

Information Geometry and Model Reduction in Oscillatory and Networked Systems

Benjamin Lane Francis

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Mark Transtrum, Chair
David Neilsen
Sean Warnick
Manuel Berrondo
Aleksandar Stanković

Department of Physics and Astronomy

Brigham Young University

Copyright © 2020 Benjamin Lane Francis

All Rights Reserved

ABSTRACT

Information Geometry and Model Reduction in Oscillatory and Networked Systems

Benjamin Lane Francis
Department of Physics and Astronomy, BYU
Doctor of Philosophy

In this dissertation, I consider the problem of model reduction in both oscillatory and networked systems. Previously, the Manifold Boundary Approximation Method (MBAM) has been demonstrated as a data-driven tool for reducing the parametric complexity of so-called sloppy models.

To be effective, MBAM requires the model manifold to have low curvature. I show that oscillatory models are characterized by model manifolds with high curvature in one or more directions. I propose methods for transforming the model manifolds of these models into ones with low curvature and demonstrate on a couple of test systems.

I demonstrate MBAM as a tool for data-driven network reduction on a small model from power systems. I derive multiple effective networks for the model, each tailored to a specific choice of system observations. I find several important types of parameter reductions, including network reductions, which can be used in large power systems models.

Finally, I consider the problem of piecemeal reduction of large systems. When a large system is split into pieces that are to be reduced separately using MBAM, there is no guarantee that the reduced pieces will be compatible for reassembly. I propose a strategy for reducing a system piecemeal while guaranteeing that the reduced pieces will be compatible. I demonstrate the reduction strategy on a small resistor network.

Keywords: model reduction, parameter inference, oscillatory systems, networks, information geometry, power systems, piecemeal reduction

ACKNOWLEDGMENTS

I am deeply grateful for the guidance and mentorship of my advisor, Mark Transtrum. He continually encouraged me to do more and to think deeper while simultaneously being sensitive to my individual needs, challenges, and abilities. He helped me to grow by prodding me towards more clarity when my ideas or arguments were soft and by giving me time and space to come at things in my own way and work at my own pace.

The faculty and staff of the Department of Physics and Astronomy have been incredibly helpful and supportive along the way, encouraging deep thinking and diligent effort, answering questions, and providing guidance and insight. I have also greatly benefited from working and collaborating with many of the other graduate students in classes along the way. In particular, it has been my pleasure to share office space with Alden Pack, who has given both a listening ear and useful feedback from time to time.

I also express my sincere appreciation for the influence and support of family members along my journey. Specifically, I want to acknowledge my mother and father, who, besides continually providing for me and loving me the best they knew how, set the example of higher education in our home and encouraged learning and the pursuit of knowledge; my brother, Timothy, who was the one who originally prompted me to consider majoring in the sciences, rather than the other options I was considering at the time; and my sister Trish and brother-in-law Jared, whose examples of loving-kindness and inclusivity have been a constant source of light in my life.

Finally, I'd like to acknowledge the support of many other individuals and friends whom I cannot name. Without their hands reaching out again and again, I never could have accomplished anything.

This work was supported in part by the US National Science Foundation under Award EPCN-1710727. Resources at the BYU Office of Research Computing were used and are gratefully acknowledged.

Contents

Table of Contents	iv
1 Information Geometry and Model Reduction in Oscillatory and Networked Systems	1
1.1 Introduction	1
1.2 Previous Work	2
1.3 Overview	5
1.3.1 Similarity Measures in Oscillatory Systems	5
1.3.2 Network Reduction in Power Systems	5
1.3.3 Piecemeal Reduction of Large Networks	6
2 Background and Methodology	7
2.1 Information Geometry	7
2.1.1 Least squares regression	7
2.1.2 The model manifold	8
2.1.3 The Fisher Information Matrix and sloppiness	9
2.1.4 Information topology: manifold boundaries	12
2.2 Power Systems Modeling	14
2.2.1 Equations of motion	15
2.2.2 Power flow analysis	17
2.2.3 Transient stability analysis	20
3 Similarity Measures in Dynamical Systems	23
4 Data-Driven Model Reduction for Network Simplification and Effective Equivalents	39
5 Piecemeal Reduction of Large Systems	51
5.1 Introduction	51
5.2 Navigating the Hasse diagram	55
5.2.1 Prime posets	58
5.2.2 Partial navigation of the Hasse diagram	61
5.3 Piecemeal Reduction Strategy	61
5.3.1 Fidelity of closed-loop behavior	64

5.3.2	Respecting conservation laws and maintaining physical interpretability . .	65
5.3.3	Preserving connectivity	65
5.3.4	Proposed strategy	68
5.4	Resistor Network Example	69
5.5	Discussion and conclusions	72
6	Conclusion	75
	Bibliography	79

Chapter 1

Information Geometry and Model Reduction in Oscillatory and Networked Systems

1.1 Introduction

Models are used as a tool for studying, understanding, and predicting the behavior of a system under various conditions. They simplify communication by summarizing knowledge that has been accumulated about the system. In engineering contexts, they can be used to evaluate system performance before the system is even built, making it cheaper and more efficient to explore various designs. In other contexts ranging from medicine to economics to meteorology to public health, they give the modeler an opportunity to explore the likelihood of certain events before they happen and examine the impact of potential interventions before they are put into practice.

These promises of modeling assume that the models being used are a sufficiently accurate representation of reality. This is especially true for quantitative mathematical models, where care

must be taken to validate them against measurement data and properly infer model parameters that cannot be measured directly. Efforts to improve mathematical models in many fields have given rise to increasingly detailed models of system components, which require both greater computational power to use and better measurement data to validate. Advances in computing continue to push the boundaries of the size and complexity of models that can be handled, and new sensors and measurement techniques are allowing data to be accumulated at unprecedented rates. But even when a large model is within computational limits and sufficient measurements are available to identify all of its parts and validate it, excessive detail obfuscates understanding of the relationships among system components - understanding that acts as a springboard for the insight necessary to solve difficult problems. This motivates reducing models in order to coarse-grain away the “microscopic” detail in order to clarify the “macroscopic” view one is interested in.

For some applications people already intuitively do this. For example, in circuit analysis, details of surface effects in a current-carrying wire that would require Maxwell’s equations to model are ignored and only the bulk current is usually considered. For simple circuits involving a few resistors, inductors, and/or capacitors, this gives a clearer picture of how different elements interact to produce the overall behavior of the circuit. But what about a large circuit (like the power grid of a region), where the number of components and lines in the circuit is in the thousands? Or what about models of the brain involving millions of neurons, complicated “networks” of chemical reactions involved in cellular processes or combustion, or large ecosystems with hundreds or thousands of interacting species and resources? How do you appropriately coarse-grain away detail in systems like these?

1.2 Previous Work

The idea of coarse-graining was introduced by Kadanoff [1] in the field of statistical mechanics. Kadanoff, in studying the behavior of the Ising model of spins on a lattice, wanted a way to

connect the microscopic picture of the behavior of individual spins in the lattice with the observed macroscopic behavior of a magnet near its critical temperature. He proposed treating groups of spins as units, using the total magnetization of the unit in place of individual spins and replacing the interactions of individual spins with an effective interaction between units that is similar to the one between spins but with adjusted parameters. This process is formalized by the Renormalization Group. More recently, the coarse-graining idea has been used as a method for model reduction in molecular dynamics simulations, where atoms are grouped and replaced with effective particles whose properties are determined from the underlying group, and the interactions between atoms are replaced with effective potentials between these particles [2–5]. Usually these methods require modelers to make choices about how the atoms are grouped and what form the effective potentials will take, based on expert intuition and available measurement data.

In the general area of dynamical systems, many previous efforts at model reduction have concentrated on linear (or linearized) systems with methods such as balanced truncation [6], singular perturbation [7, 8], Krylov subspace projection [9–12], or other methods [13–15]. Most of these attempt to reduce the order (i.e., number of differential equations) of the model, rather than its parametric complexity (i.e., number of tunable parameters) and focus on the input-output response of the system. Singular perturbation methods are usable for nonlinear systems but typically treat a specific class of approximations that distinguish slow- and fast-timescale dynamics, without considering other effectively small or large parameters or coordinated parameter effects. A commonly-used approach is to reduce the dimensionality of the system by projecting onto a subspace that approximately captures the dynamics of the full system [16–18]. Balanced truncation falls into this category, but it also includes strategies designed for nonlinear systems such as proper orthogonal decomposition (POD) with Galerkin projection and extensions of balanced truncation. The problem with all such schemes is that projection generally does not guarantee preservation of underlying physical conservation laws or symmetries, potentially introducing instabilities in the

model when reduction is performed on components in open loop. Extensions of POD or Galerkin projection methods that preserve energy, stability, or other characteristics exist [19–23], but only for certain types of systems (e.g., Hamiltonian systems). In addition, most of the above methods require researchers to choose the level of approximation introduced, although in many cases heuristics are given to guide such choices.

Most models that practitioners are interested in reducing have many tunable parameters whose values must be inferred from measurements. Multiparameter models from a wide variety of fields fall into a universality class known as *sloppy* models [24–28]. The predictions of these models are very sensitive to coordinated changes in some combinations of parameters, but very insensitive to others. Recently, a data-driven approach to parametrized model reduction for nonlinear systems has arisen in the field of information geometry, known as the Manifold Boundary Approximation Method (MBAM), that takes advantage of this parameter insensitivity by removing insensitive parameter combinations from the model [29]. Having a data-driven method allows the reduction to be tailored to the types and precision of available measurement data. Reduction using MBAM proceeds by finding and applying physically-meaningful limits in the model parameters, so any reduced models that are constructed represent the idealized behaviors of the original model in specific parameter regimes. This, in turn, means that underlying conservation laws and symmetries are respected during the reduction process.

MBAM has been applied to models from systems biology and biophysics [30–32], power systems [33,34], nuclear structure physics [35], neuroscience [36], and general linear time-invariant (LTI) systems [37]. In this dissertation, I lay the foundation for application of MBAM in general oscillatory systems, which often exhibit large parameter sensitivities and other characteristics that differentiate them from other sloppy models. I also demonstrate using MBAM as a tool for reducing networks, with specific application in power systems and electric circuits.

1.3 Overview

The outline of this work is as follows. In Ch. 2, I review concepts and methods from information geometry, and as well as modeling in power systems. This is followed by three chapters (described below) which take the form of journal articles. Concluding remarks are found in Ch. 6.

1.3.1 Similarity Measures in Oscillatory Systems

In order to quantitatively compare measured phenomena and model predictions, some kind of similarity measure must be used. For dynamical systems it is common to compare time series data to predictions using the method of least squares. I show, in Ch. 3, that this is generally inappropriate for oscillatory systems. I also define new similarity measures that address many of the issues with the usual approach.

1.3.2 Network Reduction in Power Systems

Many of the systems that scientists and engineers are interested in studying consist of large, complicated networks of many interacting components. Equations used to model the components are usually known, resulting in very large models for the entire networked system. For some applications, measurements on the whole system are impractical, if not entirely unavailable. In such cases, reducing parts of the model that are irrelevant for predicting the available measurements not only reduces the computation necessary to use the model, but also helps provide insight into the “coarse-grained” system-level behavior. In Ch. 4, I use MBAM as a tool for network reduction on a small network from the field of electric power systems. I show the relationship between the choice of observations being made on the system and the level of reduction that can be achieved in the model.

1.3.3 Piecemeal Reduction of Large Networks

The computations necessary to reduce models of very large systems using MBAM are cumbersome. New tools are needed for efficiently extending methods developed on small systems to large ones. One way to do this is to partition the system into manageable pieces, reduce them separately, and then put the pieces back together. I address issues that arise, outline a piecemeal reduction strategy, and demonstrate on a network of resistors in Ch. 5.

Chapter 2

Background and Methodology

2.1 Information Geometry

This section contains a brief review of concepts from [25, 26, 29].

Consider a parametrized model $f(t, \boldsymbol{\theta})$ of a system. This model makes predictions for the behavior $\{y_i\}_{i=1}^M$ of the system observed under various experimental conditions $\{t_i\}_{i=1}^M$ (e.g., at different times and/or for different inputs) based on the values of the parameters $\boldsymbol{\theta}$. Parameter inference consists in finding the values of $\boldsymbol{\theta}$ for which the model predictions $\{f(t_i, \boldsymbol{\theta})\}_{i=1}^M$ most closely fit the observed behavior $\{y_i\}$. Closeness is defined through the use of a metric or similarity measure (known variously as a loss, error, objective, or cost function).

2.1.1 Least squares regression

It is common in statistics to assume that the data can be reproduced by the model plus a stochastic term to account for discrepancies,

$$y_i = f(t_i, \boldsymbol{\theta}) + \zeta_i, \tag{2.1}$$

where ζ_i is a random variable that is independently distributed according to $N(0, \sigma_i)$, with σ_i being the experimental uncertainty in the measurement y_i . This leads to a probability distribution function

$$P(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{(2\pi)^{M/2} \prod_{i=1}^M \sigma_i} \exp \left(-\frac{1}{2} \sum_{i=1}^M \left(\frac{y_i - f(t_i, \boldsymbol{\theta})}{\sigma_i} \right)^2 \right), \quad (2.2)$$

which is a likelihood function for the observations $\{y_i\}$ (which collectively form the observation vector \mathbf{y}). Maximizing the likelihood leads to minimizing the sum of squares in the exponential, resulting in the following metric:

$$C(\boldsymbol{\theta}) = \frac{1}{2} \sum_i \left(\frac{y_i - f(t_i, \boldsymbol{\theta})}{\sigma_i} \right)^2, \quad (2.3)$$

It is possible to obtain other metrics if different assumptions are made about the generating process for $\{y_i\}$.

The graph of $C(\boldsymbol{\theta})$ is a surface over parameter space, so the parameter inference problem translates to one of finding the global minimum of this surface. To illustrate, consider the following rational function, which could be used to model some system for which data is available (see Fig. 2.1):

$$f(t, \boldsymbol{\theta}) = \frac{1}{1 + \theta_1 t + \theta_2 t^2 + \dots} \quad (2.4)$$

Leaving out the other parameters for the moment, I plot the cost surface over θ_1 and θ_2 in Fig. 2.2.

2.1.2 The model manifold

Being a metric or distance measure, $C(\boldsymbol{\theta})$ can be interpreted as the (squared) distance between two vectors \mathbf{y} and $\mathbf{f}(\boldsymbol{\theta})$ (which is the vector formed from the collection of predictions $\{f(t_i, \boldsymbol{\theta})\}$). These two vectors live in a space known as *data space* because it is the space of all possible data vectors \mathbf{y} that could have been obtained for the given experimental conditions [25, 26]. As $\boldsymbol{\theta}$ is varied, $\mathbf{f}(\boldsymbol{\theta})$ sweeps out a surface in this space known as the *model manifold*, which is the primary object of study in information geometry (see Fig. 2.3 for an example). In this view, the model $\mathbf{f}(\boldsymbol{\theta})$

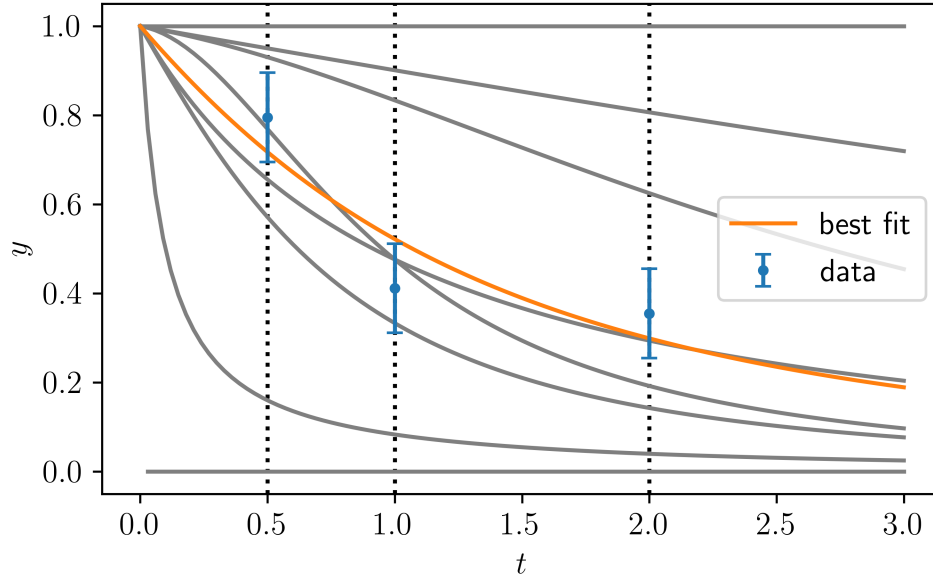


Figure 2.1 Predictions of the model in Eq. (2.4) for various choices of θ . The “best fit” curve minimizes the cost in Eq. (2.3). Each of the vertical “slices” represents a single axis in *data space* (see Fig. 2.3).

is a mapping from parameter space to data space, with the parameters acting as coordinates on the model manifold. The Jacobian $J_{i\mu} = \partial f_i / \partial \theta_\mu$ of this mapping, also known as the sensitivity matrix, characterizes the response of the model to changes in the parameters.

