

Measurement-Directed Reduction of Dynamic Models in Power Systems

Mark K. Transtrum, Andrija T. Sarić, and Aleksandar M. Stanković, *Fellow, IEEE*

Abstract—The paper describes a new model reduction procedure tailored to power systems. It uses measurement data to devise a family of reduced order nonlinear models while retaining physical interpretability of parameters and equations. The manifold boundary approximation method (MBAM) uses the Fisher information matrix calculated from measurements to identify the least relevant parameter combination in the original model. Next, it numerically constructs a geodesic on the corresponding statistical manifold originating from the initial parameters in the least relevant parameter direction until a manifold boundary is found. MBAM then identifies a limiting approximation in the mathematical form of the model and removes one parameter combination. The simplified model is recalibrated by fitting its behavior to that of the original model, and the process is repeated as appropriate. MBAM is demonstrated on the example of a synchronous generator (SG), which has been treated extensively in the literature. Implications of the proposed model reduction procedure on large power system models are illustrated on a 441-bus, 72-SG dynamical model.

Index Terms—Computational differential geometry, dynamic equivalent, system identification, information, manifold boundary approximation method.

NOMENCLATURE

Notation (Vectors and matrices are denoted in bold)

e''_d, e'_d	Generator subtransient and transient internal electromotive force in fictitious d -axis, respectively.
e''_q, e'_q	Generator subtransient and transient internal electromotive force in fictitious q -axis, respectively.
f_d, f_q	Parameters in fictitious d - and q -axis, respectively.

Manuscript received February 22, 2016; revised May 18, 2016 and August 4, 2016; accepted September 2, 2016. Date of publication September 20, 2016; date of current version April 17, 2017. This work was supported in part by the ARPA_E under contract DE-AR0000223, in part by the CURENT Engineering Research Center of the National Science Foundation and the Department of Energy under NSF Award Number EEC-1041877, and in part by the Ministry of Education and Science of the Republic of Serbia, under project III-42004. Paper no. TPWRS-00291-2016.

M. K. Transtrum is with the Department of Physics and Astronomy, Brigham Young University, Provo, UT 84602 USA (e-mail: mktranstrum@byu.edu).

A. T. Sarić is with the Department for Power, Electronic, and Communication Engineering, Faculty of Technical Sciences, University of Novi Sad, Novi Sad 21000, Serbia (e-mail: asaric@uns.ac.rs).

A. M. Stanković is with the Department of Electrical Engineering and Computer Science, Tufts University, Medford, MA 02155 USA (e-mail: astankov@ece.tufts.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TPWRS.2016.2611511

f	n_x -dimensional set of differential equations.
g	n_z -dimensional set of algebraic equations.
h	n_m -dimensional set of equations for measured variables.
i_d, i_q	Currents flowing in the fictitious d - and q -axis armature coils, respectively.
I	Fisher Information Matrix (FIM).
J, H	Jacobian and Hessian matrices, respectively.
P_g, Q_g	Generator real and reactive powers, respectively.
P_{m0}, P_m	Referent mechanical power and mechanical power inputs, respectively.
T''_{d0}, T'_{d0}	Subtransient and transient time constants in fictitious d -axis, respectively.
T''_{q0}, T'_{q0}	Subtransient and transient time constants in fictitious q -axis, respectively.
t	Time variable.
V, θ	Bus voltage magnitude and angle, respectively.
v_{f0}, v_f	Referent field voltage and field voltage inputs, respectively.
v_d, v_q	Voltages across the fictitious d - and q -axis armature coils, respectively.
$\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}$	Vectors of state, algebraic, parameter and input variables, respectively.
x''_d, x'_d, x_d	Subtransient, transient and steady-state reactances in fictitious d -axis, respectively.
x''_q, x'_q, x_q	Subtransient, transient and steady-state reactances in fictitious q -axis, respectively.
\mathbf{y}	Set of equations for measured variables (system measurement vector).
\underline{Z}_f	Fault impedance.
τ	Space variable (proportional to the arc length of the geodesic curve).
τ_m, τ_e	Net mechanical and counteracting electromagnetic torques, respectively.
ω, δ	Generator speed and generator rotor angle, respectively.
ω_m	Rotor shaft mechanical velocity.
ω_s	Synchronous speed.
$\lambda = \sigma \pm j(2\pi f)$	Eigenvalue (λ), damping (σ) and frequency (f).
ξ	Damping ratio.
Γ	Christoffel symbol.
\mathcal{P}	$\Omega_b = 2\pi f_b = 2\pi \cdot 60 = 376.9911$ Probability distribution.

Abbreviations

DAE	Differential-Algebraic Equation.
FIM	Fisher Information Matrix.
IVP	Initial Value Problem.
MBAM	Manifold Boundary Approximation Method.
POD	Proper Orthogonal Decomposition.
QoI	Quantity of Interest.
REI	Radial, Equivalent and Independent.
SG	Synchronous Generator.

I. INTRODUCTION

NEW sensing and communication capabilities are key enablers in emerging energy systems. Their promise is to compensate for contemporary trends that make the power system operation and control more challenging:

- 1) Market-driven operation leads to new flow patterns, unfamiliar to the operators.
- 2) Renewable energy sources, such as solar and wind, lead to new (and volatile) flow patterns.

In the multifaceted enterprise of power system operation and control, heterogeneous models have to match different goals and uses. In a large system with many owners and operators, models are a key means of exchanging information among various parties. They represent evolving summaries of accumulated knowledge about the system and starting points for future explorations. This is perhaps best exemplified in the case of fundamental components like synchronous generators (SGs), whose models are used in design, operation, protection, control, and fault accommodation. A particularly interesting analytical feature in power systems is that discrete structures, such as graphs, are strongly blended with continuous (and nonlinear) dynamics, resulting in network dynamics [1].

The need to manage model complexity within a physics-derived class has been present in power systems for a long time. First notable results of this type were related to structure preserving models in transient stability, where explicit inclusion of the network (and load) models resulted in both analytical advances and industrial relevance [2]. Later, work on model derivation via singular perturbation theory allowed for an extension of the notion of multiple time scales to nonlinear models of practical importance in power systems [3]. The need to further customize focus of the models has been explored in the frequency domain within the selective modal analysis, where localizability (inter- and intra-area) results have been derived for linearized models [4]. It is typically the case that the modeler has a strong preference for a class of models (often physics-derived), and the task is then to select a *gray-box* model from this class that has minimal complexity (as measured say by the number of states or parameters), or has some other desirable feature, such as ease of integration with models of other components (e.g., physical network systems [5]). An important trait of model reduction in large power systems is that it is not just a one-time interaction to solve a single abstracted computational problem [6]. Rather, it is a repeated engagement to refine and consolidate one's understanding of the model and the system.

Standard model reduction procedures from control theory are typically based on projection ideas (balanced truncation, Proper Orthogonal Decomposition (POD), Krylov methods for moment

matching of transfer matrix) and have had limited effectiveness in power systems in their original form. This is due to their disregard for the gray-box structure and for the need to return to the large model. We are focusing on model reduction procedures that can act on both component and system level. Perhaps the best known method of this type is based on singular perturbation theory, and has served as a foundation for a systematic derivation of a family of SG models [3]. This can be contrasted with predominantly system-level model reduction that keeps component models unchanged, as exemplified by transient stability studies based on concepts like coherency [7]–[11], synchrony [12] and modal analysis [13].

Attempts to use on-line data to improve dynamical models of key components, especially SGs, are a recurrent theme in power systems. The key difficulty stems from the large number of parameters in the models of interest and from the fact that observed signals in any one study are typically not sufficiently rich to reliably estimate all parameters. First efforts to employ general dynamical systems concepts like trajectory sensitivity go back more than a quarter century [14], [15]; that particular approach has been extended to hybrid systems in [16] with considerable success. Another influential approach that is based on local information extracted from the measurement Jacobian is described in [17]. To deal with ill-conditioning of the parameter estimation problem, the reference proposes in that a subset of parameters of the generator model is fixed to prior values, while estimating the remaining parameters from the available data.

