Network Reduction in Transient Stability Models using Partial Response Matching

Benjamin L. Francis*, Jacob R. Nuttall†, Mark K. Transtrum‡
Dept. of Physics and Astronomy
Brigham Young University, Provo, UT, USA
*bfrancis17@byu.edu; †jacob_nuttall@outlook.com; ‡mktranstrum@byu.edu

Andrija T. Sarić§, Aleksandar M. Stanković¶
§Dept. of Power, Electronics and Com. Engineering
§Faculty of Technical Sciences, Novi Sad, Serbia
¶Dept. of Electrical and Computer Engineering
¶Tufts University, Medford, MA, USA
§asaric@uns.ac.rs; ¶astankov@ece.tufts.edu

***Abstract*—We describe a method for simultaneously identifying and reducing dynamic power systems models in the form of differential-algebraic equations. Often, these models are large and complex, containing more parameters than can be identified from the available system measurements. We demonstrate our method on transient stability models, using the IEEE 14-bus test system. Our approach uses techniques of information geometry to remove unidentifiable parameters from the model. We examine the case of a networked system with 58 parameters using full observations throughout the network. We show that greater reduction can be achieved when only partial observations are available, including reduction of the network itself.**

***Index Terms*—Parameter Estimation, Reduced Order Systems, System Identification.**

I. Introduction

Models of dynamical phenomena in power systems have grown significantly in size and level of detail. Besides the need for increased computing power to handle them, such complicated models are increasingly difficult to interpret in terms of the system-level behavior. In addition, many models have trouble reproducing data recorded from actual events in the grid [1]. Efforts to improve a model through system identification are inhibited by the difficulty of inferring model parameters, many of which are not well-constrained by the data [2]. These efforts could be enhanced by removing from the model those parameters that are not identifiable from data. By construction, the remaining parameters are identifiable from data, thereby reducing statistical uncertainty in their inferred values and improving model predictivity. This not only makes models more manageable computationally but also improves understanding of the relationships between system components.

In this paper, we use techniques of information theory combined with differential geometry (together "information geometry") to enhance system identification by removing unidentifiable parameters. From the view of information geometry, a model is seen as a manifold embedded in the space of measurement data. This so-called model manifold contains all information about model predictions. Accordingly, it captures the global properties of the model, as contrasted with the cost surface in parameter space which condenses this information into just a single number [3].

In what follows we give only a few references with direct connections to our work. An overview of model approximation methods for dynamical systems, including Krylov methods for linear systems, is given in [4]. Use of Krylov subspace methods for reducing linear power systems models is described in [5]. Modal approaches to model reduction in linear systems are discussed in [6]. Singular perturbation theory is applied in nonlinear power systems models to obtain simplified representations in [7]. Network identification in dynamic networks with known topology is discussed in [8]. We assume known network topology and consider simultaneous identification and reduction of dynamic and network parameters in nonlinear power systems models.

This work builds on our prior efforts to explore possible reductions that can be made in a networked system with 58 unknown parameters [9]. We previously predicted that when observing only part of the system, additional reduction could be achieved. Here, we present results of performing model reduction under the conditions of one such set of reduced observations and compare the extent of the reduction with previous estimates. These results serve as a prototype for network reduction in large power system models where only parts of the system can be observed. We also consider a more realistic fault scenario (see Sec. IV-A) and much longer observation times.

The outline of the paper is as follows. In Sec. II, we formulate the system identification problem. We give an overview of our model reduction procedure in Sec. III. We describe the test system in which these methods were applied in Sec. IV-A, followed by some model reduction results for the full set of observations in Sec. IV-B. Reduction results for the partial set of observations are given in Sec. IV-C. Section V gives concluding remarks.

This work was supported in part by the NSF under grant number EPCN-1710727.

II. Problem Formulation

Transient stability models of power systems are typically cast in differential-algebraic form [10]:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{z}, \mathbf{p}, t), \\ \mathbf{0} &= \mathbf{g}(\mathbf{x}, \mathbf{z}, \mathbf{p}, t), \\ \mathbf{y} &= \mathbf{h}(\mathbf{x}, \mathbf{z}, \mathbf{p}, t). \end{aligned} \tag{1}$$

Here $\mathbf{x}$ is a vector of (differential) state variables, $\mathbf{z}$ are the algebraic variables, $\mathbf{p}$ are parameters, t is the (scalar) time variable, and $\mathbf{y}$ is the vector of system observations being made. The parameters $\mathbf{p}$ are to taken to be unknown and are to be estimated from measurements $\mathbf{y}$, although some information, such as plausible ranges for each, may be available.