2.1.3 The Fisher Information Matrix and sloppiness

The shapes of the cost surface in parameter space and of the model manifold in data space are closely related. To see this, first consider the gradient and Hessian of the cost function:

$$\nabla_\mu C(\theta) = \frac{\partial C}{\partial \theta_\mu} = -\sum_i \frac{1}{\sigma_i^2} (y_i - f_i(\theta)) \frac{\partial f_i}{\partial \theta_\mu}, \quad (2.5)$$

$$H_{\mu\nu}(\theta) \equiv \frac{\partial^2 C}{\partial \theta_\mu \partial \theta_\nu} = \sum_i \frac{1}{\sigma_i^2} \left(\frac{\partial f_i}{\partial \theta_\mu} \frac{\partial f_i}{\partial \theta_\nu} - (y_i - f_i(\theta)) \frac{\partial^2 f_i}{\partial \theta_\mu \partial \theta_\nu} \right). \quad (2.6)$$

Near the best fit parameters θ^* , as long as the fit is good it can be assumed that the deviations $(y_i - f_i(\theta^*))$ are small, so the gradient will be negligible and the Hessian, dominated by the first

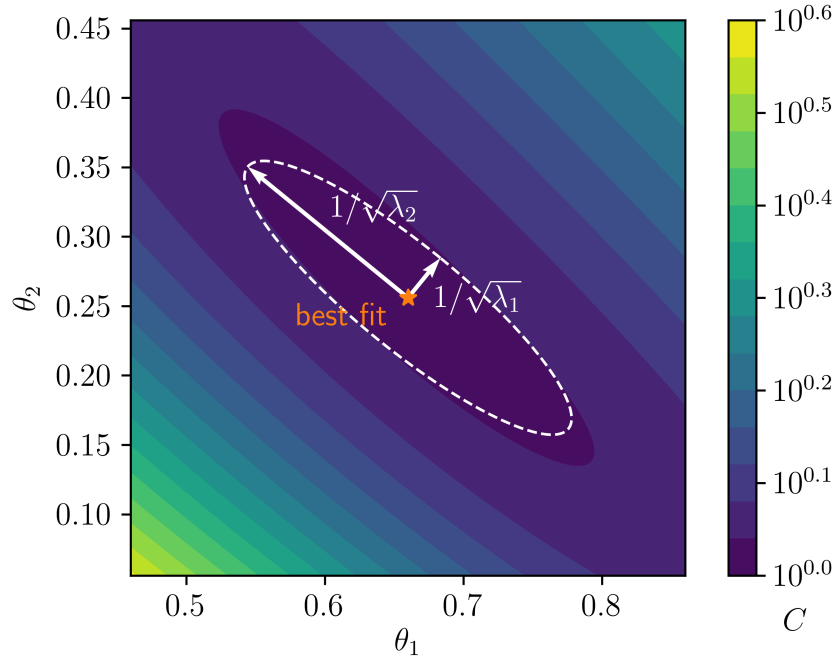


Figure 2.2 Example cost surface in parameter space for a rational function model with two parameters [see Eq. (2.4)]. Near the best fit (point of lowest cost), level contours of constant cost approximately form ellipses, whose principle axes (length and orientation) are related to the eigenvalues and eigenvectors of the Hessian of the cost.

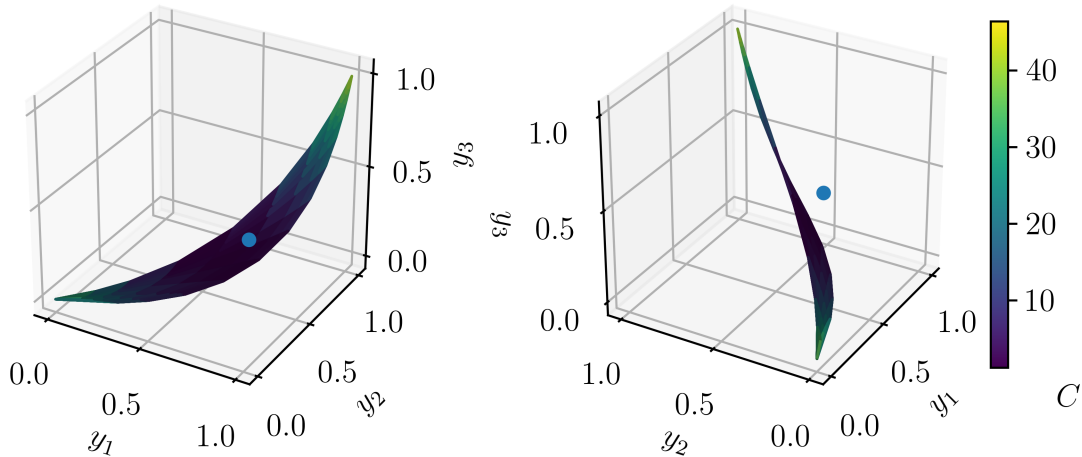


Figure 2.3 Example manifold in data space for a rational function model with two parameters [see Eq. (2.4)] being used to predict three measurements (y_1, y_2, y_3) , which are represented collectively by the plotted point. This manifold is a 2D surface because there are two parameters in the model, which act as coordinates on the model manifold. It is embedded in a 3D space because the model is trying to predict three data (see Fig. 2.1).

term in Eq. (2.6), will primarily determine the shape of the cost surface (namely an elliptical bowl; see Fig. 2.2). In fact, the expectation value of the Hessian at the best fit is an important object in statistics known as the Fisher Information Matrix (FIM):

$$I_{\mu\nu} \equiv \langle H_{\mu\nu}(\boldsymbol{\theta}^*) \rangle = \sum_i \frac{1}{\sigma_i^2} \frac{\partial f_i}{\partial \theta_\mu} \frac{\partial f_i}{\partial \theta_\nu} \quad (2.7)$$

The eigenvectors of the FIM indicate the directions of the principle axes of the hyperellipsoids which form approximate surfaces of constant cost near the best fit. The lengths of these axes are proportional to the inverse square roots of the eigenvalues.

Many multiparameter models used by practitioners exhibit a property known as *sloppiness* in which the model responds very strongly to a few so-called *stiff* combinations of parameters and is very insensitive to others (which are called *sloppy*). This is reflected in the eigenvalues of the FIM for such a model, which will be spread over many orders of magnitude (see Fig. 2.4). In the cost

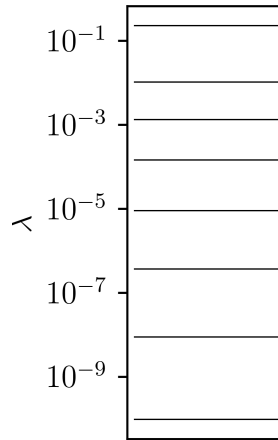


Figure 2.4 Example spectrum of eigenvalues of the FIM for a rational function model with eight parameters [see Eq. (2.4)]. Notice that the eigenvalues are spread over many orders of magnitude. This is an indication of *sloppiness*.

surface for a sloppy model, the bowl, or rather canyon, in which the best fit lies is characteristically long in the sloppy directions and thin in the stiff ones.

2.1.4 Information topology: manifold boundaries

Most observed (or desired) system behavior is finite and does not grow without bound, so the model manifolds arising out of models for these systems are typically *bounded*. Additionally, most *sloppy* models have model manifolds that exhibit a *hyperribbon* structure that is very long and thin (see Fig. 2.3). This can be understood by noting that, for analytic functions, the output of the function at neighboring points tends to be correlated. If the output at one point is fixed, the total amount of variation of other points in a neighborhood will be limited by constraints of analyticity. As the number of fixed points is increased, the amount of variation possible at interpolating points sampled from the function decreases by roughly a constant factor [26]. Applying this argument to models that are analytic functions of the parameters, cross-sections of the model manifold will likewise have a hierarchy of widths, each roughly a constant factor smaller than the previous one. The response of model predictions to changes in parameters typically reflects this property, so the eigenvalues of the FIM for a sloppy model tend to be distributed in a similar way. (For a rigorous treatment of the origin of the hyperribbon structure of model manifolds, see [38].)

Because model predictions are very insensitive to coordinated changes in some combinations of parameters (eigendirections of the FIM with very small eigenvalues), these parameters can be taken to extreme values (such as zero or infinity) while introducing very little error. This motivates an examination of the limiting behaviors of the model, which are reflected in the boundary structure of the model manifold. Topologically, for a model with N parameters, the model manifold is an N -dimensional volume which is bounded by a set of $(N - 1)$ -dimensional *faces*, which in turn are bounded by $(N - 2)$ -dimensional faces, and so on. This leads to a *partially ordered set* (specifically, a *graded poset*, and in many cases an *abstract polytope*) of boundary cells that can be schematically represented with a *Hasse diagram*. For example, the model manifold in Fig. 2.3 is a two-dimensional surface that is bounded by one-dimensional edges, which come together at zero-dimensional corners. The Hasse diagram for the boundary complex of this model manifold is shown in Fig. 2.5.

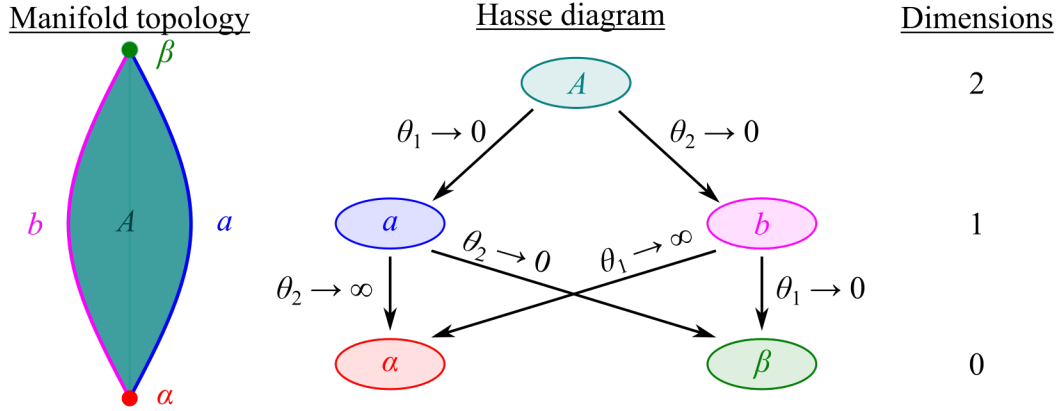


Figure 2.5 Hasse diagram for the boundary complex of the model manifold in Fig. 2.3. The topological structure of this manifold is that of a *digon* [left], a two-dimensional shape with only two edges. The Hasse diagram [center] schematically represents the boundary structure of this shape. Moving down the diagram corresponds to going from a given N -dimensional face to one of the $(N - 1)$ -dimensional faces that form its boundary. In terms of the model, this corresponds to evaluating a limit in the parameters. Doing so reduces the number of parameters (and hence the dimensionality) by one [right].

Each succeeding face in the hierarchy of boundary cells has one less dimension, so it is characterized by one less parameter. Boundary cells thus represent limiting approximations of the model with lower complexity. When the behavior of interest is close to the boundary of the model manifold, an appropriate boundary cell with fewer parameters may be used to model the behavior without introducing significant error. These boundary cells can be found by exploring the model manifold using *geodesics*, which are the analog of straight lines for curved surfaces. This method of approximating a model manifold by an appropriate boundary cell is known as the Manifold Boundary Approximation Method (MBAM) [29]. A summary of the method is as follows:

1. Calculate the FIM and identify the eigendirection with smallest eigenvalue.
2. Numerically construct a geodesic from the initial parameters in this direction until a boundary cell is encountered.

3. Identify the appropriate mathematical limit corresponding to the boundary cell and evaluate it in the model.
4. Fit the reduced model to the original and check for fidelity.
5. Repeat until the reduced model is no longer faithful to the original.

2.2 Power Systems Modeling

The following is an overview of some important power systems concepts. A couple of good introductory texts are [39] and [40], while some of the more technical details of modeling can be found in [41–43].

Power systems are the combination of components and networks designed to produce and deliver electrical power to consumers. Broadly speaking, power systems can be thought of as having three main parts: generation, transmission, and distribution. Modeling of power systems is done at many levels and serves a variety of purposes. Typical elements of a power systems model at the transmission level include generators, buses, transmission lines, and loads, each of which may be modeled with varying levels of detail. A bus (short for busbar) is a point of connection for the other components, i.e., generators and loads are connected to buses and transmission lines connect buses to each other. In the network structure of a power system, buses serve as nodes and transmission lines serve as edges (see Fig. 2.6).

The general scenario is that electrical power is produced by generators and then *injected* into the buses to which these generators are connected. From there, power *flows* through the network of transmission lines to all the other buses, where it is ultimately *consumed* by loads. *Power flow analysis* consists in numerically solving for this flow of power when the system is in steady state (i.e., producing AC electricity at 60 Hz). This usually provides the starting point for *transient stability analysis*, in which the response of the system to various *faults* (abnormal currents or voltages, e.g.,

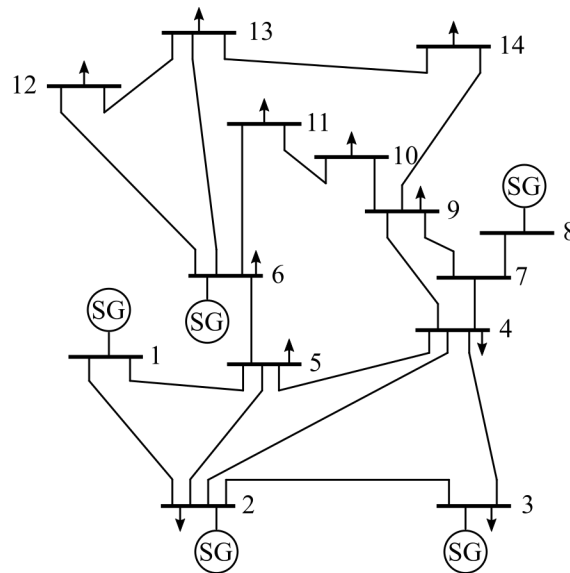


Figure 2.6 Example power systems network. Buses (horizontal lines, numbered) are connected via transmission lines. Generators (circles marked “SG”) “inject” or supply power at buses. Power flows through the network to loads (arrows), where it is “consumed”. This network is known as the IEEE 14-bus test system.

due to a short-circuit) is analyzed. In a transient stability model, the system is initialized in steady state, a fault is introduced, and the resulting *transient* in the variables of interest is observed. This transient consists of the oscillating deviations of the system around equilibrium, which presumably damp out over time (assuming the system is stable).

2.2.1 Equations of motion

The equations used to model power system phenomena reflect the fact that these phenomena cover an incredibly broad range of time scales, from 10^{-7} s (lightning effects) to 10^5 s (day-to-day power demand) [42, 43]. In principle, most of the devices in a system can be modeled with continuous dynamics, resulting in ordinary differential equations (ODEs). However, the dynamics of some variables in the system are likely to be considerably faster or slower than the behavior one is primarily interested in. In practical modeling situations, variables with slow dynamics are approximately

constant and can be taken as control parameters or inputs, whereas those with fast dynamics are approximately instantaneous and can be taken as *algebraic*, i.e., defined by an algebraic relationship to the other variables.

