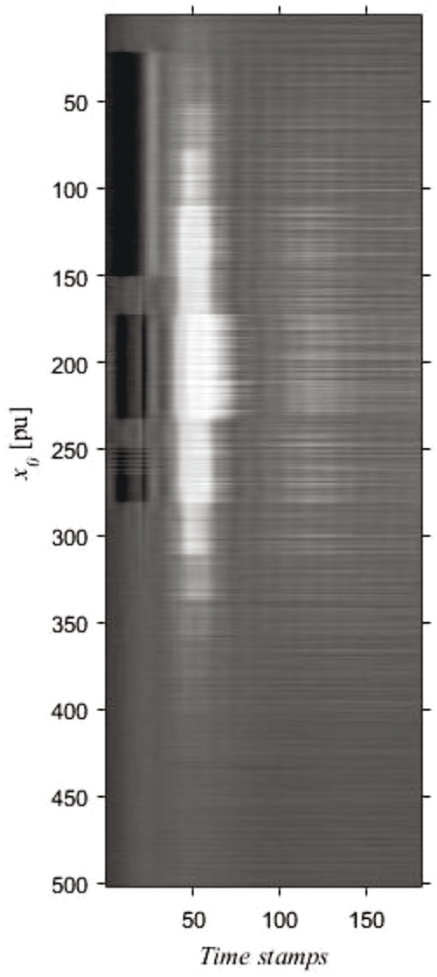

Fig. 5. Grayscale transformation of initial condition (x_0).

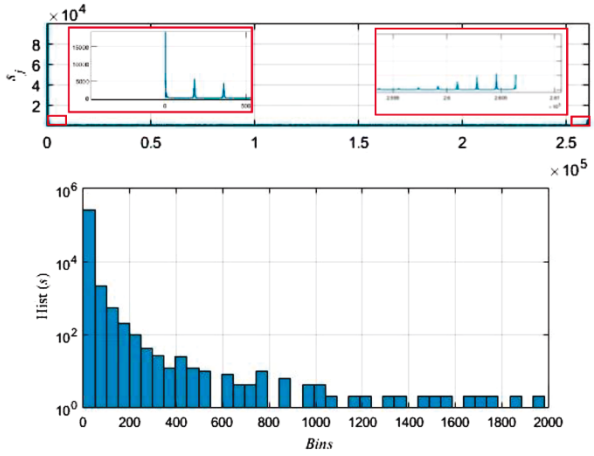

Fig. 6. ℓ_1-minimum norm solution (values in first panel, and histogram in second panel) to the fully determined problem (11).

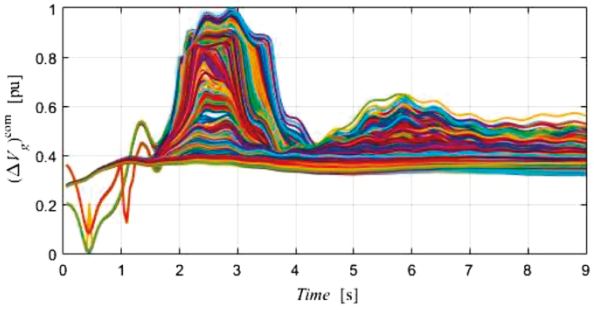

Fig. 7. Reconstructed differences (between measured and calculated values by the reduced model) of bus voltage magnitude for fully determined problem (scaled to the range 0-1).

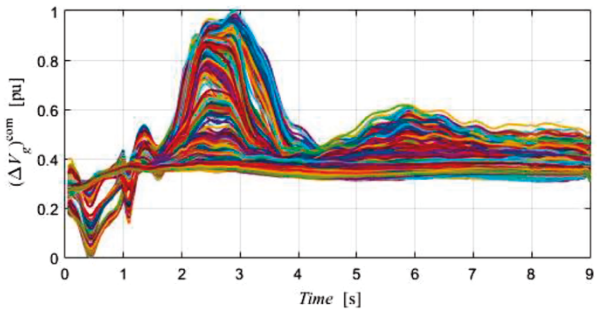

Fig. 8. Reconstructed differences (between measured and calculated values by the reduced model) of bus voltage magnitudes for underdetermined problem (scaled to the range 0-1) with truncation down to 5 %.

Reduction of sampling points can be performed in two axes: (1) time axis, reducing the number of timestamps for transient responses, and (2) initial condition axis, reducing the number of analyzed initial conditions.

After additional initial point reduction, where only every second point is retained, the reconstructed differences (between measured and calculated values by the reduced model) of bus voltage magnitudes are shown in Fig. 9. For reduction in both axes [panel b)], only 25 % of original points are used for the measurement reconstruction.