Often, available measurements are insufficient to identify all parameters in a large, complicated model [3]. Even when all of the parameters are identifiable in principle, the model's predictions may be insensitive to changes in certain combinations of parameters, making some parameters *practically* unidentifiable [11]. Such models can often be simplified while preserving the model's predictive capabilities.

Parameter identifiability can be analyzed using the Fisher Information Matrix (FIM), which is constructed from the parametric sensitivities. These, in turn, can be calculated by solving the sensitivity equations, which are obtained by differentiating (1):

$$\begin{aligned} \frac{d}{dt}\frac{\partial \mathbf{x}}{\partial \mathbf{p}} &= \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial \mathbf{f}}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{f}}{\partial \mathbf{p}}, \\ \mathbf{0} &= \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial \mathbf{g}}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{g}}{\partial \mathbf{p}}, \\ \frac{\partial \mathbf{y}}{\partial \mathbf{p}} &= \frac{\partial \mathbf{h}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial \mathbf{h}}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \mathbf{p}} + \frac{\partial \mathbf{h}}{\partial \mathbf{p}}. \end{aligned} \tag{2}$$

The eigenvalues of the FIM measure the sensitivity of model predictions to coordinated changes in various combinations of parameters. Figure 1a shows the eigenvalues of the FIM for the model discussed in Sec. IV-A. Relative sensitivity of different parameter combinations is seen in the large spread of eigenvalues; small eigenvalues indicate insensitive parameters. By systematically removing the associated parameter combinations from the model, the insensitivity can be removed (see Fig. 1b). Our method for removing these parameter combinations is discussed in the next section.

III. Model Reduction using the Manifold Boundary Approximation Method

We remove unidentifiable parameter combinations using the Manifold Boundary Approximation Method (MBAM) [12]. MBAM was first used in the context of power systems in [13]; here we provide only a brief summary. The model is reinterpreted as a mapping from the space of parameters to a second space known as data space (see Fig. 2). This mapping defines a manifold in data space, with parameters acting as coordinates on the manifold. Typically, model manifolds are bounded, with a hierarchical structure like a polygon (faces, edges, etc.). Each boundary cell corresponds to a simplifying approximation of the model. We identify these approximations by using computational differential geometry to construct geodesics on the model manifold (the analogs of straight lines on curved surfaces).

Fig. 1. Eigenvalues of the FIM for a) the original, 58-parameter model, observing all generator variables and bus voltages and angles (see Sec. IV-B); b) the 37-parameter reduced model, same observations as a; c) the original model, observing only generator variables on Bus 1 and voltages and angles on Buses 1 & 14 (see Sec. IV-C); and d) the 20-parameter reduced model, same observations as c. Our reduction method (see Sec. III) effectively removes only the smallest eigenvalues in either case (from a to b or from c to d). Limiting the region of observation (going from a to c) makes more of the parameters unidentifiable (note the difference in vertical axes); accordingly, it allows for additional reduction to be carried out in the model.

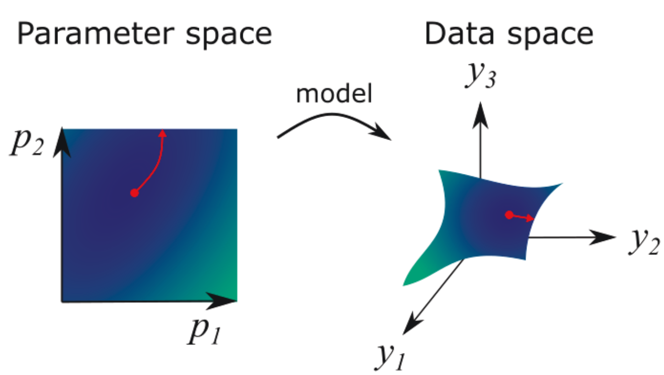

Fig. 2. Illustration of parameter space (on the left), colored by the objective function used for system identification, and the model manifold in data space (on the right). The model itself defines the mapping from parameter space to data space. The geodesic (red path) marked on the model manifold corresponds to a particular trajectory through parameter space. When the geodesic reaches the manifold boundary, some parameter (or perhaps set of parameters) goes to extreme values (see Fig.3).