From the large set of power system examples, we comment on studies with direct relevance to our approach. Ref. [18] uses a simple, but well matched predetermined model to successfully describe oscillations between groups of generators. Ref. [19] aims to identify models for adaptive protection, and uses another carefully selected simple model. We are interested in gray-box models whose very structure depends on the measured data. Ref. [20] advances the Krylov subspace linear model reduction, and explores the applications to a very large system of over 12,000 buses. Ref. [21] considers interesting and effective alternatives to coherence-based approaches in linear model reduction, while [22] establishes links between synchrony and controllability/observability. Ref. [23] has pioneered the use of Krylov techniques, and has generated a copious progeny.

Our procedure addresses nonlinear differential-algebraic (DAE) models and is aimed at applications in large systems. The outline of the paper is as follows: Section II describes the theoretical foundation of physics-based power system model reduction; Section III describes the Manifold Boundary Approximation Method (MBAM) for model and its information geometric foundations; Section IV details system-level concerns in SG model simplification; Section V presents the physics-guided power system model reduction; Section VI shows results obtained in a 441-bus benchmark example, and Section VII lists our recommendations and conclusions.

II. PHYSICS-BASED POWER SYSTEM MODEL REDUCTION

In the power systems literature there exists a division between the model- (or physics-) based approaches and the data-driven (or measurement-based) procedures. Each thrust has

its advantages—engineering interpretability and portability to future studies for physics-based methods versus customizability and scalability for the measurement-based ideas. This paper aims to explore the middle stratum and to establish contact points between the two classes of methods. Specifically, we re-evaluate (both locally and globally) physics-based models (with the SG as our key example) in light of their identifiability from typical measurements. On the next step, we propose a systematic, recursive procedure that removes one individual parameter which is hardest to identify from available measurements (and thus reduces and reparametrizes the original model). The reduced model is evaluated next in terms of matching the signals of interest (Quantities of Interest—QoIs), and, if satisfactory, a new single parameter model reduction is attempted. Our aim is to avoid “black boxes” whenever possible, as such models do not explain connections to the underlying physics that guide engineering, nor provide explanatory insights for future studies.

There are several challenges to systematic parameter reduction. First, it is often parameter combinations rather than individual parameters to which the model is insensitive. Typically, a model is sensitive to all parameters individually and the insensitivity only arises because of their compensatory nature. The relevant parameter combination typically is a nonlinear function of individual parameters. Consequently, it depends strongly on the parameter values, which are in turn very sensitive to noisy data.

To overcome these challenges, we use an information-theoretic approach. Consider an arbitrary probability distribution $\mathcal{P}(\mathbf{y} | \mathbf{p})$ for observing a vector of random variables \mathbf{y} given a parameter vector \mathbf{p} . The range of physically allowed parameter values describes a family of models. We seek a reduced model with fewer parameters that can approximate the full family described by the original. We anticipate such a reduction to exist for models with more parameters than “effective degrees of freedom” in their observations. Qualitatively, such original models have more complexity in their description than in their predictions, which leads to large uncertainties in inferred parameter values.

The sensitivity of model predictions to changes in parameters is measured locally by an eigenvalue decomposition of the Fisher Information Matrix (FIM), $I_{\mu\nu} = -\langle \frac{\partial^2 \log \mathcal{P}}{\partial p_\mu \partial p_\nu} \rangle$, where $\langle \cdot \rangle$ denotes the expectation. Often models (known as sloppy models) have eigenvalues exponentially distributed over many orders of magnitude, quantifying an extreme insensitivity to coordinated changes in the original parameters [24]. The family described above is equivalent to a manifold of potential models, known as the model manifold, with metric tensor given by the FIM. When a model has many more parameters than effective degrees of freedom in its emergent behavior, the manifold is bounded with a hierarchy of widths, qualitatively described as a hyper-ribbon [25], [26]. Indeed, the widths of this hyper-ribbon are a measure of the number of effective degrees of freedom in the model. If the narrowest widths are sufficiently small, then it can be accurately approximated by a low-dimensional, reduced model, analogous to approximating a long, narrow ribbon by either a two-dimensional surface, or a one dimensional curve (see Fig. 1). Note that the model manifold is a statistical mani-

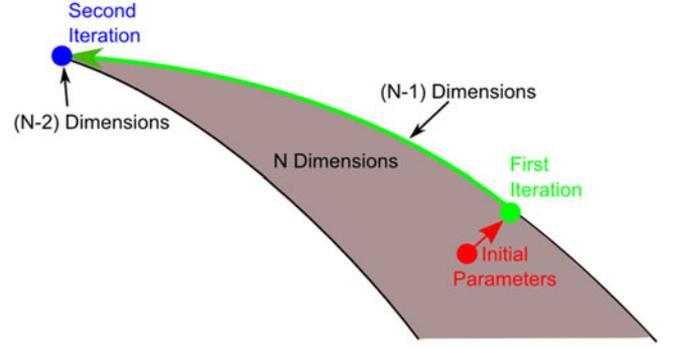


Fig. 1. Approximating the manifold by its boundary [24]. A high dimensional, bounded manifold may be approximated by a low-dimensional manifold. Parametric degrees of freedom are systematically removed, one at a time, by approximating the full manifold by its boundary. After several approximations, the reduced model is represented by a hyper-corner of the original manifold that preserves most of the original model’s behavior.

fold, i.e., each point on the manifold is a probability distribution for an observation vector. This is distinct from the slow-manifold familiar in dynamical systems theory. The slow-manifold aims to approximate the system with lower dynamical order. Here, we seek to approximate a system with fewer parameters. Although distinct, these concepts are not unrelated as we discuss at length below.

III. INFORMATION GEOMETRY FOR COMPONENT MODEL REDUCTION

Models of power systems are typically written in DAEs form:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t); \quad (1)$$

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

where \mathbf{x} is the vector of (differential) state variables, \mathbf{z} are the algebraic variables, \mathbf{p} are parameters, \mathbf{u} are inputs (typically assumed to be known in estimation studies) and t is the (scalar) time variable.

System measurement vector is assumed to be of the form:

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

It turns out that the parametric sensitivities have dynamics described by the following equations (see e.g. [27]):

$$\frac{d}{dt} \frac{\partial \mathbf{x}}{\partial \mathbf{p}} = \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{z}} \cdot \frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{p}}; \quad (4)$$

$$\mathbf{0} = \frac{\partial \mathbf{g}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial \mathbf{g}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{z}} \cdot \frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{g}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{p}}; \quad (5)$$

$$\frac{\partial \mathbf{h}}{\partial \mathbf{p}} = \frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{z}} \cdot \frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{h}(\mathbf{x}, \mathbf{z}, \mathbf{p}, \mathbf{u}, t)}{\partial \mathbf{p}}. \quad (6)$$

These equations are linear in terms of sensitivities, but the matrices involved do vary along a system trajectory. Details about DAEs modeling of SG are provided in *Appendix*. Note that (4)–(6) are derived by differentiating (1)–(3) with respect to the parameters and applying the chain rule to account for the implicit dependence of the dynamic, algebraic, and observed variables on the parameters. We also make use of the second order sensitivities below, the equations for which can be derived in like manner. The problem of calculating parametric sensitivities for dynamical systems is well-known, with a long history [28]. However, deriving expressions for the first and second order sensitivities by hand can be tedious and error prone (particularly for large models). We use automatic differentiation [29], [30] to simplify the process.

In this section, the goal is to construct a physically meaningful representation that reveals the simple ‘theory’ that is hidden in the model. *The Manifold Boundary Approximation Method (MBAM)* [24] is an approach to model approximation whose basic idea is to approximate a high-dimensional, but thin model manifold by its boundary. The boundary is identified by numerically computing a geodesic on the model manifold and using the information from this calculation to identify an approximate model with fewer parameters.

Central to the MBAM procedure is computational differential geometry in the form of numerical geodesics. *Geodesics are the analogs of straight lines generalized to curved surfaces*, and are calculated numerically as the solution to a second order ordinary differential equation in parameter space (while utilizing quantities from the data space):

$$\frac{\partial^2 p^i}{\partial \tau^2} = \sum_{j,k} \Gamma_{jk}^i \frac{\partial p^j}{\partial \tau} \cdot \frac{\partial p^k}{\partial \tau}; \quad \Gamma_{jk}^i = \sum_{\ell,m} (\mathbf{I}^{-1})^{i\ell} \frac{\partial y_m}{\partial p_\ell} \cdot \frac{\partial^2 y_m}{\partial p_j \partial p_k}, \quad (7)$$

where Γ are the so-called Christoffel symbols [31], containing curvature information about the mapping between parameter space and data space, which are expressed in terms of the first and second order parametric sensitivities (4)–(6) and \mathbf{I} is the FIM (introduced in Section II). The parameter τ parameterizes the geodesic and is proportional to the arc length of the geodesic as measured on the model manifold, i.e., in data space. Solving (7) gives a parameterized curve $\mathbf{p}(\tau)$ in parameter space that is used to reveal a limiting behavior in the model as we demonstrate below.