To illustrate, consider the following ODE:

$$\begin{aligned} T_x \dot{x} &= f(x, z, u, t), \\ T_z \dot{z} &= g(x, z, u, t), \\ T_u \dot{u} &= \varphi(x, z, u, t). \end{aligned} \tag{2.8}$$

The time constants T_x , T_z , and T_u represent the speed of the dynamics of the variables x , z , and u . Assume the timescale of the behavior of interest is $O(T_x)$. If $T_u \gg T_x$, then the dynamics of u are much slower than those of x and the system can be approximated with the limit $T_u \rightarrow \infty$, which leads to $\dot{u} = 0$ or $u(t) = u(0) = u_0$ (hence constant). If $T_z \ll T_x$, then the dynamics of z are much faster than those of x and the system can be approximated with the limit $T_z \rightarrow 0$, which leads to $0 = g(x, z, u, t)$, which is an algebraic constraint for the variable z . The result is that (2.8) becomes a *differential-algebraic* equation (DAE):

$$\begin{aligned} T_x \dot{x} &= f(x, z, u_0, t), \\ 0 &= g(x, z, u_0, t). \end{aligned} \tag{2.9}$$

In most cases, \mathbf{x} and \mathbf{z} will actually be vectors of state variables. These often differ from the variables \mathbf{y} that can actually be observed, so the system in (2.9) is usually supplemented with a set of *observation* equations $\mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{z}, \mathbf{u}_0, t)$. In addition, the functions \mathbf{f} , \mathbf{g} , and \mathbf{h} will depend on various device or system parameters \mathbf{p} .¹ A general power systems model thus takes the following

¹ In the context of power systems, I will use \mathbf{p} for parameters, rather than $\boldsymbol{\theta}$, to avoid confusion with the conventional use of $\boldsymbol{\theta}$ as a phase angle for complex quantities. I will also follow the convention of using j , rather than i , for the imaginary unit.

form:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{z}, \mathbf{p}, t), \\ \mathbf{0} &= \mathbf{g}(\mathbf{x}, \mathbf{z}, \mathbf{p}, t), \\ \mathbf{y} &= \mathbf{h}(\mathbf{x}, \mathbf{z}, \mathbf{p}, t),\end{aligned}\tag{2.10}$$

where both T_x and \mathbf{u}_0 have now been absorbed into the functions \mathbf{f} , \mathbf{g} , and \mathbf{h} .

2.2.2 Power flow analysis

In an AC system, the voltage V and current I are sinusoidally-varying quantities, which may or may not be in phase. Consequently, they are typically represented as complex quantities known as *phasors*, having both magnitude and angle, e.g.,

$$V = |V|e^{j\theta}.\tag{2.11}$$

Current through and voltage difference across an electrical line are related through the complex *impedance* Z :

$$\Delta V = ZI.\tag{2.12}$$

Impedance $Z = R + jX$ generalizes the resistance R of an electrical line through the addition of an imaginary component X called *reactance*, which indicates the tendency of the line to cause the voltage and current to be out of phase. The inverse of the impedance is the *admittance* $Y = 1/Z$, which may be split into real and imaginary parts as $Y = G + jB$, where $G = R/(R^2 + X^2)$ is the *conductance* and $B = -X/(R^2 + X^2)$ is the *susceptance* of the line.

The equations used for power flow analysis begin with conservation of current at each bus of the network:

$$I_{\text{inj},i} = \sum_{k=1}^{N_{\text{bus}}} I_{ik},\tag{2.13}$$

where $I_{\text{inj},i}$ is the total current injected into Bus i (e.g. by devices connected to Bus i), N_{bus} is the number of buses in the network, and I_{ik} is the current flowing from Bus i to Bus k . The latter can be

expressed in terms of the voltages V_i and V_k at each bus and the *admittance* Y_{ik} of the line between them:

$$I_{ik} = Y_{ik}(V_i - V_k) \quad (2.14)$$

One can substitute (2.14) into (2.13) and rewrite it to obtain

$$\begin{aligned} I_{\text{inj},i} &= \sum_{k=1}^{N_{\text{bus}}} Y_{ik}(V_i - V_k) \\ &= Y_{i1}(V_i - V_1) + Y_{i2}(V_i - V_2) + \dots \\ &= -Y_{i1}V_1 - Y_{i2}V_2 + \dots + (Y_{i1} + Y_{i2} + \dots)V_i + \dots \\ &= \begin{bmatrix} -Y_{i1} & -Y_{i2} & \dots & \sum_k Y_{ik} & \dots \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_i \\ \vdots \end{bmatrix} \end{aligned} \quad (2.15)$$

Defining the vectors $\mathbf{I}_{\text{inj}} \equiv \begin{bmatrix} I_{\text{inj},1} & I_{\text{inj},2} & \dots \end{bmatrix}^T$ and $\mathbf{V} \equiv \begin{bmatrix} V_1 & V_2 & \dots \end{bmatrix}^T$ and the matrix

$$\mathbf{Y}_{\text{bus}} \equiv \begin{bmatrix} \sum_k Y_{1k} & -Y_{12} & \dots & -Y_{1i} & \dots \\ -Y_{21} & \sum_k Y_{2k} & \dots & -Y_{2i} & \dots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ -Y_{i1} & -Y_{i2} & \dots & \sum_k Y_{ik} & \dots \\ \vdots & \vdots & \ddots & \vdots & \ddots \end{bmatrix}, \quad (2.16)$$

one can summarize (2.15) in the single matrix equation

$$\mathbf{I}_{\text{inj}} = \mathbf{Y}_{\text{bus}} \mathbf{V}. \quad (2.17)$$

The equations embodied in (2.17) are known as the *current balance* equations.

A related set of equations can be obtained for the (complex) power

$$S_{\text{inj},i} = V_i I_{\text{inj},i}^* \quad (2.18)$$

injected at each bus. Defining $\mathbf{S}_{\text{inj}} \equiv \begin{bmatrix} S_{\text{inj},1} & S_{\text{inj},2} & \dots \end{bmatrix}^T$, (2.18) can be expressed in matrix form as

$$\begin{aligned} \mathbf{S}_{\text{inj}} &= \text{diag}(\mathbf{V}) \mathbf{I}_{\text{inj}}^* \\ &= \text{diag}(\mathbf{V}) \mathbf{Y}_{\text{bus}}^* \mathbf{V}^*. \end{aligned} \quad (2.19)$$

It is often more practical to use these equations in real (rather than complex) form. Letting $S_{\text{inj},i} = P_{\text{inj},i} + jQ_{\text{inj},i}$ (P and Q are *real* and *reactive* power, respectively), $V_i = |V_i|e^{j\theta_i}$, and $Y_{\text{bus},ik} = G_{\text{bus},ik} + jB_{\text{bus},ik}$, the real part of (2.18) is

$$0 = -P_{\text{inj},i} + \sum_k |V_i||V_k|(G_{\text{bus},ik} \cos(\theta_i - \theta_k) + B_{\text{bus},ik} \sin(\theta_i - \theta_k)) \quad (2.20)$$

and the imaginary part is

$$0 = -Q_{\text{inj},i} + \sum_k |V_i||V_k|(G_{\text{bus},ik} \sin(\theta_i - \theta_k) - B_{\text{bus},ik} \cos(\theta_i - \theta_k)). \quad (2.21)$$

Equations (2.20) and (2.21) constitute the *power balance* equations.

In power flow analysis, the goal is to find values of $P_{\text{inj},i}$, $Q_{\text{inj},i}$, $|V_i|$, and θ_i that satisfy (2.20) and (2.21). Given that there are two equations for each bus, two of these four variables must be specified at each bus to obtain a consistent system. Loads are specified in terms of real $P_{l,i}$ and reactive $Q_{l,i}$ power consumed (*PQ loads*). The voltage magnitude $|V_i|$ is specified on all buses with generators, and all generators except one produce a specified amount of real power $P_{g,i}$ (*PV generators*). The real power on the remaining generator (called the *swing* or *slack generator*) cannot be specified in advance; it must be free to vary to account for losses in transmission lines. In fact, because of conservation of energy, (2.20) and (2.21) overconstrain the injected power (one equation is redundant) while the voltages remain underconstrained because no reference angle is specified. Accordingly, (2.20) is omitted for the slack generator and instead the voltage angle on the slack bus is specified.

2.2.3 Transient stability analysis

In transient stability analysis, the power flow equations are supplemented with equations for the dynamics of the various devices in the system. It is usually assumed that the dynamics of power flow are fast compared to the transients being studied, so the power flow equations form part of the algebraic relations $\mathbf{0} = \mathbf{g}(\mathbf{x}, \mathbf{z}, \mathbf{p}, t)$ in (2.10). Relevant model parameters that come from the transmission network usually include line conductances G_{ik} and susceptances B_{ik} .

Generators come in various types, including synchronous generators (SGs; used to model conventional power plants), doubly-fed induction generators (DFIGs; used to model certain types of wind power), and direct-drive synchronous generators (DDSGs; used for solar and other types of wind). In addition, each generator is often connected to other devices that regulate its output (such as *turbine governors* (TGs) for regulating the input mechanical torque and providing speed/frequency control, and *automatic voltage regulators* (AVRs) for controlling the output voltage). This review will focus on SGs and introduce some of the basic equations and parameters. A synchronous generator consists of a rotating component, appropriately called the *rotor*, and a stationary component, called the *stator*. Both components have electrical windings which carry current during operation of the generator. The windings on the rotor carry a DC current which produces a magnetic field by Ampere's law. As the rotor spins, the rotating magnetic field induces an AC voltage in the stator windings by Faraday's law. These stator windings are usually arranged so that the voltages produced are 120 degrees apart, producing three symmetrical *phases* of AC voltage. The basic equations for a synchronous generator describe the relationships between these various fields, voltages, and currents. More sophisticated models include the effects of dampening, eddy currents in the rotor, magnetic saturation, etc.

It is common practice in power systems modeling to consider all quantities from a rotating frame of reference in which sinusoidally-varying quantities appear constant in time. This is accomplished via a transformation known as the $dq0$ transformation or Park's transformation (for

details, see [41, 43]). This refers to the *direct* axis, which is aligned with the north pole of the rotating magnetic field; the *quadrature* axis, which is 90 degrees ahead of the direct axis; and the *homopolar* axis, perpendicular to both. In addition, most quantities are scaled by a set of base values using a dimensionless “per-unit” (pu) system in order to further simplify the equations.

The following equations come from applying Newton’s second law to the motion of the rotor:

$$\dot{\delta} = \Omega_b(\omega - \omega_s) \quad (2.22)$$

$$\dot{\omega} = \frac{1}{2H}(\tau_m - \tau_e - D(\omega - \omega_s)); \quad (2.23)$$

where δ and ω are the rotor angle in rad and speed in rad/s, respectively; Ω_b is the base rotational frequency (e.g., 360π rad/s in a 60 Hz system); ω_s is the reference frequency in pu (usually 1); H is known as the *inertia constant*; τ_m and τ_e are the mechanical and electromagnetic torque, respectively; and D is a damping coefficient. Ω_b , ω_s , H , and D are parameters (constants). τ_m is an input which may either be set to a reference value or be determined by a turbine governor. τ_e is usually given by

$$\tau_e = (r_a i_q + v_q) i_q + (r_a i_d + v_d) i_d, \quad (2.24)$$

where r_a is the armature resistance (a parameter), i_d and i_q are machine output currents, and v_d and v_q are machine output voltages. The last two are linked to the bus voltage through the equations

$$v_d = |V| \sin(\delta - \theta) \quad (2.25)$$

$$v_q = |V| \cos(\delta - \theta), \quad (2.26)$$

whereas the power injected by the generator into the bus is

$$P_g = v_d i_d + v_q i_q \quad (2.27)$$

$$Q_g = v_q i_d - v_d i_q. \quad (2.28)$$

Equations (2.22), (2.23), and (2.25)-(2.28) are the same for all SG models. Minimally, they must be supplemented by two more algebraic relations between the machine voltages and currents

for all quantities to be well-defined. These vary depending on the *order* (number of differential variables) of the model used. The simplest model is the second-order *classical* model, which only adds the following two equations to the foregoing:

$$0 = v_q + r_a i_q - e'_q + x'_d i_d \quad (2.29)$$

$$0 = v_d + r_a i_d - x'_d i_q, \quad (2.30)$$

where e'_q is a constant emf (voltage source) and x'_d is the d -axis *transient reactance* (a parameter).

For comparison, a slightly more sophisticated (third-order) model includes the following differential equation for e'_q :

$$\dot{e}'_q = \frac{1}{T'_{d0}} (-e'_q - (x_d - x'_d) i_d + v_f), \quad (2.31)$$

where T'_{d0} is the d -axis *open-circuit transient time constant*, x_d is the d -axis *synchronous reactance*, and v_f is the *field voltage*. T'_{d0} and x_d are both parameters; v_f is an input determined either by a reference value or by an automatic voltage regulator. For this model, (2.29) and (2.30) must be replaced by

$$0 = v_q + r_a i_q - e'_q + x'_d i_d \quad (2.32)$$

$$0 = v_d + r_a i_d - x_q i_q, \quad (2.33)$$

where x_q is the q -axis synchronous reactance (another parameter).

Chapter 3

Similarity Measures in Dynamical Systems

The field of dynamical systems covers a broad range of models. Many of these can be categorized by the behavior they exhibit: transient behavior that settles into a steady state, cyclic behavior that repeats at regular intervals (i.e., periodic oscillatory motion), cyclic behavior that is irregular (e.g., chaotic oscillatory motion), and stochastic behavior, for example. These categories are more than simply convenient; they reflect fundamental differences between different types of systems (or phases of a single system).

In this chapter, I explore the differences between (nonstochastic) oscillatory and nonoscillatory sloppy models. I show that oscillatory models have characteristic features that manifest in the cost function used for parameter inference and that these features are a reflection of model manifolds that are qualitatively different than those of nonoscillatory models. Specifically, oscillatory models have manifolds with a high effective dimensionality, rather than the low effective dimensionality typical of other sloppy models. The differences also extend to the sensitivities of model predictions to changes in parameters, which scale as a function of the total observation time in ways that are characteristic of the model type.

I develop alternative metrics that can be used for oscillatory models which transform the model manifold from one of high to one of low effective dimensionality. This is an important step

toward being able to use tools such as MBAM for reduction of models of complex oscillatory processes, because the primary assumption of MBAM is that the model manifold has a low effective dimensionality that allows it to be approximated by a boundary cell of lower dimensionality without introducing significant bias.

While much of the development of ideas and writing of this work were done collaboratively with my advisor, I performed the computations, which resulted in the figures of the paper, in addition to fleshing out much of the mathematical theory.

This chapter was published as an article in the journal *Physical Review E* [44]. I hereby confirm that the use of this article is compliant with all publishing agreements.

Unwinding the model manifold: Choosing similarity measures to remove local minima in sloppy dynamical systems

Benjamin L. Francis and Mark K. Transtrum*

Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602, USA



(Received 5 June 2018; revised manuscript received 26 February 2019; published 11 July 2019)

In this paper, we consider the problem of parameter sensitivity in models of complex dynamical systems through the lens of information geometry. We calculate the sensitivity of model behavior to variations in parameters. In most cases, models are sloppy, that is, exhibit an exponential hierarchy of parameter sensitivities. We propose a parameter classification scheme based on how the sensitivities scale at long observation times. We show that for oscillatory models, either with a limit cycle or a strange attractor, sensitivities can become arbitrarily large, which implies a high effective dimensionality on the model manifold. Sloppy models with a single fixed point have model manifolds with low effective dimensionality, previously described as a “hyper-ribbon.” In contrast, models with high effective dimensionality translate into multimodal fitting problems. We define a measure of curvature on the model manifold which we call the *winding frequency* that estimates the density of local minima in the model’s parameter space. We then show how alternative choices of fitting metrics can “unwind” the model manifold and give low winding frequencies. This prescription translates the model manifold from one of high effective dimensionality into the hyper-ribbon structures observed elsewhere. This translation opens the door for applications of sloppy model analysis and model reduction methods developed for models with low effective dimensionality.