When a geodesic encounters a boundary cell, some parameters are taken to extreme values (such as infinity or zero; see Fig. 3). This identifies a limit that can be evaluated in the equations of the model, removing one parameter and producing a simplified model. In addition to known limiting approximations (for example, a singular perturbation in which a time constant for subtransients is pushed to zero), this allows us to identify novel reductions that had not previously been considered, such as merging adjacent buses in a network (see Sec. IV-C). We construct a new model by evaluating the limit identified by the geodesic and then tune the parameters of the new model to match the predictions of the original model. This process is repeated until the predictions of the reduced model no longer faithfully reproduce those of the original. In this way, all behavior of the model that is measurably significant (e.g., participating modes) is preserved in the reduced model.

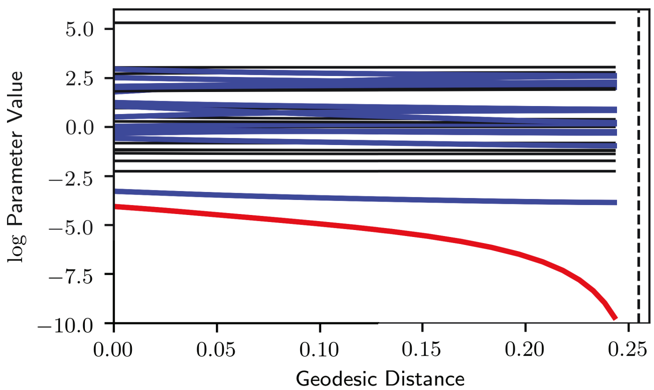

Fig. 3. Parameter values along a geodesic in the "supplier-consumer" model reduction (see Sec. IV-C). Some parameters remain mostly unchanged (black); others adjust to new values (blue); and one goes to zero (negative infinity in log values; red), encountering a singularity (dashed line, location approximate) when the geodesic reaches a manifold boundary. This indicates the limit $T''_{d0} \to 0$, which is a singular perturbation removing the d-axis subtransient.

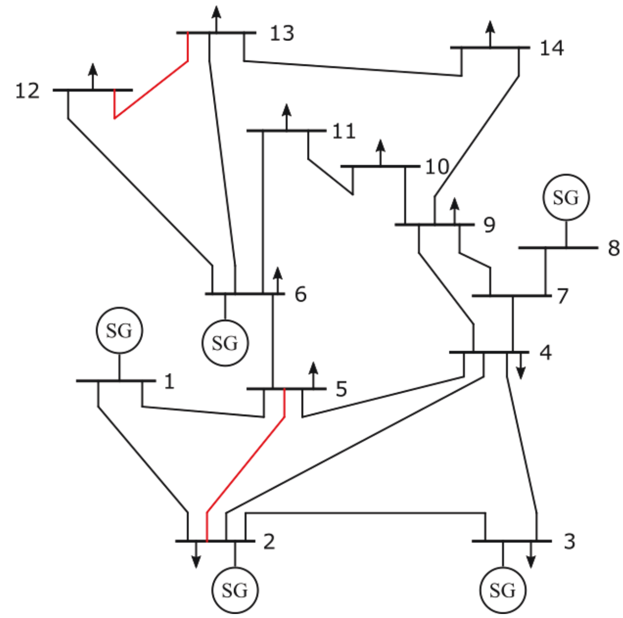

Fig. 4. IEEE 14-bus test system. Branches marked in red were removed by MBAM (see Sec. IV-B).

IV. Application

A. Test system

We use the IEEE 14-bus test system with five synchronous generators (SG) on Buses 1, 2, 3, 6, and 8 (Fig. 4). The generator in Bus 1 is implemented with a fourth-order model, including rotor angle, speed, and transient electromotive forces in the d- and q-axes. The generators in Buses 2 and 3 are implemented with a classical, second-order model for the generator speed and rotor angle. The generators in Buses 6 and 8 are both modeled with a detailed, sixth-order model, including both transient and subtransient dynamics in the d- and q-axes. We fix many parameters that are not to be estimated from transient dynamics (such as rotor moments of inertia) to predetermined values, leading to a model with 38 tunable parameters for both generator and controller elements.

To allow for network simplification as part of the model reduction procedure, we take the susceptance of each network edge as a tunable parameter and model the conductances as proportional to the susceptances. We motivate this choice by noting that both would be dependent on the length of the line being modeled, so this is effectively equivalent to letting the line lengths be tunable parameters. This produces 20 network parameters, for a total of 58 parameters in the whole model.