Equation (7) is an ordinary differential equation that we solve as an Initial Value Problem (IVP). Here, we take the model’s “true” parameter values as the starting point of the geodesic. The initial “velocity” is taken to be the least sensitive parameter direction as measured by the FIM, as we describe below. As an aside, solving multiple geodesics (using different starting points and directions) is an efficient and informative way of globally characterizing the parameter space.

The MBAM procedure can be summarized as a five step algorithm. Here, we describe the algorithm and give a summary in Fig. 2. We give relevant equations and technical details below, followed by an illustrative example.

Step a: The least sensitive parameter combination is identified from an eigenvalue decomposition of the FIM

Given:

1. DAEs model, eqs. (1), (2)
2. “True” parameter values

Step a:

1. Solve eqs. (1)–(6)
2. Calculate FIM, $\mathbf{I} = \mathbf{J}_p^T \mathbf{J}_p$
3. Find initial geodesic velocity, $\partial \mathbf{p} / \partial \tau$ (eigenvector of \mathbf{I})

Step b:

1. Solve eq. (7) to find $\mathbf{p}(\tau)$
2. Identify limit of geodesic

Step c:

- Evaluate limit to find new DAEs model, i.e. new eqs. (1), (2)

Step d:

- Fit the approximate model to the original by least-squares to find the new “true” parameter values

Step e:

Iterate Steps a-d

Fig. 2. Flow-chart summarizing the steps of the MBAM algorithm as described in the text.

which becomes the initial geodesic velocity, $\partial \mathbf{p} / \partial \tau$. The geodesic acceleration $\partial^2 \mathbf{p} / \partial \tau^2$ is given by (7). If $\sum_i (\partial p^i / \partial \tau) (\partial^2 p^i / \partial \tau^2) < 0$, then reverse the initial velocity: $\partial \mathbf{p} / \partial \tau \rightarrow -\partial \mathbf{p} / \partial \tau$. This heuristic resolves the ambiguous direction associated with the eigenvalue.

Step b: A geodesic on the model manifold is constructed by numerically solving (7) using the “true” parameter values and the velocity $\partial \mathbf{p} / \partial \tau$ calculated in **Step a** as initial conditions. The limiting behavior of this curve identifies the boundary of the model manifold, as we illustrate with examples below.

Step c: Having found the edge of the model manifold, the corresponding model is identified as an approximation to the original model. By inspecting which values of the parameter vector become infinite, we identify the boundary as a limiting approximation in the model. We evaluate this limit to construct the approximate model with a reduced parameter count.

Step d: The parameter values for this approximate model are calibrated by fitting the approximate model to the behavior of the original model by least squares regression.

Step e: The procedure is repeated until the reduced model is unable to approximate quantities of interest.

The MBAM process is described in detail and several examples are given in references [24]–[26]. Here, we briefly illustrate with a simple power systems example. Consider a model of a SG with two parameters that makes predictions for generator rotor angle, generator speed, real and reactive powers at several times after some disturbance. (In order to focus on the MBAM procedure, we postpone specific details of this simulation, such as the nature of the disturbance, until the next section.) We take as parameters the transient time-constants in the d - and q -axes.

The process begins by numerically solving both the model equations (1)–(3) and the sensitivities (4)–(6) to construct

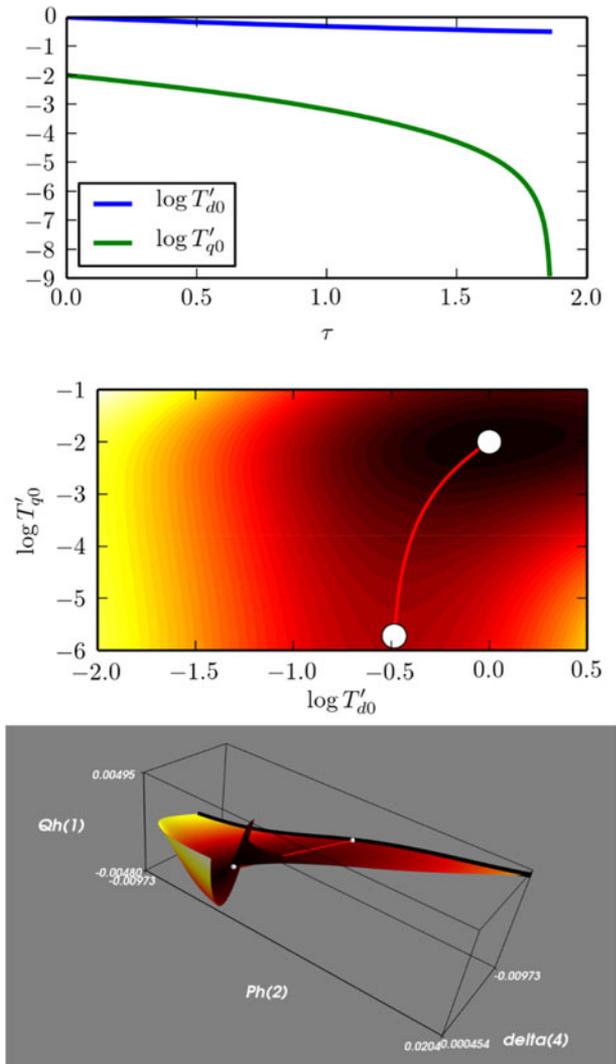


Fig. 3. Solving the geodesic equation (7) generates a sequence of parameter values as a function of τ (top panel) that parameterizes a curve through parameter space (middle panel). This curve corresponds to a path on a model manifold (bottom panel) that is approximately a straight line. This curve encounters a singularity near $\tau = 1.8$ that corresponds to the boundary of the model manifold (black line in bottom panel).

the Jacobian matrix, $\mathbf{J}_p(t) = \partial \mathbf{h}(t) / \partial \mathbf{p}$ and the FIM matrix, $\mathbf{I} = \mathbf{J}_p^T \mathbf{J}_p$ (Step a).

Next, we numerically solve (7) (Step b); a solution requires an initial condition and initial derivative. We take these to be the “true” parameter values and the eigenvector of the FIM with smallest eigenvector, respectively, found in Step a. The procedure generates a sequence of parameter values as a function of τ (Fig. 3, top panel), that parameterize a curve through parameter space (Fig. 3, middle panel). Colors in Fig. 3, middle panel represent contours of least squares cost measuring deviation of the model behavior from the true parameter values. Notice how the geodesic naturally curves through parameter space to construct the path of smallest least squares cost. This path can similarly be interpreted as a path on a manifold (Fig. 3, bottom panel). In this example, the model manifold is a two-dimensional surface (corresponding to the 2 parameters of the model) embedded in

a 1680 dimensional space (1680 = four measurements at 420 time points each). In order to visualize this manifold, we have (somewhat arbitrarily) chosen three axes corresponding to three measurements (reactive power at $t = 1$, real power at $t = 2$, and rotor angle at $t = 4$). These measurements were selected in order to make features of the manifold visually clear. Colors on the model manifold match the corresponding parameter values in the middle panel.

Notice that the geodesic is approximately a straight line through the data space (bottom panel in Fig. 3), where the length of this curve is proportional to τ . The model manifold has a boundary (black curve, bottom panel). Notice that the geodesic equation exhibits a singularity at a finite value of τ (near 1.8) that corresponds to the geodesic encountering this boundary. Inspecting the top and middle panels in Fig. 3, we infer that this boundary corresponds to the limiting approximation $T'_{q0} = 0$. Evaluating this limit in the model (Step c) is equivalent to a singular perturbation analysis, and removes one dynamical variable along with the parameter.