It is important to note that the reduced cases are not based on interpolation for reduced points, but the optimization problem is solved with the full dimension of vector ***s*** and a reduced number of constraints in (11).

6.3. Koopman analysis

Results obtained for predicted differences (between measured and calculated values by the reduced model) of reactive power by the compressed sensing and the Koopman modes are shown in Fig. 10; reconstructed results correspond to truncation down to only 4 %.

7. Conclusions

In this paper, we describe the analytical and computational environment needed to achieve coordinated physics- and data-driven modeling of power system dynamics. It combines three ingredients:

- Computational Information Geometry based reduction,
- Optimization-based compressed sensing,
- Approximation of the Koopman operator based prediction.

In future work, we plan to further quantify scaling properties of the overall algorithm, explore optimization in compressed sensing, and evaluate deep neural network-based environments for predicting dynamics in power systems.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

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.

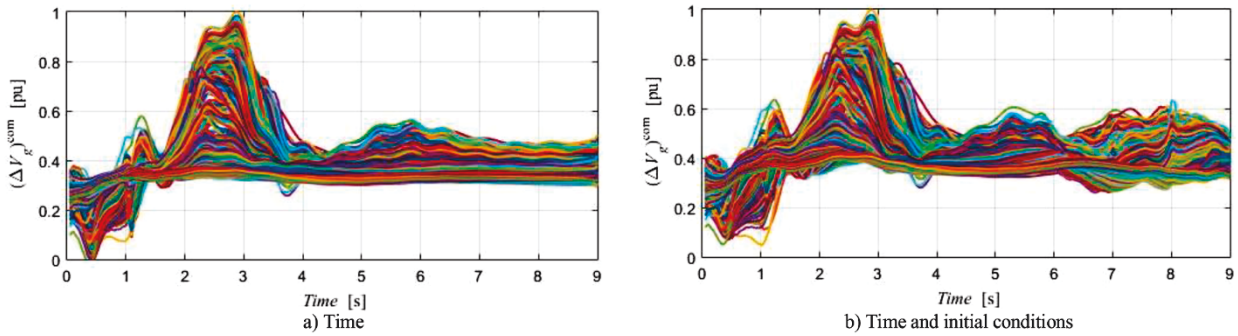

Fig. 9. Reconstructed differences (between measured and calculated values by the reduced model) of bus voltage magnitude with reduced time and initial condition sampling points (scaled to the range 0-1).

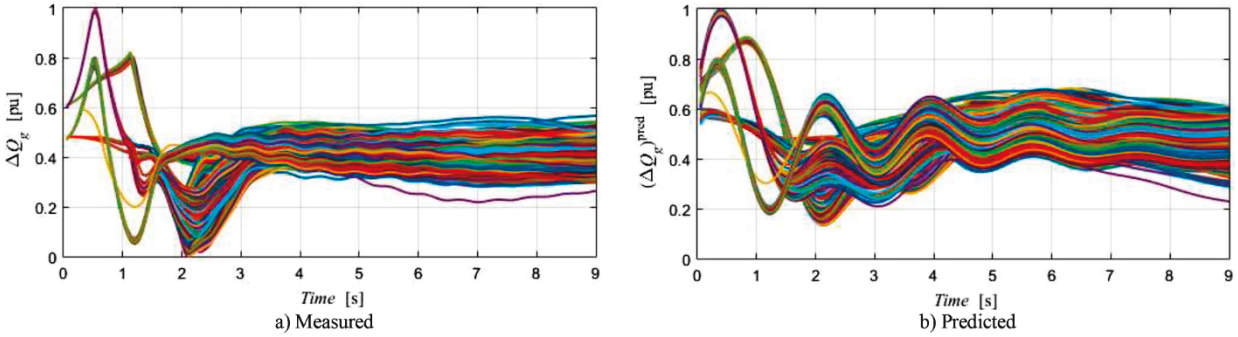

Fig. 10. Measured and predicted differences (between measured and calculated values by the reduced model) of reactive power obtained by compressed sensing and Koopman operator for different initial conditions (scaled to the range 0-1).