We assume the system is initially in steady state and perturb it at $t = 1$ s with a short circuit in Bus 14, which is subsequently cleared at $t = 1.25$ s. This is in contrast to [9], where the mechanical power seen by each generator was simultaneously increased at $t = 0$. We also consider much longer observation times (out to $t = 100$ s) than in [9], allowing us to include the long-time decay of transients. Transients out to 100 s are not typical in power systems modeling; here we have very slow dynamics, so long observation times are needed to capture them.

In this paper, two sets of observations of the resulting transients are considered. In the first set of observations, we include rotor angle, speed, and real & reactive powers in all generators as well as voltage magnitude and angle in all buses. Our reduction results for this "full" set of observations are discussed in Sec. IV-B. In the second set, we include the above four generator variables for the generator on Bus 1 only and voltage magnitude and angle on Buses 1 & 14 only. These observations are indicative of a supplier-consumer relationship and reflect what might be expected for sparse observations in a very large power system. We discuss these observations further in Sec. IV-C.

B. Full system identification

We used MBAM to reduce the original model, with the "full" set of observations, from 58 parameters down to 37. Most of the reductions we encountered occurred in generator and controller components on Buses 6 & 8. A complete list of reduction steps is given in Table I. We found several new types of limits which were not encountered in [9], including time constants going to infinity (rather than zero), controller parameters going to zero, and reactance limits not paired with a corresponding time constant limit (as would be the case in a singular perturbation approximation). In addition, this reduction did not remove the same network branches as in [9].

TABLE I
MODEL REDUCTION STEPS WITH "FULL" OBSERVATIONS.

Step	Reduction	Location	Step	Reduction	Location
1	$T'_{q0} \to \infty$	Bus 8	12	$B_{2,5} \to 0$	Line 2-5
2	$T'_{q0} \to \infty$	Bus 6	13	$T'_{d0} \to \infty$	Bus 6
3	$x_q \to \infty$	Bus 8	14	$x_d \to x'_d$	Bus 6
4	$T'_{d0} \to \infty$	Bus 8	15	$K_e \to 0$	Bus 6
5	$x_q \to \infty$	Bus 6	16	$x''_q \to 0$	Bus 6
6	$x_d \to \infty$	Bus 8	17	$K_a \to 0$	Bus 8
7	$K_e \to 0$	Bus 8	18	$K_a \to 0$	Bus 6
8	$B_{12,13} \to 0$	Line 12-13	19	$x''_d \to 0$	Bus 8
9	$x''_d \to 0$	Bus 6	20	$x_d \to x'_d$	Bus 1
10	$T''_{q0} \to 0$	Bus 8	21	$T''_{d0} \to 0$	Bus 6
11	$x_q \to x'_q$	Bus 8			

Transients for both the original and 37-parameter models are shown in Fig. 5 for Bus 6 (bus and generator variables). The discrepancy between the reduced and original models indicates that all parameters are now identifiable and the model reduction procedure is complete.

We solved the sensitivity equations (2) and calculated the FIM for this set of observations for both the original and 37-parameter models; eigenvalues are shown in Fig. 1, a and b. The large spread of the eigenvalues of the FIM in the original model, with many being very small, indicated that many of the parameters in the model were likely to be unidentifiable due to insensitivity. This is borne out by the eigenvalues calculated for the reduced model, where it is apparent that MBAM has effectively removed many of the smallest ones.

C. Partial response matching

In many contexts, only portions of the system being modeled are available for observation. This often makes many fewer parameters in the model identifiable from the available observations. In [9], we predicted that under such circumstances, additional reduction could be achieved.

Here we present results of continuing the reduction in Sec. IV-B with only partial observations – specifically, observing only Buses 1 & 14 (both bus and generator variables), as opposed to all buses. This set of observations characterizes a supplier-consumer relationship: a single generator bus and a single load bus elsewhere in the network. Changing to this partial set of observations causes many of the eigenvalues of the FIM to drop significantly (compare Fig. 1, a vs. c), indicating that many more parameters can be removed. In fact, we were able to reduce the number of parameters to 20 using MBAM – as few as predicted in [9]. Using sensitivity analysis to calculate the FIM and its eigenvalues for the 20-parameter model, we find that, as before, MBAM has removed only the smallest eigenvalues (see Fig. 1, c and d).