DOI: [10.1103/PhysRevE.100.012206](https://doi.org/10.1103/PhysRevE.100.012206)

I. INTRODUCTION

An essential part of the modeling process is selecting a similarity metric that quantifies the extent to which a model mimics the system or phenomenon of interest [1]. The choice of similarity metric informs nearly all aspects of the modeling process: model selection, data fitting, model reduction, experimental design, model validation, etc. Here, we consider the question of similarity metrics for dynamical systems, particularly oscillatory ones. Although a common choice, the least squares metric comparing model outputs at selected times may lead to models with a *high effective dimensionality*. In addition to posing technical challenges (e.g., ill-posed, multimodal fitting problems), we argue that a high effective dimensionality reflects a more fundamental issue: that the choice of metric does not accurately capture the phenomenon of interest. In this paper, we use sloppy model analysis and information geometry to identify parameter combinations in models of dynamical systems that lead to high effective dimensionalities (Secs. II and III). We then use methods of signal processing to construct similarity measures that “unwind” the model manifold and lead to well-posed inference problems (Sec. IV).

Some have already observed that one’s choice of metric is a critical aspect of parameter space exploration [2,3]. The relationship between model behavior and parameters is (locally) captured by sensitivity analysis. Previous studies have decomposed the sensitivities of periodic signals into independent parts that control amplitude, period, and other

features [4–6]. In chaotic systems, it has been found that the dynamics exhibit an exponential sensitivity to parameters [7,8]. In such cases, it is common to use measures of the statistical distribution in phase space, rather than time series [3,7,9–11]. The present work combines these insights with tools of sloppy model analysis and information geometry.

Sloppy models are a broad class of models whose behavior exhibits an exponential hierarchy of parameter sensitivities [12–20]. Using an information geometric approach, it has been shown that the local sensitivity analysis reflects a global property, i.e., a low effective dimensionality described as a *hyper-ribbon* [21–24]. It has been suggested that this hyper-ribbon structure is why simple effective (i.e., low-dimensional) theories of collective behaviors exist for systems that are microscopically complicated [23,24].

The effective dimensionality of sloppy models has important statistical implications. Information criteria (such as Akaike or Bayes) are used in model selection and penalize those with too much fitting flexibility. A model’s fitting power is most easily estimated in the asymptotic limit, in which it is simply approximated by the number of parameters, i.e., the dimension of the model manifold. For hyper-ribbons, these formulas greatly overestimate the fitting power of a model [25,26]. However, it is also possible for models to exhibit a *high effective dimensionality*, i.e., have model manifolds whose fitting power is much larger than that suggested by the number of parameters. As we show in Secs. II and III, these models will exhibit extreme multimodality when fit to data, and have model manifolds with large curvatures that tend to fill large volumes of behavior space.

The challenge of multimodality in fitting problems has been noted in many fields [9,10,27–29]. Proposals for

*mktranstrum@byu.edu

addressing multimodality have included global search methods [9,28,30–32], increasing the size of the parameter space in order to escape local minima [29,33], and changing the parameter landscape through an alternative choice of metric [3,9].

In Sec. II, we introduce the least squares metric under consideration and use model sensitivity analysis at long times to classify parameter combinations. In turn, we classify models based on which parameter types they include and show that some classes of models exhibit an anomalous statistical dimension, that is, the effective dimensionality of the model may be either much more or less than the number of parameters. In Sec. III, we argue for a deeper theoretical implication of this phenomenon. Using an information geometric approach, we relate the effective statistical dimension to the curvature on the model manifold. In Sec. IV, we explicitly demonstrate that alternative metrics can lead to different effective dimensions and present a prescription for how models of high effective dimension can be regularized through an appropriate choice of metric.

II. MODEL AND PARAMETER CLASSIFICATIONS

A. Similarity measure

Consider a parametrized model of time $y(t; \theta)$ (which could be generated, for example, as the solution to a system of differential equations), where θ is a vector of parameters (which could include initial conditions) and y is either a scalar or vector of observables. We wish to quantify the similarity of the model behavior for different values of θ . The most common metric in the literature is least squares regression, in which case the distance (or cost) between two models, with parameters θ and θ_0 , takes the form

$$C(\theta) = \frac{1}{2T} \int_0^T dt [\delta y(t; \theta)]^2, \quad (1)$$

$$\delta y(t; \theta) \equiv y(t; \theta_0) - y(t; \theta). \quad (1a)$$

We are interested in the sensitivity of model predictions at different time scales. By increasing the total time T , this cost function $C(\theta)$ defines a coarse-graining in the effective sampling rate followed by a renormalization so that the total number of effective data points is constant. When measuring the distance to observed data y_i at times t_i (with uncertainties σ_i used as weights), the integral becomes a sum,

$$C(\theta) = \frac{1}{2T} \sum_i \left(\frac{y_i - y(t_i; \theta)}{\sigma_i} \right)^2. \quad (2)$$

Being a distance measure, C defines a metric on the space of data and model predictions known as *data space* [21,22]. We interpret the model predictions $y(t_i; \theta)$ and observations y_i as components of two vectors in data space which we denote $\mathbf{y}(\theta)$ and \mathbf{y} , respectively. By varying θ , $\mathbf{y}(\theta)$ sweeps out a surface in data space known as the *model manifold*. With this

notation, Eq. (2) may be written as

$$C(\theta) = \frac{1}{2T} \delta \mathbf{y}^\top \Sigma^{-1} \delta \mathbf{y}, \quad (3)$$

$$\delta \mathbf{y} \equiv \mathbf{y} - \mathbf{y}(\theta), \quad (3a)$$

where Σ denotes the (diagonal) covariance matrix for the observation vector \mathbf{y} .

B. Sensitivity analysis and parameter classification

To quantify the sensitivity to parameters of model predictions at different time scales, we consider derivatives of the cost with respect to θ . Dropping the t and θ dependence for clarity, the gradient of Eq. (1) is

$$\frac{\partial C}{\partial \theta_\mu} = -\frac{1}{T} \int_0^T dt \left(\delta y \frac{\partial y}{\partial \theta_\mu} \right), \quad (4)$$

and the Hessian is

$$H_{\mu\nu} \equiv \frac{\partial^2 C}{\partial \theta_\mu \partial \theta_\nu} = \frac{1}{T} \int_0^T dt \left(\frac{\partial y}{\partial \theta_\mu} \frac{\partial y}{\partial \theta_\nu} - \delta y \frac{\partial^2 y}{\partial \theta_\mu \partial \theta_\nu} \right). \quad (5)$$

Note that because $\delta y(t; \theta_0) = 0$, the gradient at θ_0 is also 0 and the Hessian at θ_0 simplifies to

$$H_{\mu\nu}(\theta_0) = \frac{1}{T} \int_0^T dt \left(\frac{\partial y}{\partial \theta_\mu} \frac{\partial y}{\partial \theta_\nu} \right). \quad (6)$$

This is also approximately valid when $\theta \approx \theta_0$. For Eq. (3), the Hessian at θ_0 takes the form

$$H(\theta_0) = \frac{1}{T} \frac{\partial \mathbf{y}^\top}{\partial \theta} \Sigma^{-1} \frac{\partial \mathbf{y}}{\partial \theta}. \quad (7)$$

Although the gradient and Hessian may be evaluated at other points, $H(\theta_0)$ is particularly important because it is the Fisher information metric (FIM) for this measurement process and acts as a Riemannian metric on the model manifold. We are interested in the eigenvalues of H and their dependence on T .

Figure 1 plots a cross section of C (as a surface over θ), the eigenvalues of H , and a three-dimensional projection of the model manifold for three models (details of these models are found in Appendix A).

The first model is characterized by a transient decay to a steady state. As illustrated in Fig. 1(b), for large T , the model becomes increasingly insensitive to parameter combinations that control transient behavior, scaling as $O(T^{-1})$. The parameter that determines the steady state scales as $O(1)$. These scaling behaviors can be motivated as follows. We assume that parameter combinations which control the transient dynamics have sensitivities that decay to zero at long times,

$$\frac{\partial y}{\partial \theta_\mu}(t \rightarrow \infty; \theta) \sim 0, \quad (8)$$

while those that control the steady state are asymptotically constant:

$$\frac{\partial y}{\partial \theta_\mu}(t \rightarrow \infty; \theta) \sim \text{const.} \quad (9)$$

In light of Eq. (6), this leads to the $O(T^{-1})$ and $O(1)$ scaling behaviors observed. Note that as the total sampling time T is increased past the transient dynamics, the only new information obtained is information about the final steady state. Our

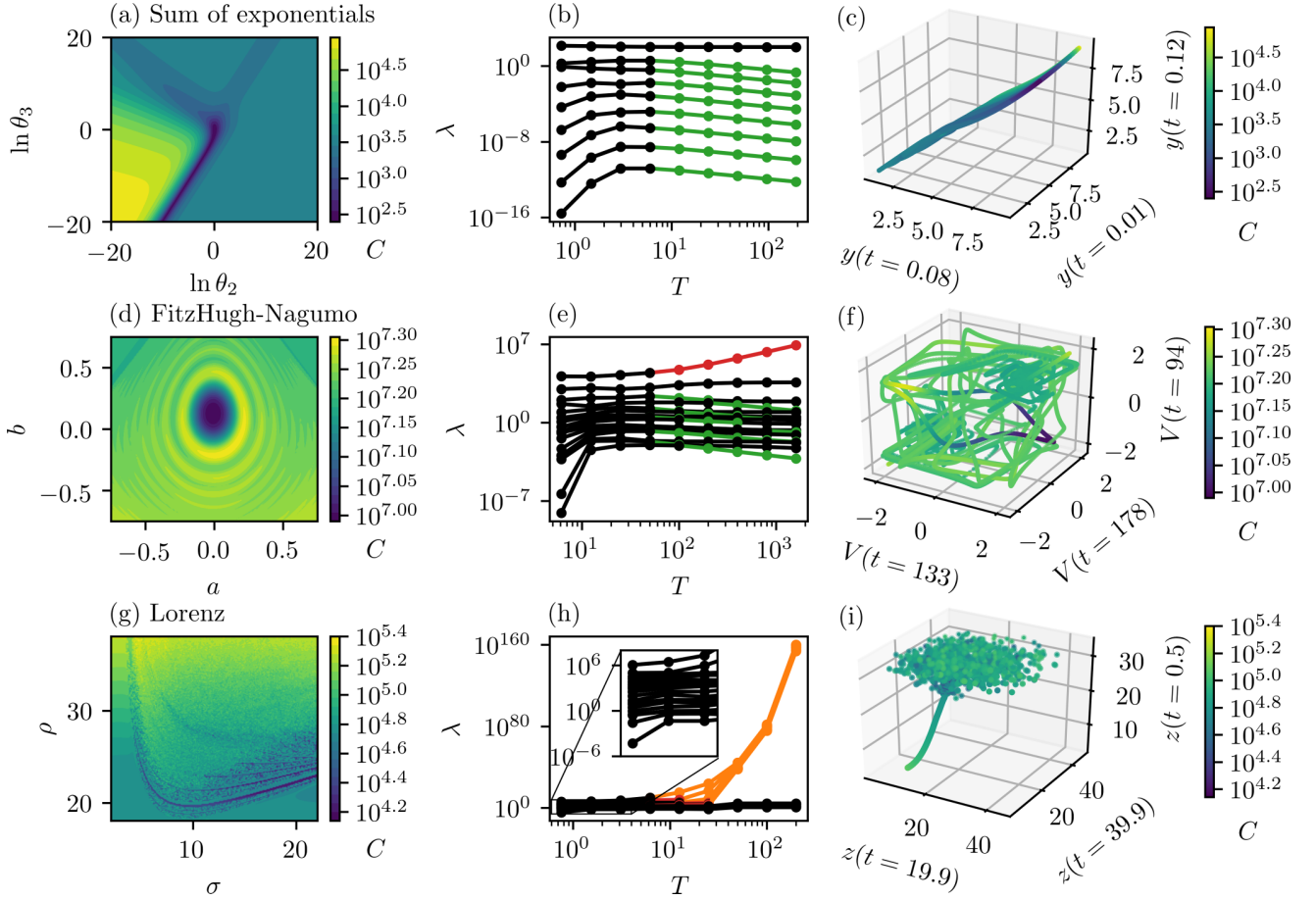


FIG. 1. Model classes. [(a), (d), (g)]: Cross sections of $C(\theta)$ [Eq. (1)] for three prototype models (see Appendix A, models A, D, and G). Contrast the canyons in (a) with the ripples in (d) and the roughness in (g). [(b), (e), (h)]: Hessian eigenvalues as a function of sampling time (same models). Colors differentiate scaling behaviors at long times. [(c), (f), (i)]: Projections of the model manifold (same models). In (c), a three-ball in parameter space was mapped to the nearly one-dimensional region of prediction space shown (low effective dimensionality). By contrast, for (f) and (i) a single parameter was varied producing a one-dimensional (1D) (space-filling) curve in prediction space (high effective dimensionality). Note that in (i), the model goes through a bifurcation where the manifold begins to oscillate rapidly. The sampling required to see continuity is prohibitive, so the points plotted become scattered.

choice of normalization keeps the effective number of data points constant, so increasing T results in an effective loss of information about the transient dynamics but no information loss for the steady state.

The second model exhibits a periodic limit cycle. As shown in Fig. 1(e), parameter combinations controlling features of the attractor scale as $O(1)$, those that control the transient approach to the attractor scale as $O(T^{-1})$, and the combination controlling frequency scales as $O(T^2)$. Motivation for the scaling behavior of the parameter combinations controlling the transient approach to the attractor follow as in the previous case. To motivate the other two scaling behaviors, we consider the steady state of the model and expand in a Fourier series:

$$y(t \rightarrow \infty; \theta) = \sum_{k=-\infty}^{\infty} c_k(\theta) e^{ik\omega(\theta)t} = y(t \rightarrow \infty; c(\theta), \omega(\theta)). \quad (10)$$

There is an intermediate dependence of the steady state on the amplitude coefficients c_k and the oscillatory frequency ω .

This allows us to decompose the parameter sensitivities of the steady state into two parts:

$$\frac{\partial y}{\partial \theta_\mu}(t \rightarrow \infty; \theta) = \sum_{k=-\infty}^{\infty} \frac{\partial y}{\partial c_k} \frac{\partial c_k}{\partial \theta_\mu} + \frac{\partial y}{\partial \omega} \frac{\partial \omega}{\partial \theta_\mu}. \quad (11)$$

Because c_k and ω are time independent by construction, the time dependence of these sensitivities is due entirely to the coefficients

$$\frac{\partial y}{\partial c_k}(t \rightarrow \infty; \theta) = e^{ik\omega(\theta)t}, \quad (12)$$

which is bounded by a constant, and

$$\frac{\partial y}{\partial \omega}(t \rightarrow \infty; \theta) = \sum_{k=-\infty}^{\infty} ikt c_k(\theta) e^{ik\omega(\theta)t} \sim t, \quad (13)$$

which grows linearly with time. The amplitude sensitivities $(\partial y / \partial c_k)(\partial c_k / \partial \theta_\mu)$ control the shape and amplitude of the steady state and give rise to $O(1)$ scaling behavior [referring

TABLE I. Parameter classification.

Eigenvalue scaling behavior	Dynamics controlled
$O(T^{-1})$	transient
$O(1)$	steady state
$O(T^2)$	frequency
$O(e^T)$	chaotic behavior

again to Eq. (6)]. By contrast, the frequency sensitivity $(\partial y / \partial \omega)(\partial \omega / \partial \theta_\mu)$ results in $O(T^2)$ scaling behavior. Other studies have focused on the sensitivity to period, rather than frequency, but the temporal scaling behavior is the same for both [4–6].

Finally, the third model is chaotic; parameters controlling its dynamics exhibit exponential sensitivities,

$$\frac{\partial y}{\partial \theta_\mu}(t \rightarrow \infty; \theta) \sim e^{\lambda_\mu t}, \quad (14)$$

leading to the exponential scaling behavior in Fig. 1(h).

We classify parameter combinations in a model according to their scaling behavior. This classification is summarized in Table I. LaMont and Wiggins have also proposed a classification of model parameters, based on the *complexity* of a given parameter combination [25]. In the case of dynamical models, our analysis illustrates the mechanisms that give rise to the complexities of each class.

C. Model classification

The different scaling behaviors for the Hessian eigenvalues are accompanied by different structures in both the cost surface and the model manifold (first and third columns of Fig. 1). The cost surface of the first model is characterized by a single, highly anisotropic basin. Its model manifold is similarly anisotropic; the long, narrow hyper-ribbon structure is common for models with low effective dimensionality [21,22]. In contrast, the second cost surface has many local minima and a model manifold with high curvature. The third cost surface exhibits a fractal-like roughness (although for finite T the derivative with respect to parameters formally exists everywhere). Its model manifold is even more highly curved and space filling.