Appendix

Differential (motion and electrical) and algebraic equations for DFIG are described respectively as:

$$\boldsymbol{f} \Rightarrow \begin{cases} \dot{\omega}_m = \frac{\tau_m - \tau_e}{2H_m} \\ \dot{\theta}_p = \frac{1}{T_p}(K_p\phi(\omega_m - \omega_{ref}) - \theta_p) \\ \dot{i}_{rd} = K_V(V - V_{ref}) - \frac{V}{x_\mu} - i_{rd} \\ \dot{i}_{rq} = \frac{1}{T_e}\left(-\frac{x_s + x_\mu}{x_\mu V}\cdot\frac{p_w^*(\omega_m)}{\omega_m} - i_{rq}\right) \end{cases} \tag{A1}$$

$$\boldsymbol{g} \Rightarrow \begin{cases} p_w^*(\omega_m) = \begin{cases} 0 & if \quad \omega_m < 0.5 \\ 2\omega_m - 1 & if \quad 0.5 \le \omega_m \le 1 \\ 1 & if \quad \omega_m > 1 \end{cases} \end{cases} \tag{A2}$$

where:

$\tau_m = \frac{P_w}{\omega_m}$; $\tau_e = x_\mu(i_{rq}i_{sd} - i_{rd}i_{sq})$;

$P_g = v_{sd}i_{sd} + v_{sq}i_{sq} + v_{rd}i_{rd} + v_{rq}i_{rq}$; $Q_g = -\frac{x_\mu V i_{rd}}{x_s + x_\mu} - \frac{V^2}{x_\mu}$;

$i_{sq} = \frac{r_s}{r_s^2 + (x_s + x_\mu)^2}[\frac{x_s + x_\mu}{r_s}(-x_\mu i_{rq} + v_{sd}) - x_\mu i_{rd} - v_{sq}]$;

$i_{sd} = \frac{(x_s + x_\mu)i_{sq} + x_\mu i_{rq} - v_{sd}}{r_s}$;

$v_{sd} = -Vsin\theta$; $v_{sq} = Vcos\theta$;

$v_{rd} = -r_r i_{rd} + (1 - \omega_m)[(x_s + x_\mu)i_{rq} + x_\mu i_{sq}]$;

$v_{rq} = -r_r i_{rq} - (1 - \omega_m)[(x_s + x_\mu)i_{rd} + x_\mu i_{sd}]$.

References

[1] G. Wang, A.M. Stanković, "Equation-free system-level modeling and analysis of series resonant DC/DC converters," *Electrimacs*, Toulouse, France, Jul. 2017.

[2] G. Wang, "Equation-free system-level modeling and analytics in energy processing systems," *Ph.D. Dissertation*, Dept. of Electrical and Computer Engineering, Tufts University, May 2019.

[3] A.T. Sarić, M.K. Transtrum, A.M. Stanković, "Data classification and parameter identification in power systems by manifold learning," *IEEE PowerTech*, Milano, Italy, Jun. 2019.

[4] A.T. Sarić, M.K. Transtrum, A.M. Stanković, Information geometry for model identification and parameter estimation in renewable energy – DFIG plant case, IET Gen., Transm. Distr. 12 (Mar. (6)) (2018) 1294–1302.

[5] I.G. Holiday, Kevrekidis, Equation-free analysis of a dynamically evolving multi-graph, Eur. Phys. J. Spec. Top. 225 (Sep. (6-7)) (2016) 1281–1292.

[6] M.K. Transtrum, A.T. Sarić, A.M. Stanković, Information geometry approach to verification of dynamic models in power systems, IEEE Trans. Power Syst. 33 (Jan. (1)) (2018) 440–450.

[7] M.K. Transtrum, A.T. Sarić, A.M. Stanković, Information geometry for model verification in energy systems, The 19th Power Systems Computation Conference (PSCC), Genoa, Italy, 2016.

[8] S.L. Brunton, J.N. Kutz, Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control, Cambridge University Press, 2019.

[9] E. Candés, J. Romberg, "ℓ_1-magic: recovery of sparse signals via convex programming," 2005. Online available: https://statweb.stanford.edu/~candes/l1magic/.

[10] H. Arbabi, I. Mezić, Ergodic theory, dynamic mode decomposition and computation of spectral properties of the Koopman operator, SIAM J. Appl. Dyn. Syst. 16 (4) (2017) 2096–2126.

[11] Y. Susuki, I. Mezić, Nonlinear Koopman modes and power system stability assessment without models, IEEE Trans. Power Syst. 29 (Mar. (2)) (2014) 899–907.

[12] J.H. Tu, On dynamic mode decomposition: theory and applications, J. Comp. Dyn. 1 (2) (2014) 391–421.

[13] K.K. Chen, J.H. Tu, C.W. Rowley, Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses, J. Nonlinear Sci. 22 (Dec. (6)) (2012) 887–915.

[14] F. Milano, Power System Modelling and Scripting, Springer, 2010.

[15] M. Grant, S. Boyd, Y. Ye, "CVX: Matlab software for disciplined convex programming," Version 2.1, Dec. 2018. Online available: http://cvxr.com/cvx/.

[16] M. Gavish, D.L. Donoho, The optimal hard threshold for singular value is $4/\sqrt{3}$, IEEE Trans. Information Theory 60 (Aug. (8)) (2014) 5040–5053.