A full list of reduction steps is shown in Table II. The most notable difference from the reduction with the "full" set of observations is in the network reductions. Not only are many more branches cut, but several buses are also *merged* (indicated by the line susceptance going to infinity). This can be understood intuitively by noting that the limit $B_{ik} \to \infty$

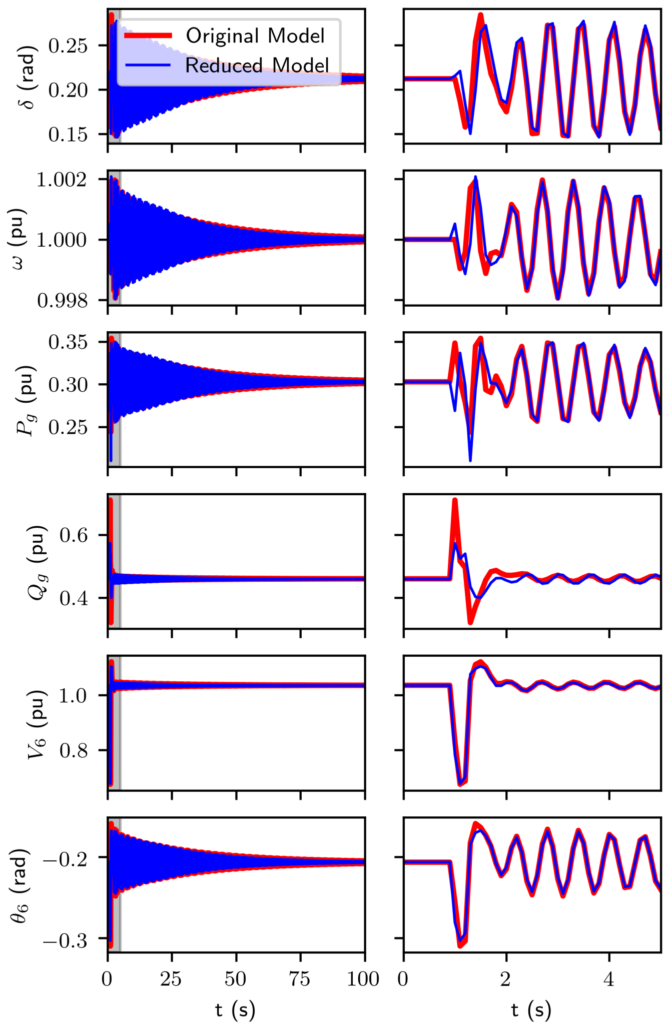

Fig. 5. Transients predicted in generator variables on Bus 6, as well as bus voltage and angle, following a short circuit in Bus 14 at $t = 1$ s (cleared at $t = 1.25$ s). The shaded regions from the left-hand plots are plotted on the right for detail.

can be achieved, for example, by letting the line length go to zero. For a mathematical proof, see the appendix. The resulting reduced network is shown in Fig. 6. Although many of the unobserved buses have merged, it is interesting to note that no buses have merged with the observed Buses 1 & 14.

Transients for all observables in this set are shown in Fig. 7 for both the 20-parameter and original models. There is still very good agreement between the two, except for the reactive power in the generator on Bus 1, which has begun to deviate significantly (again, indicating that all parameters are identifiable).

V. CONCLUSION

In this paper, we present results of using the Manifold Boundary Approximation Method (MBAM) to simultaneously reduce a dynamic power systems model and perform parameter identification for two choices of system measurements using the IEEE 14-bus test system. Compared to previous studies

TABLE II
MODEL REDUCTION STEPS WITH "SUPPLIER-CONSUMER" OBSERVATIONS.

Step	Reduction	Location	Step	Reduction	Location
1	$T'_{q0} \to \infty$	Bus 8	20	$x'_q, e'_d \to \infty$	Bus 8
2	$T'_{q0} \to \infty$	Bus 6	21	$B_{7,9} \to \infty$	Line 7-9
3	$x_q \to \infty$	Bus 8	22	$x''_q \to 0$	Bus 6
4	$T'_{d0} \to \infty$	Bus 8	23	$B_{7,8} \to \infty$	Line 7-8
5	$x_q \to \infty$	Bus 6	24	$T''_{d0}, x'_d, e'_q \to \infty$	Bus 8
6	$x_d \to \infty$	Bus 8	25	$x_d \to x'_d$	Bus 6
7	$K_e \to 0$	Bus 8	26	$B_{9,14} \to 0$	Line 9-14
8	$B_{12,13} \to 0$	Line 12-13	27	$T'_{d0} \to \infty$	Bus 6
9	$B_{10,11} \to \infty$	Line 10-11	28	$K_e \to 0$	Bus 6
10	$B_{9,10} \to \infty$	Line 9-10	29	$K_a \to 0$	Bus 6
11	$B_{6,12} \to \infty$	Line 6-12	30	$B_{4,7} \to \infty$	Line 4-7
12	$B_{2,5} \to 0$	Line 2-5	31	$B_{6,13} \to \infty$	Line 6-13
13	$T''_{q0} \to 0$	Bus 8	32	$T''_{d0} \to 0$	Bus 6
14	$x''_q \to 0$	Bus 8	33	$K_e \to 0$	Bus 3
15	$B_{4,9} \to 0$	Line 4-9	34	$B_{3,4} \to 0$	Line 3-4
16	$B_{6,11} \to 0$	Line 6-11	35	$B_{1,5} \to 0$	Line 1-5
17	$x''_d \to 0$	Bus 6	36	$x_d \to x'_d$	Bus 1
18	$B_{4,5} \to 0$	Line 4-5	37	$B_{5,6} \to \infty$	Line 5-6
19	$K_a \to 0$	Bus 8	38	$x'_d \to 0$	Bus 3