Having found the functional form of the reduced model, parameter values are determined by fitting the behavior of the approximate model to that of the original model by least squares (Step d). For concreteness, let \mathbf{p}' denote the parameters in the reduced model, so that \mathbf{y}_p and $\mathbf{y}_{p'}$ are the predictions of the original and approximate models respectively, then we select \mathbf{p}' to minimize $\sum (\mathbf{y}_{p'} - \mathbf{y}_p)^2$. This step both determines the new parameter values while answering the larger question: “How accurate is the reduced model?” In practice, we iterate MBAM (Step e) as long as the sum of squares error between true and approximate models is “acceptably small”. What constitutes “acceptably small” depends on the context and application, but a convenient stopping criterion is when the error in the approximation is comparable to the experimental noise in measurements.

In general, having models with many parameters prevents visualizations, such as in Fig. 3, but the basic procedure can be executed to identify a sequence of simplifying limits. These limits need not always correspond to singular perturbations, however. Indeed, the limits and the process for evaluating them will vary with each iteration of the process and require some theoretical work (motivated by insights from computational differential geometry) to construct the simplified model.

Notice how the model provides the connection between the parameter space and data space through the Jacobian matrix $\mathbf{J}_p(t) = \partial \mathbf{h}(t) / \partial \mathbf{p}$ as calculated in (4)–(6). Further note that the Christoffel symbols $[\Gamma$ in (7)] involve the second order sensitivities that are found by taking another derivative in (4)–(6). We omit an explicit formula for these sensitivities as the derivation is straightforward and the result is lengthy and not illuminating. Because we evaluate these sensitivities using automatic differentiation [29], [30], these expressions are not explicitly needed. There is a technical subtlety in the evaluation of (7) that is critical for our approach to be tractable for large models. Because the second derivative of the observation vector is contracted twice with the geodesic velocity vector (i.e., the sums over indices j and k in (7) form two “dot products” with the geodesic velocities and the array of second derivatives), only a

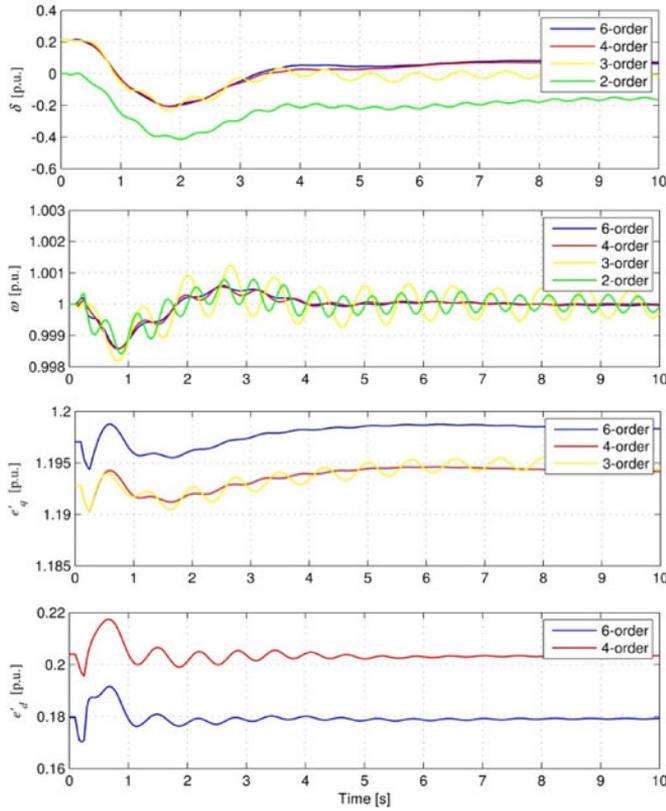


Fig. 4. Time responses of state variables for SG in bus 2 for different model orders.

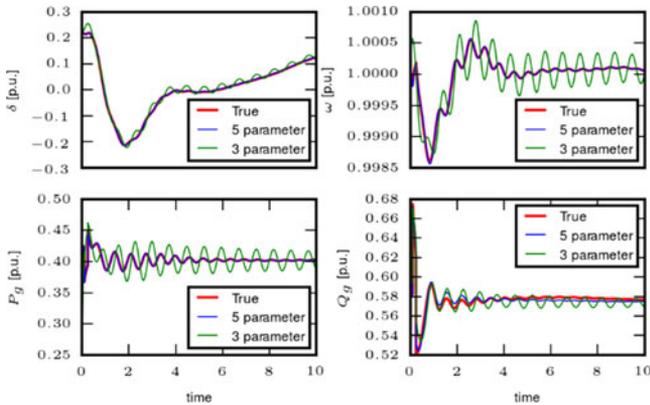


Fig. 5. Measurements compared with the 3- and 5-parameter models.

directional second derivative is needed, which can be calculated efficiently as in [24]–[26].

Solutions to the geodesic equation (7) are calculated using standard methods for numerically integrating IVPs, as we have done to generate Figs. 3 and 6 later [24]–[26]. The geodesics tend to be highly nonlinear (reflecting the incompleteness of the local analysis), but align with the local curvature of the cost surface. In this way, the geodesics systematically explore the non-local structure of the parameter space.

By constructing several orthogonal geodesic paths, one can identify different cross sections of the model manifold and use the geodesic distance (τ) to measure the width of the manifold. For many models, including those considered here, it is empirically observed that the entire manifold is bounded, and often

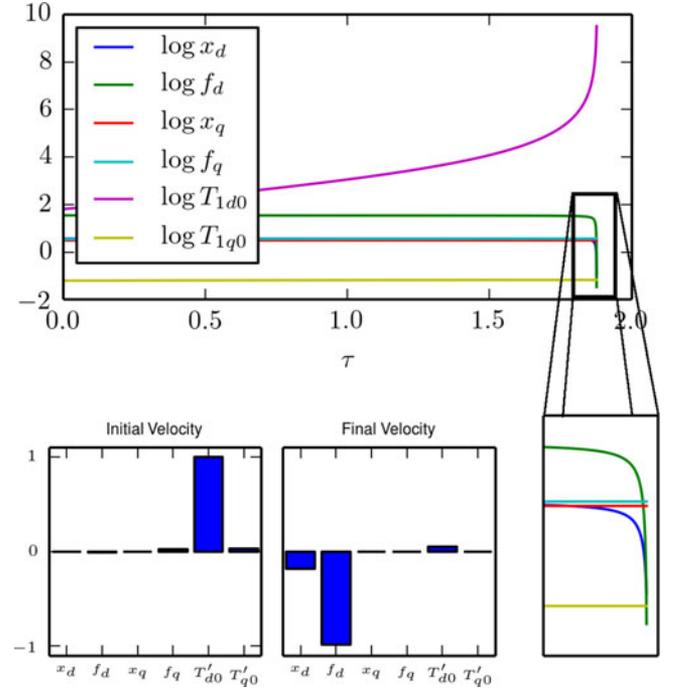


Fig. 6. Top: Parameter values along geodesic path for the 6-parameter model; the fifth parameter (T'_{d0}) initially becomes large, the geodesic then rotates to give $f_d = 0$ at the manifold boundary.

Bottom: Components of the initial and final geodesic velocities; the initial geodesic velocity dominated by T'_{d0} rotates to reveal the limit $f_d = 0$.

highly anisotropic with widths typically forming an exponential hierarchy, reminiscent of the hierarchy of FIM eigenvalues revealed by the local analysis [25], [26].

The geodesics give a global characterization of the model behaviors. When geodesic curves extend parameter values to zero or infinity in a finite distance on the model manifold, i.e., finite value of τ , the corresponding parameter combination is susceptible to identifiability problems. These limiting cases correspond to boundaries of the model manifold that are structurally simpler models. Furthermore, these reduced models are a feature of the model manifold that are invariant to certain changes in the observation vector that we discuss later, giving a *topological* (i.e., global) description of the manifold [32].

The result of the 5-step procedure removes the single least-identifiable parameter combination from the model. Iterating MBAM therefore produces a series of models of decreasing complexity that explicitly connect the microscopic components to the macroscopic behavior. These models correspond to hypercorners of the original model manifold.