These three models are prototypes of three model classes, distinguished by the scaling behavior of the largest eigenvalue for large T . For the first class, $\lambda_{\max} \sim O(1)$ is bound by a constant. For the second class, $\lambda_{\max} \sim O(T^n)$ is bound by a polynomial. For the third class, $\lambda_{\max} \sim O(e^T)$ is exponential. We plot the eigenvalues of the Hessian (at large, fixed T) for the three prototype models and for two additional models from each class in Fig. 2 (details of these models are found in Appendix A). All nine models are considered *sloppy*; that is, the eigenvalues of the Hessian are spread over many orders of magnitude. Accordingly, we refer to these model classes as *sloppy models of the first, second, and third kinds*, respectively.

III. MANIFOLD CURVATURE

The large sensitivities of sloppy models of the second and third kinds are necessarily associated with large curvature and

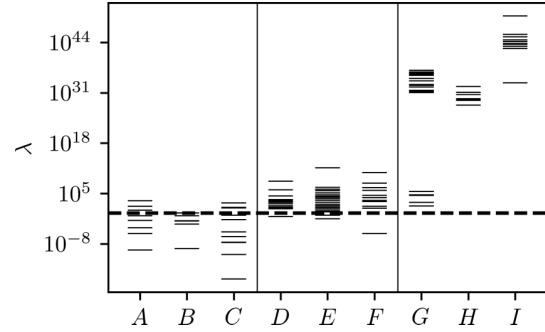


FIG. 2. Eigenvalues of $H(\theta_0)$ for the following models (see Appendix A for details): A: sum of exponentials; B: rational polynomial; C: biological adaptation; D: FitzHugh-Nagumo; E: Hodgkin-Huxley; F: Wnt oscillator; G: Lorenz; H: Hindmarsh-Rose; I: damped, driven pendulum. $\{A, B, C\}$ are nonoscillatory models, $\{D, E, F\}$ are periodic, and $\{G, H, I\}$ are chaotic.

high effective dimensionality on the model manifold. This can be understood by noting that the absolute variation in the model behavior is bounded (the models oscillate within a finite range and do not grow). This restricts the model manifold to a finite region of data space. Large parameter sensitivities indicate that the model manifold is very long in the associated parameter directions. The only way to fit something very long into a finite region is for it to curve, fold, or wind. The combination of large parameter sensitivities and bounded predicted behavior necessitates large manifold curvature. For large T , there will be enough winding that the manifold effectively fills a volume of higher dimension than that of the manifold itself. This high effective dimensionality is the opposite effect of the low effective dimensionality argued for in previous studies of sloppy models [23,24].

To quantify this effect, we introduce a quantity that we call the *winding frequency*, as follows. The extrinsic curvature associated with parameter direction v is given by the geodesic curvature $k(v) = 1/R$ (as in Ref. [22]), where R is the radius of curvature of the circle tangent to the manifold along direction $(\partial y / \partial \theta^\mu) v^\mu$ (sum over μ implied; see Fig. 3). We define

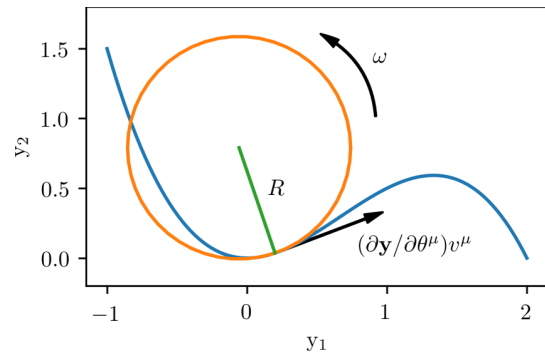


FIG. 3. Illustration of winding frequency. The “s”-shaped curve represents a possible 1D cross section of a model manifold (obtained, for example, by varying just one parameter combination) in a simple 2D data space. Also shown are the tangent or velocity vector $(\partial y / \partial \theta^\mu) v^\mu$ (sum over μ implied), the tangent circle with radius R , and the winding frequency ω defined in Eq. (15).

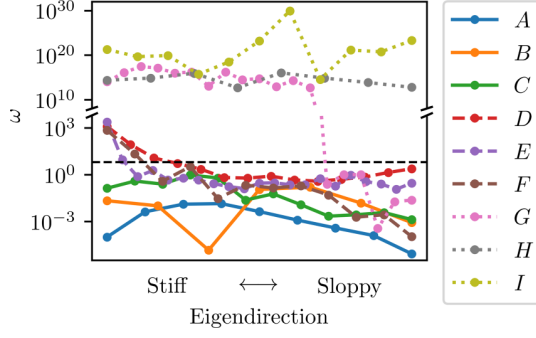


FIG. 4. Winding frequencies along Hessian eigendirections for the models from Fig. 2, ordered from left to right by magnitude of the corresponding eigenvalue. “Stiff” refers to eigendirections with large eigenvalues, while “sloppy” refers to eigendirections with small eigenvalues. The black dashed line at $\omega = 2\pi$ roughly distinguishes low from high winding frequencies.

the winding frequency as

$$\omega(v) \equiv \left| \frac{\partial \mathbf{y}}{\partial \theta^\mu} v^\mu \right| k(v), \quad (15)$$

which is the angular velocity at which the manifold locally winds around the tangent circle, such that $f = \omega/2\pi$ is the number of windings of the manifold per unit change in parameters. Because C is a distance measure for the manifold embedding space, each winding of the manifold results in a local minimum of C , so f is also the frequency of local minima in C as we move along parameter direction v .

Figure 4 shows winding frequencies along Hessian eigendirections for the models from Fig. 2. Notice that sloppy models of the first kind (i.e., hyper-ribbons) have low winding frequencies. Sloppy models of the second kind have high winding frequency in the stiffest direction, which controls frequency. Sloppy models of the third kind have high winding frequencies in more than one direction.

The effective dimensionality, estimated by the winding frequencies, depends on the metric of the model manifold embedding space, i.e., Eq. (1). We now show that alternative choices for embedding the model manifold can lead to different effective dimensionalities.

IV. ALTERNATIVE METRICS

A. Analytic signal (AS)

The high effective dimensionality of sloppy models of the second kind is due entirely to the parameter combination controlling frequency. Varying this parameter combination causes model predictions to pass in and out of phase with each other, resulting in local minima in the cost (see Figs. 5 and 6). We avoid this aliasing by defining the phase of oscillation as a monotonically increasing function of time and comparing model behaviors at the same phase.

Many definitions of instantaneous frequency and phase have been considered in the literature [34–36]. We use the *analytic signal* approach [37], which is discussed in Sec. IV A 1. Some alternatives are discussed in Appendix B. We propose a metric for oscillatory systems in Sec. IV A 2. We discuss the

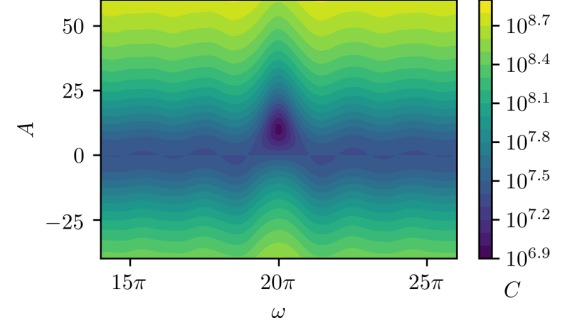


FIG. 5. Cost C for the model $y(t) = A \cos(\omega t)$, treating A and ω as parameters. The cost has been rescaled to make the local minima apparent.

Hessian and FIM in Sec. IV A 3. Results of applying our metric to the FitzHugh-Nagumo model are found in Sec. IV A 4. Comparing model predictions with observational data in this paradigm requires calculating the phases of the observations, which will have uncertainty. We propagate uncertainty and derive appropriate covariance matrices in Appendix C. Calculation of winding frequencies requires second-order parameter sensitivities (specifically, when calculating the geodesic curvature κ); we derive the necessary formulas in Appendix D.

1. Phase definition

The analytic representation $z(t)$ of an oscillatory signal $y(t)$ is a complex function defined as

$$z(t) \equiv y(t) + iH[y](t) = A(t)e^{i\phi(t)}, \quad (16)$$

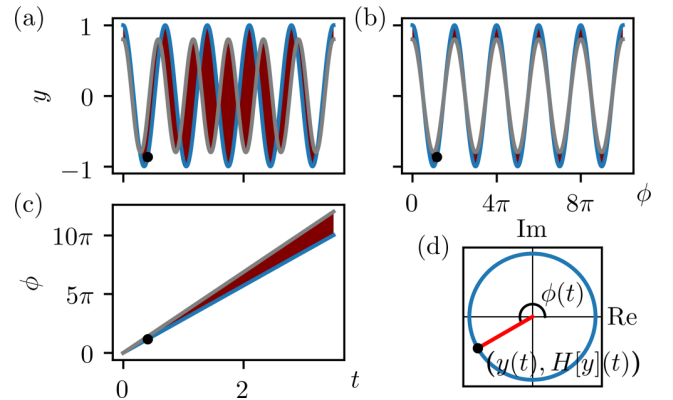


FIG. 6. Decoupling amplitude from phase. (a) Signal vs time for two signals with mismatched amplitude and frequency; their difference is indicated by the shading between the curves. The mismatch in frequency causes a large difference δy when the two signals are out of phase ($t \approx 1.8$) but little or no difference when they are in phase ($t \approx 3.5$). (b) Signal vs phase for the same signals. The difference-at-constant-phase $\delta \tilde{y}$ is consistent throughout (see Sec. IV A 2). (c) Phase vs time for the two signals. The difference in phase $\delta \phi$ simply grows linearly (see Sec. IV A 2). (d) Analytic representation in the complex plane of the black point marked in the other three panels (see Sec. IV A 1).

where $H[y](t)$ is the Hilbert transform of $y(t)$,

$$H[y](t) \equiv \frac{1}{\pi} \text{P. V.} \int_{-\infty}^{\infty} \frac{y(\tau)}{t - \tau} d\tau, \quad (17)$$

and the magnitude $A(t)$ and argument $\phi(t)$ of $z(t)$ are

$$A(t) \equiv \sqrt{y^2(t) + H^2[y](t)}, \quad (18)$$

$$\phi(t) \equiv \tan^{-1} \left(\frac{H[y](t)}{y(t)} \right) \quad (19)$$

[see Fig. 6(d)].

In light of Eq. (16), we reinterpret $y(t)$ in terms of amplitude and phase as

$$y(t) = \text{Re}\{z(t)\} = A(t) \cos(\phi(t)). \quad (20)$$

We then define a new signal \tilde{y} as a function of phase:

$$\tilde{y}(\phi(t)) \equiv y(t) = A(t) \cos(\phi(t)). \quad (21)$$

As long as $\phi(t)$ is monotonically increasing, the relationship between ϕ and t is invertible. Hence, we may also write

$$\tilde{y}(\phi) = A(t(\phi)) \cos(\phi). \quad (22)$$

If $y(t)$ is a vector (rather than scalar) function of time, then $\phi(t)$ will also be a vector function of time. That is, for each scalar component of $y(t)$, the preceding prescription for constructing the phase should be applied separately. If this is not possible or does not produce a set of monotonically increasing phases, it may be applied to a single scalar component of y and the resulting phase used for all of the components. For other alternatives that avoid using the Hilbert transform, see Appendix B.

As a final note, a necessary condition for $\phi(t)$ to be monotonically increasing is that the signal $y(t)$ oscillate around 0. If it does not, the time average $\langle y(t) \rangle = (1/T) \int_0^T y(t) dt$ should be subtracted from $y(t)$ prior to calculating the phase. $H[y](t)$ will be unaffected, as the Hilbert transform of a constant is 0.

2. New cost using phase

We want to construct a cost that compares models at the same phase rather than the same time. Actually, we can go one step further and construct a cost that also retains the phase information while still eliminating the aliasing of oscillations that results in local minima. We use an approximation of Eq. (1a) that arises from the propagation of uncertainty considered in Appendix C [see Eq. (C16)]. In the discrete case [comparing a model $y(t_i; \theta)$ with observational data y_i], Eq. (1a) is

$$\delta y_i \equiv y_i - y(t_i; \theta). \quad (23)$$

We define the deviations of the phases ϕ_i of the observations from the phases $\phi(t_i; \theta)$ predicted by the model as

$$\delta \phi_i \equiv \phi_i - \phi(t_i; \theta) \quad (24)$$

[see Fig. 6(c)], and the deviations of the observations from the predictions at constant phase as

$$\delta \tilde{y}_i \equiv y_i - \tilde{y}(\phi_i; \theta) \quad (25)$$

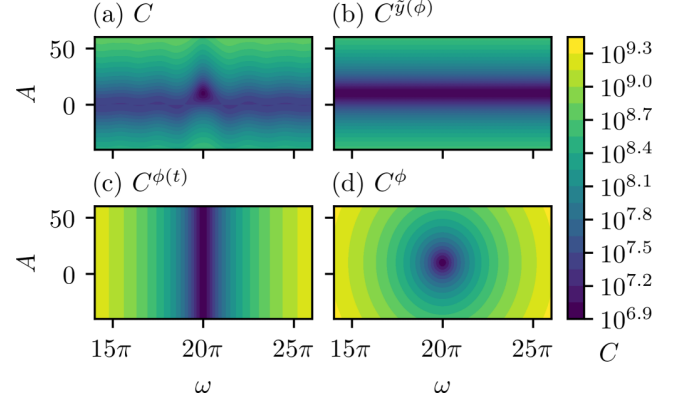


FIG. 7. Cost decomposition of the model $y(t) = A \cos(\omega t)$. (a) Same as Fig. 5. (b) $\tilde{y}(\phi) = A \cos(\phi)$ is insensitive to changes in ω and varies linearly with A , resulting in a quadratic dependence of $C^{\tilde{y}(\phi)}$ on A only. (c) $\phi(t) = \omega t$ is insensitive to changes in A and varies linearly with ω , resulting in a quadratic dependence of $C^{\phi(t)}$ on ω only. (d) Cost using Eq. (27). The ripples in (a) have been replaced with a quadratic basin.

[see Fig. 6(b)]. The approximation we use for oscillatory systems is

$$\delta y_i \approx \delta \tilde{y}_i + \left(\frac{\partial \tilde{y}}{\partial \phi} \right)_i \delta \phi_i \equiv \delta \hat{y}_i. \quad (26)$$

The first term captures changes in amplitude while the second term captures changes in phase or frequency, so both pieces of information are retained (see Sec. IV A 3). At the same time, because this approximation is first order in $\delta \phi_i$, it eliminates the nonlinear dependence on frequency that results in ripples in the cost (refer back to Fig. 6), which we will demonstrate in Figs. 7 and 8.

We define a new cost function by replacing $\delta \mathbf{y}$ in Eq. (3) with the approximation $\delta \hat{\mathbf{y}}$ defined according to Eq. (26):

$$C^\phi(\theta) \equiv \frac{1}{2T} \delta \hat{\mathbf{y}}^\top \Sigma^{-1} \delta \hat{\mathbf{y}}. \quad (27)$$

Using Eq. (26), this may be decomposed into three pieces representing the amplitude contribution, the phase contribution, and a cross term:

$$C^\phi(\theta) = C^{\tilde{y}(\phi)}(\theta) + C^{\phi(t)}(\theta) + C^X(\theta), \quad (27a)$$

$$C^{\tilde{y}(\phi)}(\theta) \equiv \frac{1}{2T} \delta \tilde{\mathbf{y}}^\top \Sigma^{-1} \delta \tilde{\mathbf{y}}, \quad (27b)$$

$$C^{\phi(t)}(\theta) \equiv \frac{1}{2T} \Delta \phi^\top \Sigma^{-1} \Delta \phi, \quad (27c)$$

$$C^X(\theta) \equiv \frac{1}{2T} (\delta \tilde{\mathbf{y}}^\top \Sigma^{-1} \Delta \phi + \Delta \phi^\top \Sigma^{-1} \delta \tilde{\mathbf{y}}), \quad (27d)$$

$$\Delta \phi_i \equiv \left(\frac{\partial \tilde{y}}{\partial \phi} \right)_i \delta \phi_i. \quad (27e)$$

We compare C [Eq. (3)], $C^{\tilde{y}(\phi)}$, $C^{\phi(t)}$, and C^ϕ in Fig. 7.