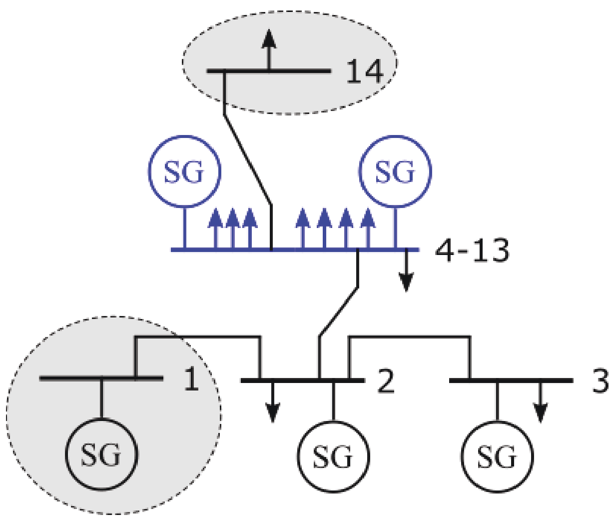

Fig. 6. Reduced network for "supplier-consumer" observations (shaded regions). Components marked in blue are from buses that were merged during the reduction.

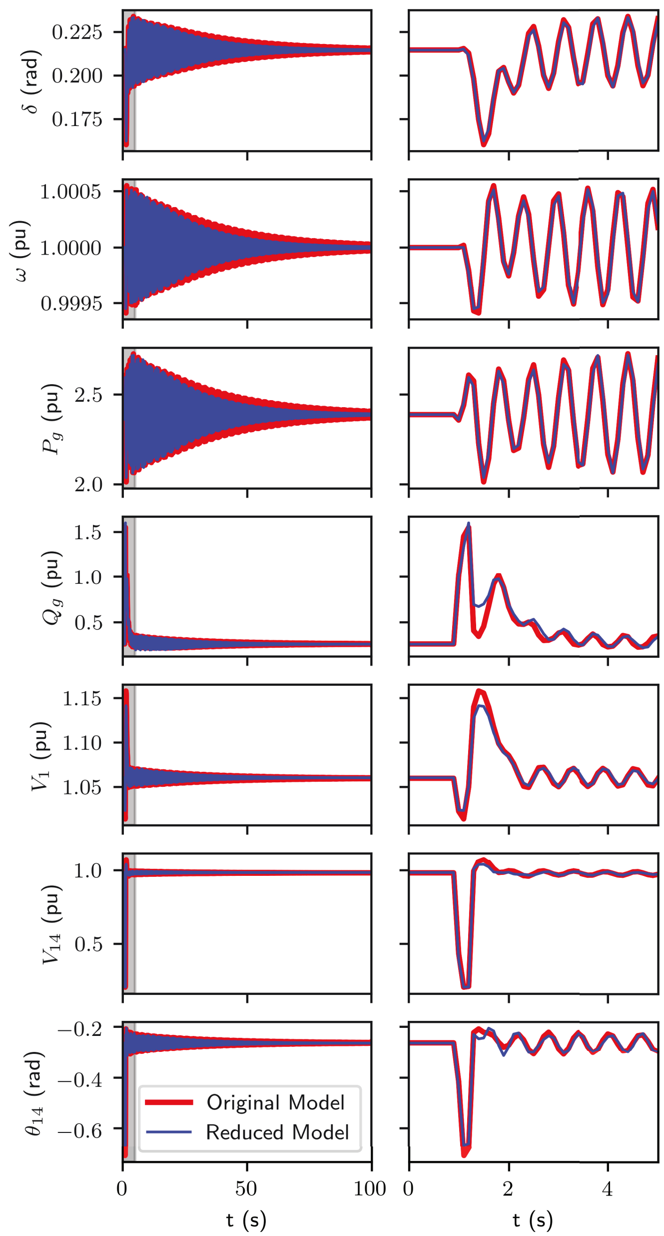

Fig. 7. Transients predicted in generator variables on Bus 1 as well as voltages and angles in Buses 1 & 14 ("supplier-consumer" observations; see Sec. IV-C). The shaded regions from the left-hand plots are plotted on the right for detail. θ_1 has been omitted because it was used as the reference angle.

[9], our model includes a more realistic fault scenario and longer observation times. In addition to implementing a more realistic model, our results go beyond those of previous studies in several important ways.

First, we have leveraged the data-driven nature of MBAM by reducing the system under only partial observations, in addition to the full set studied previously. We show that when a sparse set of observations is used, much greater reduction can be achieved (20 parameters remaining out of 58 for partial observations, as opposed to 37 out of 58 for full). In particular, we find that unobserved parts of the network can be greatly simplified. These results have important implications for large, networked power system models in which only part of the system is under observation, where we would expect to see similar network simplification. This also suggests that the method could be used to derive dynamic equivalents and other types of effective models in complex power networks.

We have also identified new types of approximations in the form of parameter limits that were not encountered in previous studies. Identifying types of approximations that are amenable to power systems is an important and necessary step to scaling up these methods. A catalog of potential parameter limits (derived from small or moderate-sized systems) can replace the expensie geodesic calculation on larger models that have similar mathematical structure [14]. Future work will focus on scaling up these methods to larger power systems models, using a combination of network decomposition, computational improvements, and theoretical insights gained from studies on smaller models such as this.

References

[1] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa *et al.*, "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1922–1928, 2005.