MBAM is a very general approximation scheme, making no assumptions about the mathematical form of the model, or underlying physics of the system. It requires only that the model manifolds have a hierarchy of boundaries. The existence of these edges was first noted in the context of data fitting [33] and Markov-Chain Monte Carlo sampling of Bayesian posterior distributions [24]–[26]. It was noted that algorithms tend to ‘evaporate’ parameters, i.e., allow them to drift to extreme, usually infinite, values. These parameter values correspond to limiting behaviors in the model, i.e., manifold boundaries, which we use as reduced models.

IV. MODEL REDUCTION FOR SYNCHRONOUS GENERATOR (SG)

We have developed a Matlab-derived and PSAT-based simulation environment built around the IEEE 14-bus test system (see [34, Fig. 2.4 and *Appendix D*] for detailed input data). The test system is modified to include direct-drive SG (used by industry to model solar plants and a new generation of wind) in bus 6, doubly-fed induction generator (capturing prevalent type of wind plants today) in bus 8 and SGs in buses 1, 2 and 3 (describing conventional units and interconnections).

We have added our code for evaluation of measurement sensitivities and for computational differential geometry (in Julia). Our Matlab code is fully general in the sense that it allows for a variety of measurements (generator rotor angle and generator speed, nodal active and reactive power injections, nodal voltage magnitude and angle, branch active and reactive flows, and branch current magnitude). The right-hand sides of the sensitivity equations are found using Julia’s “DualNumbers” package for forward automatic differentiation [35]. The differential equations for both the model (1)–(6) and the geodesic (7) are solved using the legacy FORTRAN solver VODE [36].

MBAM takes a parameter-centered approach in which there are the following SG’s electrical parameters in the model (the models are named after the number of state variables):

6-order (A1a-d): $x'_d, x'_q, x_d, x_q, T'_{d0}, T'_{q0}, T''_{d0}$, and T''_{q0} .

4-order (A1a-c): $x'_d, x'_q, x_d, x_q, T'_{d0}$, and T'_{q0} .

3-order (A1a-b): x'_d, x'_q, x_d, x_q , and T'_{d0} .

2-order (A1a): x'_d, x'_q, x_d , and x_q .

We take as measured QoI for our model to be the generator rotor angle (δ), generator speed (ω), real (P_g) and reactive powers (Q_g), and seek an approximate model that preserves the predictions of the original parameter model but with fewer parameters. Our analyses are performed on SG connected to bus 2 in IEEE 14-bus test example [27], [34]. For simplicity, we assume that interface variables for SG in bus 2 (P_{m2} , v_{f2} , V_2 and θ_2) are known functions of time, but independent of the parameters considered. This, of course, is an approximation for a multi-generator system, but it allows direct comparison with numerous references that focus on a single SG. This example of single-machine multi-bus system can be compared with examples in [3], [38], [39].

In Fig. 4 time responses of state variables for SG in bus 2 are shown for different model orders (the interface variables are held constant). The responses were obtained for three-phase short circuit in bus 4 in $t = 0.1$ s, cleared after 150 ms.

Depending on the model order (from 2- to 6-order) the following differential equations are used, respectively: (A1a) + (A1b) + (A1c) + (A1d). Similarly, the following algebraic equations are used: (A2a)+(A2b') [or (A2b''), or (A2b''')]. Depending on the model order, the initial conditions (and transients) are different.

A. 4-Order SG Model

We start with the 4-order (two-axis, d - q , two electrical states, 6-parameter) model, as it is very widely used in power system analyses. It is also the model considered in the seminal reference on singular perturbation-based SG model derivation and

reduction [3]. We consider a more detailed 6-th order model (with two damper windings) later. Throughout the model reduction we enforce the constraints arising from the physical nature of parameters, namely $x_q > x'_q$ and $x_d > x'_d$ by introducing the parameters f_q and f_d defined by the relation $x'_q = x_q/(1 + f_q)$ and similarly for f_d . We use f_q and f_d (which are positive) as parameters in place of x'_q and x'_d .¹ The QoI during the transient of interest are shown in Fig. 5.

Iteration 1: We first calculate the FIM and consider its eigenvalues and eigenvectors. The eigenvector with smallest eigenvalue is aligned primarily with the fifth parameter, which is T'_{d0} (see Fig. 6). This direction becomes the initial parameter space velocity of the geodesic.

The geodesic initially moves toward the limit in which T'_{d0} becomes infinite. Inspecting the equations, however, reveals that in this limit the variable e'_q becomes constant and the parameter combination $x_d - x'_d$ no longer appears in the model. Therefore, as T'_{d0} becomes very large, the parameter combination $x_d - x'_d$ becomes the least identifiable parameter combination and the geodesic appropriately rotates to give the limit in which f_d becomes zero, so that $x_d = x'_d$ at the manifold boundary.

Please note that the computational effort is reasonable—the geodesic shown in Fig. 6 took 182.5 sec to calculate and all geodesics together took 329.43 sec². Also, the calculation scales favorably with system size and can be efficiently parallelized on large models.

In the next step (**Step d**), the remaining parameters are retuned, and the only notable differences are the 14% reduction of x_d (effectively taking the previous value of x'_d) and a very large increase in the value for T'_{d0} . The fit of this model is very good (as seen on Fig. 5), with notable discrepancies only in the reactive power measurement.

Iteration 2: On the second iteration of MBAM, the geodesic reveals the manifold boundary to be T'_{d0} becoming infinite. The resulting 4-parameter (x_d, x_q, x'_q , and T'_{q0}) model has 3 differential equations (δ, ω and e'_d), and the parameter retuning with the largest change coming from parameter x_d , which essentially takes on the previous value of x'_d . The fit for this model is practically indistinguishable from the fit shown in Fig. 5.

Iteration 3: On this iteration of MBAM, the geodesic reveals the manifold boundary to be $x_q \rightarrow x'_q$. At this point, the differential equation for e'_d becomes autonomous, with the right-hand side equal to e'_d/T'_{q0} . The parameter recalibration results in largest changes in T'_{q0} (nearly 98% reduction). The fit for this model is, however, not nearly as good, as can be seen in Fig. 5.

Iteration 4: The geodesic identifies the next limit as $T'_{q0} \rightarrow 0$, which results in the so-called classical SG model (2-order), with only mechanical states retained. When recalibrated, the remaining two parameters (x_d and x_q) do not change much

¹The positivity of f_d and f_q is a way to enforce the physical constraints $x_q > x'_q$ and $x_d > x'_d$. Without this contrivance, the procedure may veer into non-physical parts of the parameter space. Equations with f_d and f_q appear even more nonlinear, but are better suited to our method. The model manifold and geodesics are invariant to re-parameterization such as this.

²Intel(R) Core(TM) i7-3520M CPU @2.90GHz, 64-bit Operating System, 16 GB RAM.

TABLE I
PARTICIPATION FACTOR MATRIX

Eigenvalue	δ	ω	e'_q	e'_d
λ_1	0.04	0.04	0.92	0.00
$\lambda_{2,3}$	0.48	0.48	0.04	0.00
λ_4	0.00	0.00	0.00	1.00

from the previous iteration, and the match is similar to the one in Fig. 5, with even more pronounced oscillations.

B. Discussion of the Two-Axis Model Reduction

Our results suggest that, for the available measurements, the four parameter model (obtained after *Iteration 2*) offers a good tradeoff between model complexity and response fidelity. However, this is not the “single-axis” model obtained via singular perturbation [3] which keeps the dynamics of the slow electrical variable e'_q , while e'_d becomes an algebraic quantity. This discrepancy has several contributing factors:

- 1) Our transients are 10 sec long, but most of the transient is over in 7–8 sec. Note that the “true” value of T'_{d0} (used in Fig. 4) is 6.1 sec, thus quite close to the effective duration of the transient. Thus, e'_q does not change appreciably, and approximating it as constant is not unreasonable.
- 2) This reasoning can be made quantitative locally using modal analysis and participation factors. The eigenvalues do not vary much during the transient, and approximately are at $\lambda_1 = -0.37$, $\lambda_{2,3} = -0.73 \pm j9.46$ and $\lambda_4 = -3.58$, with the participation factors matrix shown in Table I. This table clearly identifies the state e'_q as the dominant contributor to the slowest eigenvalue λ_1 . At the same time, the sensitivity of λ_1 to parameter T'_{d0} is 0.066, which is about 200 smaller than the sensitivity of λ_4 to “its” time constant T'_{d0} that equals 12.966.
- 3) Another view is that the MBAM results can be obtained from the analysis of the measurement Hessian [17], [27]. Starting with the Jacobian matrix, $\mathbf{J}_p(t) = \partial \mathbf{h}(t) / \partial \mathbf{p}$, first partial derivatives of system measurement vector (3) with respect to the parameter vector (\mathbf{p}), and considering the Hessian matrix, $\mathbf{H}_p(t) = \mathbf{J}_p^T(t) \mathbf{J}_p(t)$ (for small increments), we calculate the eigenvalues (for the whole transient). We find a wide spread of the eigenvalues—the condition number is 347430.1, and the smallest eigenvalue is 8.36. Parameter T'_{d0} has the highest participation in this eigenvalue (0.41). This validates the conclusion that T'_{d0} is hard to estimate from the transient considered. Interestingly, the second parameter with second highest participation in small eigenvalues of the Hessian is x_d , which is evaporated in the first iteration. Notice that the local sensitivity analysis incorrectly ranks the order of importance for the parameters T'_{d0} and x_d , a problem remedied by the non-local geodesic analysis.
- 4) The composition of the measurement set (QoI) also influences the MBAM procedure. A closer examination of the

two-axis model reveals that P_g , δ and ω are all strongly affected by e'_d , while only Q_g is affected by e'_q . Indeed, if only the reactive power (Q_g) is used as QoI, then our MBAM procedure evaporates T'_{q0} first, and essentially repeats the steps of a model reduction based on singular perturbation [3]. For all other choices of QoI, the model reduction retraces the steps outlined previously.

C. 6- and 3-Order SG Models

In the case of a 6-order SG model (one damper winding in each of d - and q -axis), the MBAM procedure starts by evaporating the time constants corresponding to damper windings. This reduces the model to the 4-order and is consistent with singular perturbation. Later, the process follows steps described earlier in this section.

When starting with a 3-order model (δ , ω , and e'_q —see Fig. 4), MBAM first evaporates the time constant T'_{d0} corresponding to the electrical equation, and reduces the model to the classical one [only mechanical differential equations (A1a)]; this is also consistent with singular perturbation.

D. Sensitivity to Types of Transients

We have considered the case when SG models in one node are varied, while the rest of the system remains unchanged. To explore how our results depend on specific details, we have repeated our calculation for a variety of conditions. Specifically, we repeated our calculations from several choices of “true” parameter values and for different transient dynamics (generated by different locations of short circuit, locations of SG for reduction, power system’s loading levels, and contingencies), always observing generator rotor angle, generator speed, and real and reactive powers. This is a fairly standard set for identification studies; in practice, the generator rotor angle is estimated from other local measurements, such as the terminal current and voltage. In all cases, MBAM has resulted in the same sequence of parameter evaporations and consequent model reduction, so we do not show the detailed results for these different transients.

The general dependence of MBAM on the choice of QoIs as has been explored theoretically in [32]. Changing the QoIs will change the FIM and by extension the geometric properties of the manifold. In many cases, such as changing the measurement times or input functions, these changes effectively “stretch” or “compress” portions of the manifold, i.e., transform the model manifold in a differentiable way—transformation, known as *diffeomorphisms*. The manifold boundaries are invariant to diffeomorphisms, because boundaries are characterized by singularities in the FIM. In other words, the boundaries are a feature of the differential topology of the family of manifolds generated by varying the QoIs. Therefore, the reduced models identified by MBAM are generally robust to changes in the QoIs, because it is exploiting a topological invariant of the model.

More extreme changes in the QoIs, such as ignoring a dynamical variable altogether can lead to more drastic changes in the model manifold that do affect the boundary structure, a phenomenon described as manifold collapse. Because they change the topological structure of the manifold, variations in

QoIs that lead to manifold collapse can also lead to different reduced MBAM models. These different reduced models allow different effective relationships among the observations.

V. PHYSICS-GUIDED POWER SYSTEM MODEL REDUCTION

Our strategy for extending the MBAM approach to *large systems* combines linearized model of the interconnection with non-linear component models:

Step 1: Determination of key modes to be included in the system model with key states (via participation factors of the system matrix). These tend to be well-understood in interconnections with challenging transients [22].

Step 2: Determination of corresponding physical components and their vital (“systemic”) parameters (via sensitivities of the system matrix to parameters).

Step 3: Measurement structure and MBAM-based model reduction to achieve good matching of QoI, while maintaining the presence of all systemic parameters.

Step 4: Provide guidance for design of additional experiments if needed.

For example, in the IEEE 14-bus test system, the least damped pole pair is at $-0.61 \pm j10.89$ (thus critical for oscillation damping). The sensitivities of these eigenvalues to parameters of SG in bus 2 vary widely, and are largest for x'_q (0.92) and x'_d (0.05); thus it makes sense to declare x'_q as a *systemic parameter*. This parameter is retained in even very low-order models (e.g., iterations 2 or 3 of MBAM), so area or SG identification with any of the proposed reduced models will suffice for the critical pole as far as SG in bus 2 is concerned.

Given the networked nature of power system dynamics, it makes sense to initiate model reduction in nodes, or across cut-sets (using the line flow data—the so-called area model reduction). When the model for a component/subsystem serves as a starting point for MBAM, the issue of its alignment with the system level model arises naturally. For example, in the IEEE 14-bus test system, the linearization of the overall model has 48 states, and we use participation factors to establish links between its eigenvalues and the 4-order model of the SG in bus 2 in the previous section. We shouldn’t expect a very precise agreement, as the system model contains dynamics of exciters, turbines and governors that have been removed from the SG model dynamics (by being treated as functions of time, known in simulations), so that we can focus MBAM on reducing the SG itself. Some modes undergo little, or no change between the two models—for example $\lambda_4 = -3.58$ (dominated by e'_d) is shifted to -3.42 . Other modes are changed more significantly—for example, the node dominated by e'_q (and the exciter) nearly maintains the frequency of λ_2 , but the damping is reduced. While engineering judgement clearly has to play a role in checking the local and system level model consistency, we believe that nodal/cut-set MBAM reduction, conjoined with the determination of systemic parameters, provides a sound foundation for model reduction in large power systems, as illustrated in the next *Section*. The physical interpretability of steps in MBAM reduction is certainly helpful in this regard.

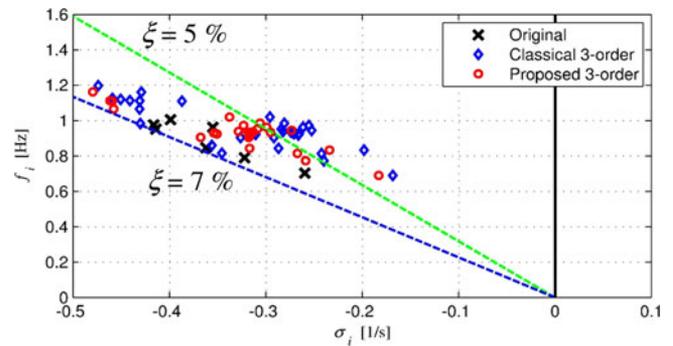


Fig. 7. Eigenvalues plots for different dynamic model reduction options.

VI. APPLICATION TO LARGE POWER SYSTEM MODEL REDUCTION

The main characteristics of the original dynamic model for real-world, large power system (Electric Power Industry of Serbia, a part of the ENTSO-E interconnection) can be summarized as: 441 buses, 655 branches (lines and transformers), 72 SGs (43 of 4-order models and 29 of 6-order models), with exciters and turbines. The model has 850 differential and 1314 algebraic variables.

In this section we show how insights gained from the dynamic reduction analysis may be used to simplify the dynamic model of a real-life power system, while maintaining remarkable fidelity of the response with very little tuning.

A model of a SG (or other dynamic components) is typically reduced and identified from nodal measurements, and then the procedure is repeated as many times as necessary, one node at a time. This aspect makes the method scalable for real-world power systems application. We also envision area- or subnetwork-wise procedure for large systems.