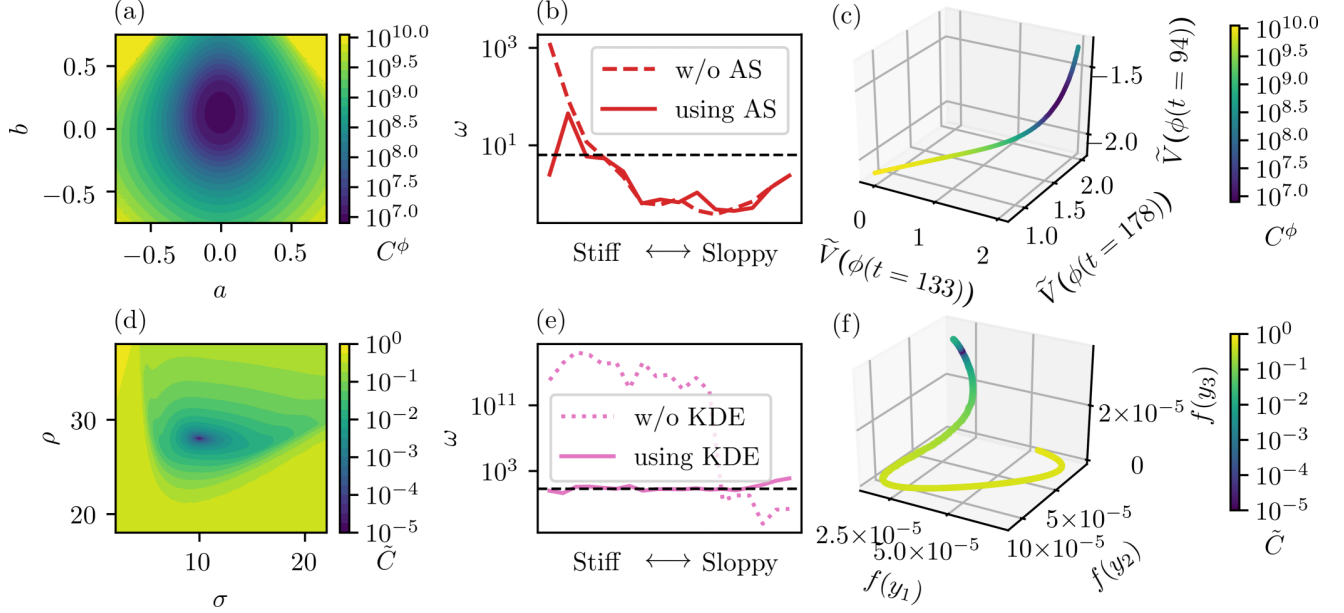


FIG. 8. Effects of alternative metrics on cost surfaces (first column), winding frequencies (second column), and manifolds (third column), (a)–(c) using analytic signal (AS) and (d)–(f) using kernel density estimation (KDE). Compare (a), (c), (d), and (f) with Figs. 1(d), 1(f), 1(g), and 1(i), respectively. Panels (b) and (e) show both the winding frequencies shown previously in Fig. 4 (“w/o ____”) and the winding frequencies that result when using our metrics (“using ____”) for comparison. Note that long time series are needed to achieve these results; see Appendix E for details.

When comparing two models with parameters θ_0 and θ , Eqs. (24)–(27) take the form

$$\delta\phi(t; \theta) \equiv \phi(t; \theta_0) - \phi(t; \theta), \quad (28)$$

$$\begin{aligned} \delta\tilde{y}(t; \theta) &\equiv \tilde{y}(\phi(t; \theta_0); \theta_0) - \tilde{y}(\phi(t; \theta); \theta) \\ &= y(t; \theta_0) - \tilde{y}(\phi(t; \theta); \theta), \end{aligned} \quad (29)$$

$$\delta\hat{y}(t; \theta) \equiv \delta\tilde{y}(t; \theta) + \frac{\partial\tilde{y}(\phi(t; \theta_0); \theta_0)}{\partial\phi} \delta\phi(t, \theta), \quad (30)$$

$$C^\phi(\theta) \equiv \frac{1}{2T} \int_0^T dt [\delta\hat{y}(t; \theta)]^2. \quad (31)$$

As we show in Sec. IV A 3, Eq. (31) is a quadratic approximation of Eq. (1) (i.e., they have the same gradient and Hessian). In other words, Eq. (31) is an isometric embedding of the model manifold. However, because changes in frequency only affect $\phi(t; \theta)$, which is unbounded, the large manifold volume is no longer constrained to a finite region of the embedding space.

3. Fisher information metric

We stated in Sec. II B that the Hessian of the cost evaluated at θ_0 is the Fisher information metric (FIM). Specifically, the FIM is related to the cost by

$$\mathcal{I}_{\mu\nu} = \left\langle \frac{\partial^2 C(\theta_0)}{\partial\theta_\mu \partial\theta_\nu} \right\rangle = \langle H_{\mu\nu}(\theta_0) \rangle. \quad (32)$$

We have already shown that

$$\mathcal{I}_{\mu\nu} = \frac{1}{T} \int_0^T dt \frac{\partial y}{\partial\theta_\mu} \frac{\partial y}{\partial\theta_\nu} \quad (33)$$

for Eq. (1) [see Eq. (6)]. We can rewrite this for oscillatory systems in light of Eq. (21),

$$y(t; \theta) = \tilde{y}(\phi(t; \theta); \theta). \quad (34)$$

Differentiating Eq. (34) with respect to θ_μ , we obtain

$$\frac{\partial y}{\partial\theta_\mu} \Big|_t = \frac{\partial\tilde{y}}{\partial\theta_\mu} \Big|_\phi + \frac{\partial\tilde{y}}{\partial\phi} \Big|_\theta \frac{\partial\phi}{\partial\theta_\mu} \Big|_t, \quad (35)$$

where the $|_x$ notation is used to indicate that the argument x is being held constant in the given derivative. This relationship is exact and shows a decoupling of the amplitude sensitivity from the phase sensitivity [similar to Eq. (11)]. Substituting Eq. (35) into Eq. (33) yields

$$\mathcal{I}_{\mu\nu} = \frac{1}{T} \int_0^T dt \left(\frac{\partial\tilde{y}}{\partial\theta_\mu} + \frac{\partial\tilde{y}}{\partial\phi} \frac{\partial\phi}{\partial\theta_\mu} \right) \left(\frac{\partial\tilde{y}}{\partial\theta_\nu} + \frac{\partial\tilde{y}}{\partial\phi} \frac{\partial\phi}{\partial\theta_\nu} \right). \quad (36)$$

We now show that Eq. (36) is also the FIM for Eq. (31). First we calculate the gradient of $C^\phi(\theta)$:

$$\frac{\partial C^\phi}{\partial\theta_\mu} = -\frac{1}{T} \int_0^T dt \left\{ \delta\hat{y} \left(\frac{\partial\tilde{y}}{\partial\theta_\mu} + \frac{\partial\tilde{y}}{\partial\phi} \frac{\partial\phi}{\partial\theta_\mu} \right) \right\}. \quad (37)$$

Next we calculate the Hessian and evaluate it at θ_0 [note that $\delta\hat{y}(\theta_0) = \delta\phi(\theta_0) = 0$]:

$$H_{\mu\nu}(\theta_0) = \frac{1}{T} \int_0^T dt \left(\frac{\partial\tilde{y}}{\partial\theta_\mu} + \frac{\partial\tilde{y}}{\partial\phi} \frac{\partial\phi}{\partial\theta_\mu} \right) \left(\frac{\partial\tilde{y}}{\partial\theta_\nu} + \frac{\partial\tilde{y}}{\partial\phi} \frac{\partial\phi}{\partial\theta_\nu} \right). \quad (38)$$

Clearly this is the same as Eq. (36). Because the FIM is preserved, the new cost [Eq. (31)] constitutes an isometric embedding of the model manifold and no information is lost (in the sense of the Fisher information).

4. Results

We implement the metric for the FitzHugh-Nagumo model as an example; results are shown in Figs. 8(a)–8(c). The local minima in the cost surface in Fig. 1(d) have been eliminated, the winding frequency of the stiffest direction is significantly reduced, and the manifold is no longer highly curved (see also Appendix E). Because the new cost function is an isometric embedding [i.e., preserves the Hessian in Eq. (6)], the curvature of the cost surface at the bottom of the bowl is the same as that in Fig. 1(d).

B. Kernel density estimation (KDE)

The high effective dimensionality of sloppy models of the third kind cannot be eliminated using the metric discussed in Sec. IV A. Adjusting the phase of a chaotic time series is insufficient to account for the variation in predictions as one moves from point to point in parameter space [resulting in the apparent roughness of the cost surface illustrated in Fig. 1(g)]. This is reflected in the exponential sensitivities of chaotic systems at long times and is connected with a fundamental difference in manifold structure between sloppy models of the second and third kinds. Note from Eq. (15) that winding frequency is directly proportional to geodesic (extrinsic) curvature. Figure 4 shows that the manifolds of sloppy models of the second kind only have high extrinsic curvature in one direction (like a scroll of paper). This high curvature can be eliminated through an isometric embedding (analogous to unwinding the scroll). By contrast, sloppy models of the third kind have high extrinsic curvature in more than one dimension. High extrinsic curvature in multiple dimensions is necessarily associated with some intrinsic curvature, and this intrinsic curvature cannot be eliminated through an isometric embedding (think of a globe, which can't be “unwound” and laid flat).

The sensitivities of chaotic time series to parameters (including initial conditions) make time series prediction in sloppy models of the third kind impractical at long times. However, model predictions $y_i(\theta)$ in phase space do give rise to a predictable distribution $f(y, \theta)$ [11]. We evolve an ensemble of initial conditions and use the result to approximate this distribution with a *kernel density estimate* [38,39],

$$f(y, \theta) \approx \frac{1}{nh} \sum_{i=1}^n K\left(\frac{y - y_i(\theta)}{h}\right), \quad (39)$$

where n is the number of predictions or observations, $K(\cdot)$ is a kernel function, and h is the kernel bandwidth. A natural metric to use for distributions is the *Hellinger distance*,

$$\tilde{C}(\theta) \equiv \frac{1}{2} \int dy (\sqrt{f(y, \theta_0)} - \sqrt{f(y, \theta)})^2, \quad (40)$$

because it is a quadratic form, which can be interpreted as a Euclidean embedding space. It also induces a metric on the model manifold that is given by the Fisher information metric.

We implement this cost for the Lorenz system; results are shown in Figs. 8(d)–8(f). The “rough” cost surface of Fig. 1(g) has been replaced with a basin, the high winding frequencies have all disappeared, and the manifold is regular (see also Appendix E).

V. CONCLUSION

Multimodality in comparing and training multiparameter models is a common problem [9,10,27–29]. Many common search algorithms find only a local minimum (a point in parameter space which locally minimizes the chosen distance measure) and not the global one. Even with global search methods the possibility of local minima reduces confidence that the global minimum will be found. Here, we have shown how the choice of distance measure affects the number of local minima. We have quantified this effect using curvature on the model manifold and introduced the *winding frequency* to estimate the density of local minima in parameter space. Finally, we have shown that through an appropriate choice of metric, the model manifold can be systematically “unwound” to remove local minima while preserving relevant physical interpretations of distance.

In this paper we have studied systems for which the relevant structures were known *a priori* (e.g., limit cycles and strange attractors). However, the metrics we propose may also be useful for identifying previously unknown structures in other complex systems.

One of the ongoing challenges for many complex systems is the development of appropriate reduced-order representations [23,40–42]. More broadly, it has been suggested that the existence of useful simplified models is often due to a systematic compression of parameter space [23]. Compressing the parameter space leads to a type of “universality class” in which models with different parameter values make statistically indistinguishable predictions. This line of work has also led to methods for constructing simplified models from more complex and complete mechanistic representations [43]. Ultimately, this compression is a consequence of the similarity metric used to compare model behaviors.

For sloppy models of the first kind (which have previously dominated the literature), the compression “squashes” some dimensions to be very thin [as in Fig. 1(a) and Refs. [21,22]], leading to a universality class of continuously connected parameters for which reduced-order models can be systematically derived [43]. In contrast, for sloppy models of the second and third kind, the manifold is wound tightly, so that a compression leads to a manifold folding in which noncontiguous regions of the manifold are identified as part of the same universality class. It is unlikely that predictive reduced-order models can be found for sloppy models with high winding frequencies as this would imply the existence, for example, of accurate long-term weather forecasts. High winding frequency is the information-geometric equivalent of sensitivity to microscopic details (such as frequency, initial conditions, or other parameters), well studied in chaotic systems. In contrast, by unwinding the model manifold using an alternative metric, the manifold is transformed into a hyper-ribbon and this extreme sensitivity to microscopic details is removed.

Understanding how effective theories emerge at long time scales is a challenging problem that has drawn on sophisticated expertise from a variety of fields, including dynamical systems [44,45], signal processing [4,6], statistics [26,46], and optimization [10,29]. In this work we have combined insights from these other domains with tools of information geometry.

Our hope is that this explicit connection will bring new tools, such as sloppy model analysis and the manifold boundary approximation method, to bear on a wide range of important, ongoing scientific problems.

ACKNOWLEDGMENTS

We thank A. Stankovic for helpful comments. Resources at the Fulton Supercomputing Lab at Brigham Young University were used for many of the computations. B.L.F. and M.K.T. were supported by the National Science Foundation through Grant No. NSF EPCN-1710727.

APPENDIX A: MODELS

Following are the models examined in Figs. 1, 2, 4, and 8. In some cases (models *D*, *G*, and *I*), additional polynomial terms (with parameters as coefficients) were added to the model equations of motion. This allows calculation of the *structural susceptibility* of the model, that is, susceptibility to perturbations of the underlying dynamics [47]. These terms can be thought of as representing details of the real system that have been left out of the model.

A: A sum of decaying exponentials leading to a steady state,

$$y(t; \theta) = \theta_1 + \sum_{n=2}^N e^{-\gamma_n t}, \quad (\text{A1a})$$

$$\gamma_n \equiv \sum_{i=2}^n \theta_i, \quad (\text{A1b})$$

where $\theta_i > 0$. Eigenvalues of the Hessian and winding frequencies [Figs. 1(b), 2, and 4] were calculated at $\theta_i = 1$; likewise for the cost surface in Fig. 1(a) except for the two parameters indicated on the axes. For the manifold projection in Fig. 1(c), $\ln \theta_2$, $\ln \theta_3$, and $\ln \theta_4$ were varied over a spherical volume of radius 20 centered on the original parameter values. Note that using θ_i as the parameters of the model, rather than using the decay rates γ_n directly, guarantees that the decay rates are ordered (i.e., $\gamma_{n+1} > \gamma_n$), breaking the symmetry between them.

B: A rational polynomial model,

$$y(t; \theta) = \frac{\theta_1 + \theta_2 t + \theta_3 t^2 + \theta_4 t^3}{1 + \theta_5 t + \theta_6 t^2 + \theta_7 t^3}. \quad (\text{A2})$$

Parameter values used were randomly chosen in the range $e^{-5} \leq \theta_i \leq e^5$.

C: We used the IFFLP model of biological adaptation described in [48].

D: The FitzHugh-Nagumo model [29,49,50] can be written as

$$\dot{V} = c \left(V - \frac{V^3}{3} + R + I + \sum_{n,m} \theta_{nm} V^n R^m \right), \quad (\text{A3a})$$

$$\dot{R} = -\frac{1}{c}(V - a + bR). \quad (\text{A3b})$$

We used a constant input current I , taken as a model parameter (in addition to the parameters a , b , c , and θ_{nm}). Initial

conditions used were $(V_0, R_0) = (-1, 1)$. Eigenvalues of the Hessian and winding frequencies [Figs. 1(e), 2, and 4] were calculated at $(a, b, c, I, \theta_{nm}) = (-0.0225, 0.135, 3.0, 0, 0)$, likewise for the cost surface in Fig. 1(d) except for the two parameters indicated on the axes. For the manifold projection in Fig. 1(f), a slice of parameter space along $-2 \leq a \leq 2$, $b = 0.2$ was used (all other parameters as above).

E: We implemented the Hodgkin-Huxley model described in [51].