[2] M. Burth, G. C. Verghese, and M. Vélez-Reyes, "Subset selection for improved parameter estimation in on-line identification of a synchronous generator," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 218–225, 1999.

[3] M. K. Transtrum, B. B. Machta, and J. P. Sethna, "Why are nonlinear fits to data so challenging?" *Physical Review Letters*, vol. 104, no. 6, p. 060201, 2010.

[4] A. C. Antoulas and D. C. Sorensen, "Approximation of large-scale dynamical systems: An overview," Tech. Rep., February 2001. [Online]. Available: https://hdl.handle.net/1911/101964

[5] D. Chaniotis and M. Pai, "Model reduction in power systems using krylov subspace methods," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 888–894, 2005.

[6] D. Bonvin and D. Mellichamp, "A unified derivation and critical review of modal approaches to model reduction," *International Journal of Control*, vol. 35, no. 5, pp. 829–848, 1982.

[7] J. Chow, J. Winkelman, M. Pai, and P. W. Sauer, "Application of singular perturbations theory to power system modeling and stability analysis," in *1985 American Control Conference*. IEEE, 1985, pp. 1401–1407.

[8] A. S. Bazanella, M. Gevers, J. M. Hendrickx, and A. Parraga, "Identifiability of dynamical networks: which nodes need be measured?" in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 5870–5875.

[9] M. K. Transtrum, B. L. Francis, A. T. Saric, and A. M. Stankovic, "Simultaneous global identification of dynamic and network parameters in transient stability studies," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2018, pp. 1–5.

[10] P. Kundur, *Power system stability and control*. McGraw-hill New York, 1994, vol. 7.

[11] M. K. Transtrum, B. L. Francis, C. C. Youn, A. T. Saric, and A. M. Stankovic, "Geometrically motivated reparameterization for identifiability analysis in power systems models," in *2018 North American Power Symposium (NAPS)*. IEEE, 2018, pp. 1–6.

[12] M. K. Transtrum and P. Qiu, "Model reduction by manifold boundaries," *Physical Review Letters*, vol. 113, no. 9, p. 098701, 2014.

[13] M. K. Transtrum, A. T. Sarić, and A. M. Stanković, "Measurement-directed reduction of dynamic models in power systems," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2243–2253, 2016.

[14] M. K. Transtrum, B. L. Francis, A. T. Sarić, and A. M. Stanković, "Parameter reduction of large-scale models," In preparation.