Based on conclusions derived in Sections IV and V, related to the dynamic reduction for sloppy model, we assume that the all classical 3-order (and higher) SG’s models (described in *Appendix*) are replaced with 3-order differential equations (δ , ω and e'_d), one algebraic equation [e'_q ; $T'_{d0} \rightarrow \infty$ in (A1b)] and 4-parameter (x'_d, x'_q, x_d , and x_q)—denoted as “Proposed 3-order” in Fig. 7, where f_i and σ_i denotes i -th eigenvalue frequency and damping, respectively. These results are compared with original and 3-order SG’s dynamic reduction models [(A1a-b) in *Appendix*], denoted in Fig. 7 as “Original” and “Classical 3-order”, respectively. Blue and green lines in Figs. 7 and 8 shows damping ratios (ξ) of 7% and 5%, respectively. Note that the proposed model outperforms the standard one in this example. While more studies are in order before such statements can be generalized, it is encouraging that the proposed model shows practical utility.

The test system is interconnected with neighboring power systems by ten 400 kV and 220 kV lines. Two interconnection lines exports energy and static (load-based) equivalents are connected in these buses. Eight interconnection lines imports energy and in these buses dynamic [SG- and Radial, Equivalent and Independent (REI-) based] equivalents are connected. Large test system is subjected to the three-phase short circuit in bus

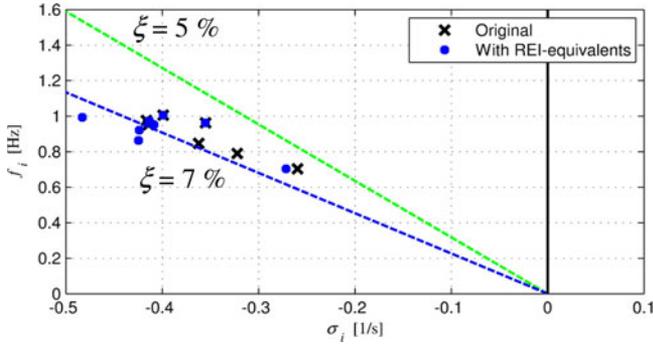


Fig. 8. Eigenvalue plots for test system with neighboring areas (modeled by REI-equivalents).

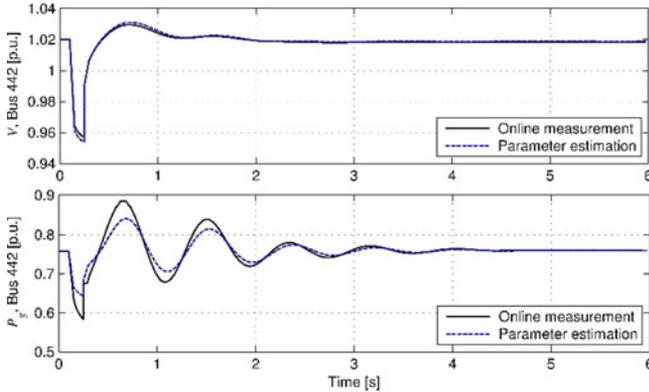


Fig. 9. Time responses in a boundary point (modeled by REI equivalents): bus voltage and SG's active power for original ('Online measurement') and reduced ('Parameter estimation') large power system models.

11 in $t = 0.1s$, which cleared after 250 ms (fault impedance is $Z_f = j0.1$ p.u.).

Based on analyses from Sections IV and V, the electrical dynamic part of REI and minimum-loss equivalents for neighboring power system areas are assumed as proposed 3-order differential equations (δ , ω , and e'_d), and one algebraic equation model [e'_q ; $T'_{d0} \rightarrow \infty$ in (A1b)]. Reactances of the REI-equivalents (x'_d , x'_q , and x_q) are obtained from boundary measurements [37], while the time constant T'_{q0} and x_d are identified as the sloppy parameters, where in simulations are assumed $T'_{q0} = 1$ s and $x_d = 1.83$ p.u.

Rightmost eigenvalue plots are shown in Fig. 8, while in Fig. 9 we show examples of time responses for bus voltage and SG's active power in a bus with REI equivalent.

VII. CONCLUSION

The paper describes a nonlinear model reduction procedure that is well-suited for power system applications. It combines physical and engineering insight with data-driven exploration to produce practically interpretable models. It turns out that the networked system structure is a key to successful application of MBAM. We have studied its properties on a small system and illustrated its potential on a much larger system from industrial practice. The procedure connects with existing methods, such as singular perturbation in interesting ways and offers a new tool to the armamentarium of a power system analyst.

APPENDIX

If for synchronous generator (SG) we assume $\omega \approx \omega_s$, (or $\omega \approx 1$ p.u.), then $\tau_m = P_m/\omega_m \approx P_m$ and $\tau_e = P_g$ common mechanical differential equations (2-order model) are [34]:

$$\mathbf{f} \Rightarrow \begin{cases} \dot{\delta} = \Omega_b(\omega - \omega_s) \\ \dot{\omega} = \frac{1}{2H}(P_m - P_g - D(\omega - \omega_s)) \end{cases}. \quad (\text{A1a})$$

For one q -axis, 3-order model to (A1a) the following differential equation is added:

$$\mathbf{f} \Rightarrow \left\{ \dot{e}'_q = \frac{1}{T'_{d0}}(-e'_q - (x_d - x'_d)i_d + v_f) \right.}. \quad (\text{A1b})$$

For one d - and one q -axis, 4-order model to (A1a) and (A1b) the following differential equation is added:

$$\mathbf{f} \Rightarrow \left\{ \dot{e}'_d = \frac{1}{T'_{q0}}(-e'_d + (x_q - x'_q)i_q) \right.}. \quad (\text{A1c})$$

For two d - and two q -axes, 6-order (Anderson-Fouad's) model [38] (note that this model can be considered as a simplification of the Sauer-Pai's model [39]) to (A1a)-(A1c) following differential equations are added:

$$\mathbf{f} \Rightarrow \begin{cases} \dot{e}''_q = \frac{1}{T''_{d0}}(-e''_q + e'_q - (x'_d - x''_d)i_d) \\ \dot{e}''_d = \frac{1}{T''_{q0}}(-e''_d + e'_d + (x'_q - x''_q)i_q) \end{cases}. \quad (\text{A1d})$$

The common algebraic equations are:

$$\mathbf{g} = \begin{cases} 0 = P_m - P_g \\ 0 = P_{m0} - P_m \\ 0 = v_{f0} - v_f \\ 0 = V \sin(\delta - \theta) - v_d \\ 0 = V \cos(\delta - \theta) - v_q \end{cases}. \quad (\text{A2a})$$

For one q -axis, 3-order model to (A2a) the following algebraic equations are added:

$$\mathbf{g} = \begin{cases} 0 = v_q + r_a i_q - e'_q + x'_d i_d \\ 0 = v_d + r_a i_d - x_q i_q \end{cases}. \quad (\text{A2b}')$$

For one d - and one q -axis, 4-order model to (A2a) the following algebraic equations are added:

$$\mathbf{g} = \begin{cases} 0 = v_q + r_a i_q - e'_q + x'_d i_d \\ 0 = v_d + r_a i_d - e'_d - x'_q i_q \end{cases}. \quad (\text{A2b}'')$$

For two d - and two q -axes, 6-order (Anderson-Fouad's) model to (A2a) the following algebraic equation are added:

$$\mathbf{g} = \begin{cases} 0 = v_q + r_a i_q - e''_q + x''_d i_d \\ 0 = v_d + r_a i_d - e''_d - x''_q i_q \end{cases}. \quad (\text{A2b}''')$$

REFERENCES

- [1] A. J. van der Schaft and B. Maschke, "Discrete conservation laws and Port-Hamiltonian systems on graphs and complexes," arXiv:1107.2006v1, Jul. 2011.