F: We used the Wnt oscillator model described in [52].

G: The Lorenz system [53] is given by

$$\dot{x} = \sigma(y - x) + \sum_{n,m,p} \theta_{nmp} x^n y^m z^p, \quad (\text{A4a})$$

$$\dot{y} = x(\rho - z) - y, \quad (\text{A4b})$$

$$\dot{z} = xy - \beta z. \quad (\text{A4c})$$

Initial conditions used were $(x_0, y_0, z_0) = (1, 1, 10)$. Model parameters include σ , ρ , β , and θ_{nmp} . Additional parameters for rescaling x , y , and z after solving the ODE [e.g., $\tilde{x} \equiv (x - x_{\text{ref}})/x_{\text{scale}}$] were also included to illustrate that all parameters in a chaotic system need not exhibit an exponential sensitivity [see Fig. 1(h)]. (In general, parameters like these could account for differences in units between the model and the observations, if there were any.) Eigenvalues of the Hessian and winding frequencies [Figs. 1(h), 2, and 4] were calculated at $(\sigma, \rho, \beta, \theta_{nmp}, x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}}, x_{\text{scale}}, y_{\text{scale}}, z_{\text{scale}}) = (10, 28, 8/3, 0, 0, 0, 0, 1, 1, 1)$; likewise for the cost surface in Fig. 1(g) except for the two parameters indicated on the axes. For the manifold projection in Fig. 1(i), a slice of parameter space along $\sigma = 10.05$, $10 \leq \rho \leq 30$ was used (all other parameters as above).

H: The Hindmarsh-Rose model [54,55] can be written as

$$\dot{x} = y - ax^3 + bx^2 - z + I, \quad (\text{A5a})$$

$$\dot{y} = c - dx^2 - y, \quad (\text{A5b})$$

$$\dot{z} = \epsilon \left(x - \frac{1}{s}(z - z_R) \right). \quad (\text{A5c})$$

Initial conditions used were $(x_0, y_0, z_0) = (-0.216\,272\dots, 0.183\,969, 0.066\,920\dots)$. Model parameters include I (taken as a constant input current), a , b , c , d , ϵ , s , and z_R . Eigenvalues of the Hessian and winding frequencies (Figs. 2 and 4) were calculated at $(a, b, c, d, I, \epsilon, s, z_R) = (1, 3, 1, 5, 0, 0.004, 4, 3.1586)$.

I: The equations of motion for a damped, driven pendulum (derivable using Newton's 2nd law) are

$$\dot{\phi} = \omega + \sum_{n,m,p} \theta_{nmp} \phi^n \omega^m \phi^p, \quad (\text{A6a})$$

$$\dot{\omega} = -\frac{\omega}{Q} - \sin(\phi) + A \cos(\phi), \quad (\text{A6b})$$

$$\dot{\phi} = \omega_D. \quad (\text{A6c})$$

Initial conditions used were $(\phi_0, \omega_0, \phi_0) = (-2, 0, 0)$. Model parameters include Q , A , ω_D , and θ_{nmp} . Eigenvalues of the Hessian and winding frequencies (Figs. 2 and 4) were calculated at $(Q, A, \omega_D, \theta_{nmp}) = (2, 1.16, 2/3, 0)$.

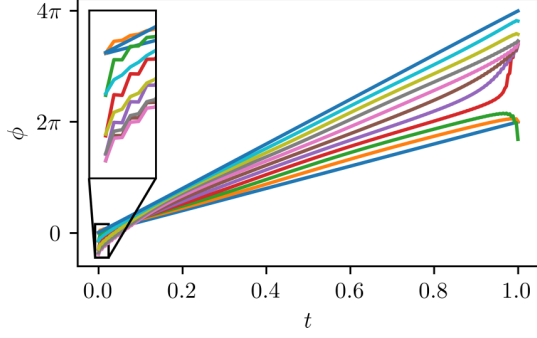


FIG. 9. Phases obtained when implementing the Hilbert transform numerically on the model $y(t, \theta) = A \cos(\omega t)$, for ω ranging from 2π to 4π . The effects of the Gibbs phenomenon can be seen near the ends for some values of ω .

APPENDIX B: ALTERNATIVES FOR OBTAINING A PHASE

In some cases, Eq. (19) cannot be used to obtain a monotonically increasing phase. For example, some oscillatory behavior does not have a unique center of oscillation. If that is the case, one approach is to decompose the signal using *empirical mode decomposition* into a number of *intrinsic mode functions*, for each of which a separate phase may then be defined [56]. However, because this method is empirical, the decomposition may not vary smoothly with the parameters of the model, leading to discontinuities in the cost function.

Even when the oscillatory behavior does have a single center of oscillation, in practice the Hilbert transform must be implemented numerically (especially for observational data). This usually involves a fast Fourier transform, which can introduce unwanted effects in the phase due to the Gibbs phenomenon (see Fig. 9). The impact of end effects can be reduced by leaving the ends out of the cost function, or through windowing.

More generally, any monotonically increasing function of time may be used for a phase, provided it has the appropriate frequency. One proposal is to use

$$\phi(t) = \omega t + \phi_0, \quad (\text{B1})$$

and to estimate a value of ω from the oscillatory signal. This may be done by fitting a line to the phase obtained from Eq. (19) or by using a Fourier transform to decompose the signal into frequency components and selecting one.

We also suggest the following method of obtaining a phase (found in [35]) that does not require the use of the Hilbert transform. It is sometimes the case that two signals, $y_1(t)$ and $y_2(t)$, can be selected from the dynamical variables $y(t)$ of a system and used to calculate a phase as follows:

$$\phi(t) = \arctan\left(\frac{y_2(t)}{y_1(t)}\right). \quad (\text{B2})$$

The only requirement is that the combined signal correspond to a *proper rotation*, which has both a definite direction and unique center of rotation, so that the phase will be monotonically increasing [36,57]. For example, in some cases, a signal

$y(t)$ and its time derivative $\dot{y}(t)$ may be used:

$$\phi(t) = \arctan\left(\frac{y(t)}{\dot{y}(t)}\right). \quad (\text{B3})$$

APPENDIX C: COVARIANCE MATRICES

We consider how uncertainty in experimental observations propagates to phases calculated using Eq. (19). First, we define more precisely the covariance matrix $\Sigma^{y(t)}$ for the observations with time as the independent variable. Let ξ_i denote random variables drawn from the normal distribution $\mathcal{N}(0, 1)$. We assume the observations y_i are random variables that are normally distributed about the predictions $y(t_i; \theta_0)$ of the model at the best fit, with standard deviation given by the uncertainties σ_i , and write

$$y_i = y(t_i; \theta_0) + \sigma_i \xi_i. \quad (\text{C1})$$

The deviations

$$\delta y_i \equiv y_i - y(t_i; \theta) \quad (\text{C2})$$

vary with the predictions of the model, but at the best fit they are random variables with mean 0 and standard deviation σ_i :

$$\delta y_i(\theta_0) = y(t_i; \theta_0) + \sigma_i \xi_i - y(t_i; \theta_0) = \sigma_i \xi_i. \quad (\text{C3})$$

The elements of the covariance matrix are defined as the expectation of the product of deviations at the best fit:

$$\Sigma_{ij}^{y(t)} \equiv \langle \delta y_i \delta y_j \rangle = \langle \sigma_i \xi_i \sigma_j \xi_j \rangle = \sigma_i \sigma_j \langle \xi_i \xi_j \rangle. \quad (\text{C4})$$

The matrix is diagonal if the deviations are independent (i.e., if $\langle \xi_i \xi_j \rangle = \delta_{ij}$).

1. Covariance matrix for phase

Next, we derive the covariance matrix for the phases. The observations y_i are assumed to have occurred at the phases $\phi(t_i; \theta_0)$ predicted by the model. These phases will differ from the phases ϕ_i calculated using Eq. (19) due to the presence of noise in the observations. We define the deviations of the phases as

$$\delta \phi_i \equiv \phi_i - \phi(t_i; \theta). \quad (\text{C5})$$

Note that, due to the presence of the Hilbert transform in Eq. (19), the phase $\phi(t)$ has a functional dependence on the signal $y(t)$, i.e., $\phi(t) = \phi[y](t)$. We use this functional dependence and Eq. (C2) to relate $\delta \phi_i$ to δy_i :

$$\begin{aligned} \phi_i &= \phi_i[y] \\ &= \phi_i[y(t; \theta) + \delta y] \\ &\approx \phi_i[y(t; \theta)] + \sum_j \frac{\partial \phi_i[y(t; \theta)]}{\partial y_j} \delta y_j \\ &= \phi(t_i; \theta) + \sum_j \frac{\partial \phi_i}{\partial y_j} \delta y_j, \end{aligned} \quad (\text{C6})$$

$$\delta \phi_i \approx \sum_j \frac{\partial \phi_i}{\partial y_j} \delta y_j. \quad (\text{C7})$$

In the fourth line we have simplified the notation for clarity, and we have kept only the first-order terms. This

approximation is valid near the best fit where δy_i is small. At the best fit, we have

$$\delta\phi_i(\theta_0) = \sum_j \frac{\partial\phi_i}{\partial y_j} \sigma_j \xi_j, \quad (\text{C8})$$

which shows that $\delta\phi_i(\theta_0)$ are random variables with mean 0.

Before proceeding, the derivative $\partial\phi_i/\partial y_j$ merits some attention. First, we note that it may be evaluated using either $y(t_i; \theta)$ or y_i to first order in δy_i :

$$\begin{aligned} \frac{\partial\phi_i[y(t_j; \theta)]}{\partial y_j} \delta y_j &= \frac{\partial\phi_i[y_j - \delta y_j]}{\partial y_j} \delta y_j \\ &= \frac{\partial\phi_i[y_j]}{\partial y_j} \delta y_j + O(\delta y^2). \end{aligned} \quad (\text{C9})$$

Second, using Eq. (19), we can derive an explicit expression for $\partial\phi_i/\partial y_j$:

$$\begin{aligned} \frac{\partial\phi_i}{\partial y_j} &= \frac{\partial}{\partial y_j} \left[\tan^{-1} \left(\frac{H_i[y]}{y_i} \right) \right] \\ &= \frac{1}{1 + (H_i[y]/y_i)^2} \left(\frac{1}{y_i} \frac{\partial H_i[y]}{\partial y_j} - \frac{H_i[y]}{y_i^2} \frac{\partial y_i}{\partial y_j} \right). \end{aligned} \quad (\text{C10})$$

($H_i[y]$ is understood to mean the i th component of the Hilbert transform of y .) To evaluate the derivative $\partial H_i[y]/\partial y_j$, we use the definition of the derivative and the linearity of the Hilbert transform:

$$\begin{aligned} \frac{\partial H_i[y]}{\partial y_j} &= \lim_{h \rightarrow 0} \frac{H_i[y + h\delta_j] - H_i[y]}{h} \\ &= \lim_{h \rightarrow 0} \frac{H_i[y] + hH_i[\delta_j] - H_i[y]}{h} = H_i[\delta_j]. \end{aligned} \quad (\text{C11})$$

(We are using δ_j to denote the vector formed by taking the j th column of the Kronecker delta δ_{ij} when considered as a matrix.) Plugging this into Eq. (C10) gives

$$\frac{\partial\phi_i}{\partial y_j} = \frac{y_i H_i[\delta_j] - H_i[y] \delta_{ij}}{y_i^2 + H_i[y]^2}. \quad (\text{C12})$$

Third, the matrix $\partial\phi/\partial y$ defined by Eq. (C12) is singular (i.e., it has at least one zero eigenvalue). As we now show, this is because changes in the amplitude of an oscillation do not affect the phase.

Theorem. The matrix $\partial\phi/\partial y$, whose ij th element is

$$\frac{\partial\phi_i}{\partial y_j} = \frac{y_i H_i[\delta_j] - H_i[y] \delta_{ij}}{y_i^2 + H_i[y]^2},$$

has at least one zero eigenvalue, corresponding to the eigenvector $\delta y^* = y$.

Proof.

$$\begin{aligned} \sum_j \frac{\partial\phi_i}{\partial y_j} \delta y_j^* &= \sum_j \frac{y_i H_i[\delta_j] - H_i[y] \delta_{ij}}{y_i^2 + H_i[y]^2} y_j \\ &= \frac{y_i H_i[\sum_j \delta_j y_j] - H_i[y] \sum_j \delta_{ij} y_j}{y_i^2 + H_i[y]^2} \\ &= \frac{y_i H_i[y] - H_i[y] y_i}{y_i^2 + H_i[y]^2} = 0. \end{aligned} \quad \blacksquare$$

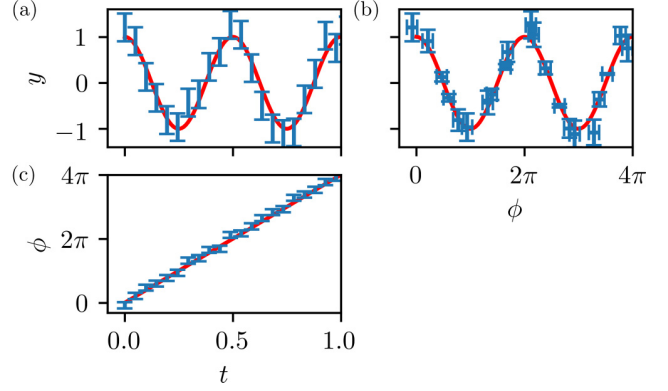


FIG. 10. Propagation of uncertainty. (a) Data (blue) simulated from the model $y(t) = A \cos(\omega t)$ (red) by adding uniform Gaussian noise. Error bars indicate uncertainty. (b) Data (blue) plotted as a function of phase compared with $y(\phi) = A \cos(\phi)$ (red). Error bars indicate the uncertainties obtained using Eqs. (C14) and (C20). (c) Phase (blue), obtained for each data point using Eq. (19), compared with $\phi(t) = \omega t$ (red). Error bars indicate the uncertainties obtained using Eq. (C14).

Any change in amplitude at constant phase is a multiple of y and thus also lies in the null space of $\partial\phi/\partial y$.

Returning to Eq. (C7), we derive an expression for the covariance matrix $\Sigma^{\phi(t)}$ for the phases:

$$\begin{aligned} \Sigma_{ij}^{\phi(t)} &\equiv \langle \delta\phi_i \delta\phi_j \rangle \\ &= \left\langle \sum_k \frac{\partial\phi_i}{\partial y_k} \delta y_k \sum_l \frac{\partial\phi_j}{\partial y_l} \delta y_l \right\rangle \\ &= \sum_{k,l} \frac{\partial\phi_i}{\partial y_k} \langle \delta y_k \delta y_l \rangle \frac{\partial\phi_j}{\partial y_l} \\ &= \sum_{k,l} \frac{\partial\phi_i}{\partial y_k} \Sigma_{kl}^{y(t)} \frac{\partial\phi_j}{\partial y_l} \end{aligned} \quad (\text{C13})$$

$$\Sigma^{\phi(t)} = \frac{\partial\phi}{\partial y} \Sigma^{y(t)} \frac{\partial\phi}{\partial y}^T. \quad (\text{C14})$$

This shows how uncertainties $\sigma_{y_i}^2 = \Sigma_{ii}^{y(t)}$ in the observations are propagated to uncertainties $\sigma_{\phi_i}^2 = \Sigma_{ii}^{\phi(t)}$ in the phases of the observations (see Fig. 10).