Appendix

Here we show that the limit $B_{ik} \to \infty$ leads to merging the two Buses i and k. We begin with the power flow equations for Bus i:

$$0 = P_{g,i} - P_{d,i} + P_{inj,i} \tag{3}$$

$$0 = Q_{g,i} - Q_{d,i} + Q_{inj,i}, \tag{4}$$

$$P_{inj,i} = \sum_j V_i V_j \left[G_{ij}\cos(\theta_i - \theta_j) + B_{ij}\sin(\theta_i - \theta_j)\right] \tag{5}$$

$$Q_{inj,i} = \sum_j V_i V_j \left[G_{ij}\sin(\theta_i - \theta_j) - B_{ij}\cos(\theta_i - \theta_j)\right]. \tag{6}$$

With the conductances G_{ij} modeled as proportional to the susceptances B_{ij},

$$G_{ij} = -cB_{ij}, \tag{7}$$

the injected power becomes

$$P_{inj,i} = -\sum_j V_i V_j B_{ij}\left[c\cos(\theta_i - \theta_j) - \sin(\theta_i - \theta_j)\right] \tag{8}$$

$$Q_{inj,i} = -\sum_j V_i V_j B_{ij}\left[c\sin(\theta_i - \theta_j) + \cos(\theta_i - \theta_j)\right]. \tag{9}$$

Now, consider the limit $B_{ik} \to \infty$. To evaluate this limit, we first isolate all terms containing B_{ik}. Because

$$B_{ii} = -\sum_j B_{ij} = \ldots - B_{ik} + \ldots, \tag{10}$$

there are three nonzero terms containing B_{ik} in each of the power flow equations, one from $j = i$ in the sum and two from $j = k$:

$$\{\ldots\} = cV_i^2 B_{ik} - V_i V_k B_{ik}\left[c\cos(\theta_i - \theta_k) - \sin(\theta_i - \theta_k)\right]$$
$$\{\ldots\} = V_i^2 B_{ik} - V_i V_k B_{ik}\left[c\sin(\theta_i - \theta_k) + \cos(\theta_i - \theta_k)\right],$$

where $\{\ldots\}$ contains all other terms in each equation. Dividing through by B_{ik} gives

$$\frac{\{\ldots\}}{B_{ik}} = V_i\left[cV_i - cV_k\cos(\theta_i - \theta_k) + V_k\sin(\theta_i - \theta_k)\right] \tag{11}$$

$$\frac{\{\ldots\}}{B_{ik}} = V_i\left[V_i - cV_k\sin(\theta_i - \theta_k) - V_k\cos(\theta_i - \theta_k)\right]. \tag{12}$$

Taking the limit $B_{ik} \to \infty$ eliminates the left-hand side of both equations. Assuming $V_i \neq 0$, we have

$$0 = cV_i - cV_k\cos(\theta_i - \theta_k) + V_k\sin(\theta_i - \theta_k) \tag{13}$$

$$0 = V_i - cV_k\sin(\theta_i - \theta_k) - V_k\cos(\theta_i - \theta_k). \tag{14}$$

Rearranging (13) gives

$$V_k\sin(\theta_i - \theta_k) = -c\left[V_i - V_k\cos(\theta_i - \theta_k)\right], \tag{15}$$

which, when substituted into (14), gives

$$\begin{aligned} 0 &= V_i + c^2\left[V_i - V_k\cos(\theta_i - \theta_k)\right] - V_k\cos(\theta_i - \theta_k) \\ &= (c^2+1)\left[V_i - V_k\cos(\theta_i - \theta_k)\right] \\ &= V_i - V_k\cos(\theta_i - \theta_k). \end{aligned} \tag{16}$$

Plugging (16) back into (15) and assuming $V_k \neq 0$, we conclude that

$$V_k\sin(\theta_i - \theta_k) = 0, \tag{17}$$

$$\theta_i - \theta_k = n\pi \tag{18}$$

where we need only consider the two cases $n = 0$ or $n = 1$. We will show momentarily that we can exclude the possibility that $n = 1$.

We return to (16) and substitute $\theta_i - \theta_k = n\pi$:

$$0 = V_i - V_k\cos(n\pi). \tag{19}$$

If $n = 1$, $\cos(\pi) = -1$ and (19) cannot be satisfied (since both V_i and V_k are positive). Hence,

$$V_i = V_k, \tag{20}$$

$$\theta_i = \theta_k; \tag{21}$$

that is, Buses i and k have identical voltages and have, in effect, been merged.