- [2] A. R. Bergen and D. J. Hill, "A structure preserving model for power system stability analysis," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 1, pp. 25–35, Jan. 1981.
- [3] P. V. Kokotovic and P. W. Sauer, "Integral manifold as a tool for reduced-order modeling of nonlinear systems: A synchronous machine case study," *IEEE Trans. Circuits Syst.*, vol. 36, no. 3, pp. 403–410, Mar. 1989.
- [4] G. N. Ramaswamy *et al.*, "Synchronic modal equivalencing (SME) for structure-preserving dynamic equivalents," *IEEE Trans. Power Syst.*, vol. 11, no. 1, pp. 19–29, Feb. 1996.
- [5] A. J. van der Schaft, "On model reduction of physical network systems," in *Proc. 21st Int. Symp. Math. Theory Netw. Syst.*, Jul. 2014.
- [6] G. C. Verghese, "Getting to the gray box: some challenges in model reduction," in *Proc. Amer. Control Conf.* [Online]. Available: acc2009.a2c2.org/GrayBoxPost3.pdf.
- [7] A. Chang and M. M. Adibi, "Power system dynamic equivalents," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 8, pp. 1737–1743, Nov/Dec. 1970.
- [8] A. J. Germond and R. Podmore, "Dynamic aggregation of generating unit models," *IEEE Trans. Power App. Syst.*, vol. PAS-97, no. 4, pp. 1060–1068, Jul./Aug. 1978.
- [9] L. Wang, M. Klein, S. Yirga, and P. Kundur, "Dynamic reduction of large power systems for stability studies," *IEEE Trans. Power Syst.*, vol. 12, no. 2, pp. 889–895, May 1997.
- [10] F. Ma and V. Vittal, "Right-sized power system dynamic equivalents for power system operation," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1998–2005, Nov. 2011.
- [11] J. H. Chow, *Power System Coherency and Model Reduction*. New York, NY, USA: Springer, 2013.
- [12] G. N. Ramaswamy, C. Evrard, G. C. Verghese, O. Fillâtre, and B. C. Lesieutre, "Extensions, simplifications, and tests of synchronic modal equivalencing (SME)," *IEEE Trans. Power Syst.*, vol. 12, no. 2, pp. 896–905, May 1997.
- [13] J. M. Undrill and A. E. Turner, "Construction of power system electro-mechanical equivalents by modal analysis," *IEEE Trans. Power App. Syst.*, vol. PAS-90, no. 5, pp. 2049–2059, Sep. 1971.
- [14] J. J. Sanchez-Gasca, C. J. Bridenbaugh, C. E. J. Bowler, and J. S. Edmonds, "Trajectory sensitivity based identification of synchronous generator and excitation system parameters," *IEEE Trans. Power Syst.*, vol. 3, no. 4, pp. 1814–1822, Nov. 1988.
- [15] S. M. Benchluch and J. H. Chow, "A trajectory sensitivity method for the identification of nonlinear excitation system models," *IEEE Trans. Energy Convers.*, vol. 8, no. 2, pp. 159–164, Jun. 1993.
- [16] I. A. Hiskens, "Nonlinear dynamic model evaluation from disturbance measurement," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 702–710, Nov. 2001.
- [17] M. Burth, G. C. Verghese, and M. Velez-Reyes, "Subset selection for improved parameter estimation in on-line identification of a synchronous generator," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 218–225, Feb. 1999.
- [18] A. Chakraborty, J. H. Chow, and A. Salazar, "A measurement-based framework for dynamic equivalencing of large power systems using wide-area phasor measurements," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 68–81, Mar. 2011.
- [19] M. A. M. Ariff, B. C. Pal, and A. K. Singh, "Estimating dynamic model parameters for adaptive protection and control in power system," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 829–839, Mar. 2015.
- [20] Z. Zhu, G. Geng, and Q. Jiang, "Power system dynamic model reduction based on extended Krylov subspace method," *IEEE Trans. Power Syst.*, to be published.
- [21] S. Wang, S. Lu, N. Zhou, G. Lin, and M. Elizondo, "Dynamic-feature extraction, attribution, and reconstruction (DEAR) method for power system model reduction," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 20149–2059, Sep. 2014.
- [22] B. Marinescu, B. Mallem, and L. Rouco, "Large-scale power system dynamic equivalents based on standard and border synchrony," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1873–1882, Nov. 2010.
- [23] D. Chaniotis and M. A. Pai, "Model reduction in power systems using Krylov subspace methods," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 888–894, May 2005.
- [24] M. K. Transtrum and P. Qiu, "Model reduction by manifold boundaries," *Phys. Rev. Lett.*, vol. 113, no. 9, pp. 098701-1–098701-6, Aug. 2014.
- [25] M. K. Transtrum, B. B. Machta, and J. P. Sethna, "Why are nonlinear fits to data so challenging?" *Phys. Rev. Lett.*, vol. 104, no. 6, pp. 060201-1–060201-4, Feb. 2010.
- [26] M. K. Transtrum, B. B. Machta, and J. P. Sethna, "Geometry of nonlinear least squares with applications to sloppy models and optimization," *Phys. Rev. E*, vol. 83, no. 3, pp. 036701-1–036701-35, Mar. 2011.
- [27] M. K. Transtrum, A. T. Sarić, and A. M. Stanković, "Information geometry for model verification in energy systems," in *Proc. 19th Power Syst. Comput. Conf.*, Genoa, Italy, Jun. 20–24, 2016.
- [28] J. R. Leis and M. A. Kramer, "The simultaneous solution and sensitivity analysis of systems described by ordinary differential equations," *ACM Trans. Math. Softw.*, vol. 14, no. 1, pp. 45–60, Mar. 1988.
- [29] L. B. Rall, *Automatic Differentiation: Techniques and Applications*. New York, NY, USA: Springer-Verlag, 1981.
- [30] C. H. Bischof, H. M. Bücker, P. Hovland, U. Naumann, and J. Utke, eds., *Advances in Automatic Differentiation*. Berlin, Germany: Springer-Verlag, 2008.
- [31] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. San Francisco, CA, USA: Freeman, 1973.
- [32] M. K. Transtrum, G. Hart, and P. Qiu, "Information topology identifies emergent model classes." [Online]. Available: <http://arxiv.org/abs/1409.6203>
- [33] R. Gutenkunst, "Sloppiness, modeling, and evolution in biochemical networks," Ph.D. dissertation, Cornell Univ., Ithaca, NY, USA, 2007. [Online]. Available: www.lassp.cornell.edu/sethna/pubPDF/GutenkunstPhD.pdf
- [34] F. Milano, *Power System Modelling and Scripting*. London, U.K.: Springer, 2010.
- [35] DualNumbers.jl, GitHub repository. [Online]. Available: <https://github.com/JuliaDiff/DualNumbers.jl>
- [36] P. N. Brown, G. D. Byrne, and A. C. Hindmarsh, "VODE: A variable-coefficient ODE solver," *SIAM J. Sci. Statist. Comput.*, vol. 10, no. 5, pp. 1038–1051, Sep. 1989.
- [37] A. T. Sarić, M. K. Transtrum, and A. M. Stanković, "Dynamic model estimation for power system areas from boundary measurements," in *Proc. IEEE Power & Energy Soc. General Meeting, Session: Power Syst. Dynamic Performance*, Boston, MA, USA, Jul. 17–21, 2016, Paper 16PESGM0376.
- [38] P. M. Anderson and A. A. Fouad, *Power System Control and Stability*. New York, NY, USA: Wiley-IEEE Press, 2002.
- [39] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*. Englewood Cliffs, NJ, USA: Prentice Hall, 1998.

Mark K. Transtrum received the Ph.D. degree in physics from Cornell University, Ithaca, NY, USA, in 2011. He then studied computational biology as a Postdoctoral Fellow at MD Anderson Cancer Center, Houston, TX, USA. Since 2013, he has been an Assistant Professor of physics and astronomy with Brigham Young University, Provo, UT, USA. His work considers representations of a variety of complex systems including power systems, systems biology, materials science, and neuroscience.

Andrija T. Sarić was born in 1962. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of Belgrade, Belgrade, Serbia, in 1988, 1992, and 1997, respectively. He is a Full Professor of electrical engineering at the Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia. His main areas of interest are power system analysis, optimization and planning, as well as application of artificial intelligence methods in these areas.

Aleksandar M. Stankovic (F'05) received the Ph.D. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1993. From 1993 to 2010, he was with Northeastern University, Boston. He currently serves as an A.H. Howell Professor at Tufts University, Medford, MA, USA. He has held visiting positions at the United Technologies Research Center (sabbaticals in 2000 and 2007) and at L'Université de Paris-Sud and Supelec (in 2004). He is a coeditor of a book series on power electronics and power systems for Springer.

Prof. Stankovic is an Associate Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS. He has previously served in the IEEE TRANSACTION ON SMART GRID in power systems and on control system technology in the same capacity, since 1996.