2. Covariance matrix for observations as a function of phase

Finally, we derive the covariance matrix $\Sigma^{y(\phi)}$ for the observations with phase as the independent variable. We define the deviations of the observations from the predictions at constant phase as

$$\delta\tilde{y}_i \equiv y_i - \tilde{y}(\phi_i; \theta). \quad (\text{C15})$$

We can relate these to δy_i and $\delta \phi_i$ using Eqs. (C5) and (21):

$$\begin{aligned}
 \delta \tilde{y}_i &= y_i - \tilde{y}(\phi_i; \theta) \\
 &= y_i - \tilde{y}(\phi(t_i; \theta) + \delta \phi_i; \theta) \\
 &\approx y_i - \tilde{y}(\phi(t_i; \theta), \theta) - \frac{\partial \tilde{y}(\phi(t_i; \theta); \theta)}{\partial \phi} \delta \phi_i \\
 &= y_i - y(t_i; \theta) - \left(\frac{\partial \tilde{y}}{\partial \phi} \right)_i \delta \phi_i \\
 &= \delta y_i - \left(\frac{\partial \tilde{y}}{\partial \phi} \right)_i \delta \phi_i. \tag{C16}
 \end{aligned}$$

In light of Eqs. (C3) and (C8), $\delta \tilde{y}_i$ also has mean 0 at the best fit. Note that, similar to $\partial \phi_i / \partial y_j$, $\partial \tilde{y} / \partial \phi$ may be evaluated using either ϕ_i or $\phi(t_i; \theta)$ to first order in $\delta \phi_i$:

$$\begin{aligned}
 \frac{\partial \tilde{y}(\phi(t_i; \theta); \theta)}{\partial \phi} \delta \phi_i &= \frac{\partial \tilde{y}(\phi_i - \delta \phi_i; \theta)}{\partial \phi} \delta \phi_i \\
 &= \frac{\partial \tilde{y}(\phi_i; \theta)}{\partial \phi} \delta \phi_i + \mathcal{O}(\delta \phi^2). \tag{C17}
 \end{aligned}$$

We can take Eq. (C16) a step further using Eq. (C7):

$$\begin{aligned}
 \delta \tilde{y}_i &= \delta y_i - \left(\frac{\partial \tilde{y}}{\partial \phi} \right)_i \delta \phi_i \\
 &= \delta y_i - \sum_j \left(\frac{\partial \tilde{y}}{\partial \phi} \right)_i \frac{\partial \phi_i}{\partial y_j} \delta y_j \\
 &= \sum_j \left[\delta_{ij} - \left(\frac{\partial \tilde{y}}{\partial \phi} \right)_i \frac{\partial \phi_i}{\partial y_j} \right] \delta y_j \\
 &\equiv \sum_j D_{ij} \delta y_j. \tag{C18}
 \end{aligned}$$

Taking the expectation of pairs of deviations $\delta \tilde{y}_i$, we obtain

$$\begin{aligned}
 \Sigma_{ij}^{y(\phi)} &\equiv \langle \delta \tilde{y}_i \delta \tilde{y}_j \rangle \\
 &= \left\langle \sum_k D_{ik} \delta y_k \sum_l D_{jl} \delta y_l \right\rangle \\
 &= \sum_{k,l} D_{ik} \langle \delta y_k \delta y_l \rangle D_{jl} \\
 &= \sum_{k,l} D_{ik} \Sigma_{kl}^{y(t)} D_{jl} \tag{C19}
 \end{aligned}$$

$$\Sigma^{y(\phi)} = D \Sigma^{y(t)} D^T. \tag{C20}$$

This gives us a way to compute the uncertainties $\sigma_{\tilde{y}_i}^2 = \Sigma_{ii}^{y(\phi)}$ of the observations when taking phase as the independent variable instead of time (see Fig. 10).

APPENDIX D: PARAMETER SENSITIVITIES

Here we derive the first- and second-order parameter sensitivities of \tilde{y} and ϕ that are used in calculating the FIM and winding frequencies for the analytic signal-based metric of

Sec. IV A. We begin with Eq. (20),

$$y(t; \theta) = A(t; \theta) \cos(\phi(t; \theta)), \tag{D1}$$

and differentiate it with respect to θ_μ :

$$\frac{\partial y}{\partial \theta_\mu} = \frac{\partial A}{\partial \theta_\mu} \cos(\phi) - A \sin(\phi) \frac{\partial \phi}{\partial \theta_\mu}. \tag{D2}$$

Comparing with Eq. (35), we now see that we have explicit expressions for $\partial \tilde{y} / \partial \theta_\mu$ and $\partial \tilde{y} / \partial \phi$ in terms of A , ϕ , and $\partial A / \partial \theta_\mu$:

$$\begin{aligned}
 \left. \frac{\partial \tilde{y}}{\partial \theta_\mu} \right|_\phi &= \frac{\partial A}{\partial \theta_\mu} \cos(\phi) & \left. \frac{\partial \tilde{y}}{\partial \phi} \right|_\theta &= -A \sin(\phi) \\
 &= \frac{y}{A} \frac{\partial A}{\partial \theta_\mu} & &= -H[y]. \tag{D3}
 \end{aligned}$$

In the second line we have used the trigonometric relationships $\cos(\phi) = y/A$ and $\sin(\phi) = H[y]/A$ which are easily derived from Eqs. (18) and (19). The second derivative of Eq. (D1) is

$$\begin{aligned}
 \frac{\partial^2 y}{\partial \theta_\mu \partial \theta_\nu} &= \frac{y}{A} \frac{\partial^2 A}{\partial \theta_\mu \partial \theta_\nu} - \frac{H[y]}{A} \left(\frac{\partial A}{\partial \theta_\mu} \frac{\partial \phi}{\partial \theta_\nu} + \frac{\partial \phi}{\partial \theta_\mu} \frac{\partial A}{\partial \theta_\nu} \right) \\
 &\quad - y \frac{\partial \phi}{\partial \theta_\mu} \frac{\partial \phi}{\partial \theta_\nu} - H[y] \frac{\partial^2 \phi}{\partial \theta_\mu \partial \theta_\nu}. \tag{D4}
 \end{aligned}$$

Because the new analytic signal-based metric involves \tilde{y} and ϕ , we use only the first term (which is $\partial^2 \tilde{y} / \partial \theta_\mu \partial \theta_\nu$) and the last term in this expression when calculating the geodesic curvature.

Expressions for the sensitivities of A and ϕ are obtained by differentiating Eqs. (18) and (19):

$$A = \sqrt{y^2 + H^2[y]} \quad \phi = \tan^{-1} \left(\frac{H[y]}{y} \right) \tag{D5}$$

$$\frac{\partial A}{\partial \theta_\mu} = \frac{1}{A} \left(y \frac{\partial y}{\partial \theta_\mu} + H[y] H \left[\frac{\partial y}{\partial \theta_\mu} \right] \right) \tag{D6}$$

$$\frac{\partial \phi}{\partial \theta_\mu} = \frac{1}{A^2} \left(y H \left[\frac{\partial y}{\partial \theta_\mu} \right] - H[y] \frac{\partial y}{\partial \theta_\mu} \right) \tag{D7}$$

$$\frac{\partial^2 A}{\partial \theta_\mu \partial \theta_\nu} = A \frac{\partial \phi}{\partial \theta_\mu} \frac{\partial \phi}{\partial \theta_\nu} + \frac{1}{A} \left(y \frac{\partial^2 y}{\partial \theta_\mu \partial \theta_\nu} + H[y] H \left[\frac{\partial^2 y}{\partial \theta_\mu \partial \theta_\nu} \right] \right) \tag{D8}$$

$$\begin{aligned}
 \frac{\partial^2 \phi}{\partial \theta_\mu \partial \theta_\nu} &= -\frac{1}{A} \frac{\partial A}{\partial \theta_\mu} \frac{\partial \phi}{\partial \theta_\nu} - \frac{1}{A} \frac{\partial \phi}{\partial \theta_\mu} \frac{\partial A}{\partial \theta_\nu} \\
 &\quad + \frac{1}{A^2} \left(y H \left[\frac{\partial^2 y}{\partial \theta_\mu \partial \theta_\nu} \right] - H[y] \frac{\partial^2 y}{\partial \theta_\mu \partial \theta_\nu} \right). \tag{D9}
 \end{aligned}$$

APPENDIX E: REGULARITY OF COST SURFACES AND MANIFOLDS

In Fig. 8, a sufficiently large number of time points was included in the cost and manifold calculations to demonstrate the results of using the new metrics in the limit of infinite time. In practice, only a finite number of time points can be included. Here we demonstrate the convergence of the

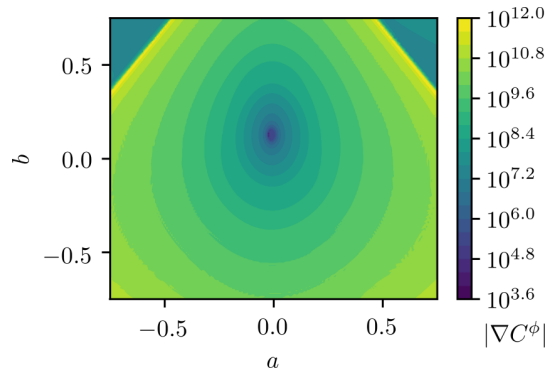


FIG. 11. Magnitude of the gradient of the FitzHugh-Nagumo cost $|\nabla C^\phi(\theta)|$. The magnitude of the gradient has only one minimum, indicating that the cost cross section shown in Fig. 8(a) has a single minimum. The minimum of $|\nabla C^\phi(\theta)|$ shown is not quite zero because the actual minimum of Fig. 8(a) is between the grid points where $|\nabla C^\phi(\theta)|$ has been calculated. Note that in the upper corners of the plot, there is a phase transition to nonoscillatory behavior, where the methods of Sec. IV A cannot be applied effectively. The sharp apparent dropoff is due to such choices as having our algorithms return zeros rather than throw errors for these regions.

FitzHugh-Nagumo manifold and the Lorenz cost as a function of the number of sampled time points. In addition, we discuss the gradient of the FitzHugh-Nagumo cost, shown in Fig. 8(a), as it relates to the regularity of the new surface.

Figure 11 shows a plot of the magnitude of the gradient of the cost cross section shown in Fig. 8(a). The significance of the gradient of the cost is that every local minimum of the cost will be a zero of the gradient. If there are multiple local minima still present in the new cost, then the gradient will have multiple zeros. We plot the magnitude of the gradient so that zeros can be found easily. It is clear from Fig. 11 that there is only one zero, so the new cost does, in fact, have a single minimum.

Figure 12 shows two projections of the FitzHugh-Nagumo manifold (signal predictions at constant phase): one calculated using about 24 time points per cycle in the original time series and the other using twice the time sampling of the first.

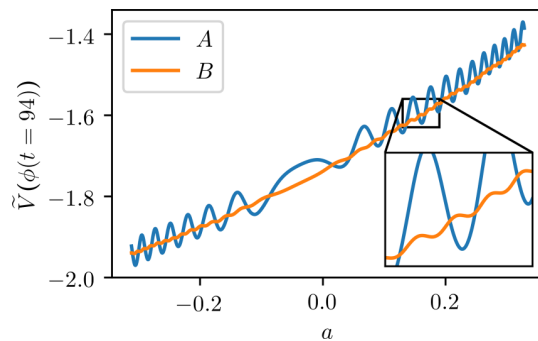


FIG. 12. FitzHugh-Nagumo manifold projection. A was calculated using about 24 time points per cycle in the original time series; B was calculated using twice the time sampling of A.

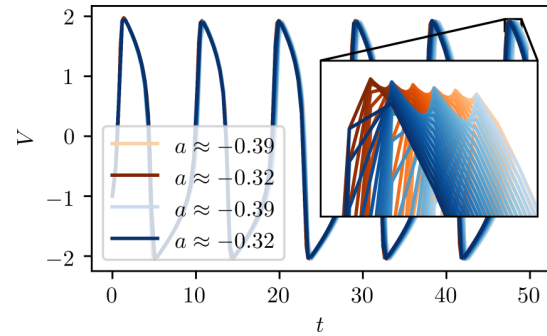


FIG. 13. FitzHugh-Nagumo amplitude oscillations. Colors are the same as in Fig. 12, with dark or light indicating the value of the parameter a . As the peak moves between sampled time points, the amplitude appears to oscillate.

The manifold itself exhibits oscillations in both cases. These oscillations are an artifact of the finite time sampling of the oscillatory signal predicted by the model. As parameters that control frequency are varied, the peak of each cycle shifts between adjacent time points and the local amplitude appears to oscillate (see Fig. 13). Hence the predicted signal values at a given (constant) phase also oscillate, resulting in the manifold oscillations observed.

As demonstrated in Fig. 12, doubling the sampling of time points doubles the frequency of these manifold oscillations, but their amplitude decreases by a factor of ~ 10 . Hence, in the limit of infinite sampling they disappear. In practice they will be negligible as long as enough time points per cycle are sampled for the amplitude of the oscillations to be small compared to the amplitude of the signal itself (and to changes effected by the parameters).

The attractors of chaotic systems have fractal structure that is realized only in the limit of infinite sampling time T . Accordingly, as more time points are included, the kernel density estimate Eq. (39) will approach the true distribution $f(y, \theta)$ asymptotically. Figure 14 illustrates the convergence of a cross section of the cost Eq. (40) for the Lorenz system as the total sampling time T is varied.

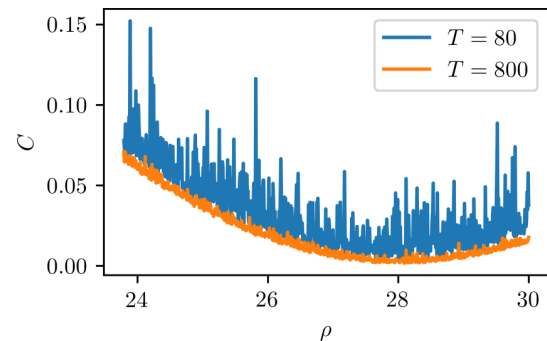


FIG. 14. Lorenz cost. As the number of sampled time points grows, the noise in the cost dies away. When fit to a parabola, the MAE between the parabola and the cost cross section shown is 0.0088 for $T = 80$ and 0.0013 for $T = 800$ (about a sevenfold reduction in noise).

- [1] Although our interest here is in mathematical models, the same is true of model organisms in biology that, by virtue of being (for example) mammals, vertebrates, eukaryotes, etc., have a similar phylogeny to a target system, e.g., humans.
- [2] M. Pospischil, M. Toledo-Rodriguez, C. Monier, Z. Piwkowska, T. Bal, Y. Frégnac, H. Markram, and A. Destexhe, *Biol. Cybern.* **99**, 427 (2008).
- [3] G. LeMasson, R. Maex, in *Computational Neuroscience: Realistic Modeling for Experimentalists*, edited by E. De Schutter (CRC, Boca Raton, 2000), Chap. 1, pp. 1–24.
- [4] M. A. Kramer, H. Rabitz, and J. M. Calo, *Appl. Math. Modell.* **8**, 328 (1984).
- [5] R. Larter, H. Rabitz, and M. Kramer, *J. Chem. Phys.* **80**, 4120 (1984).
- [6] A. K. Wilkins, B. Tidor, J. White, and P. I. Barton, *SIAM J. Sci. Comput.* **31**, 2706 (2009).
- [7] D. J. Lea, M. R. Allen, and T. W. Haine, *Tellus A* **52**, 523 (2000).
- [8] V. F. Pisarenko and D. Sornette, *Phys. Rev. E* **69**, 036122 (2004).
- [9] J. Annan and J. Hargreaves, *Tellus A* **56**, 520 (2004).
- [10] A. C. Lorenc and T. Payne, *Q. J. R. Meteorol. Soc.* **133**, 607 (2007).
- [11] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics* (Springer, New York, 2013), Vol. 97.
- [12] K. S. Brown and J. P. Sethna, *Phys. Rev. E* **68**, 021904 (2003).
- [13] K. S. Brown, C. C. Hill, G. A. Calero, C. R. Myers, K. H. Lee, J. P. Sethna, and R. A. Cerione, *Phys. Biol.* **1**, 184 (2004).
- [14] S. L. Frederiksen, K. W. Jacobsen, K. S. Brown, and J. P. Sethna, *Phys. Rev. Lett.* **93**, 165501 (2004).
- [15] J. J. Waterfall, F. P. Casey, R. N. Gutenkunst, K. S. Brown, C. R. Myers, P. W. Brouwer, V. Elser, and J. P. Sethna, *Phys. Rev. Lett.* **97**, 150601 (2006).
- [16] R. Gutenkunst, Ph.D. thesis, Cornell University, 2007.
- [17] F. P. Casey, D. Baird, Q. Feng, R. N. Gutenkunst, J. J. Waterfall, C. R. Myers, K. S. Brown, R. A. Cerione, and J. P. Sethna, *IET Syst. Biol.* **1**, 190 (2007).
- [18] R. N. Gutenkunst, F. P. Casey, J. J. Waterfall, C. R. Myers, and J. P. Sethna, *Ann. N.Y. Acad. Sci.* **1115**, 203 (2007).
- [19] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, J. P. Sethna, *PLoS Comput. Biol.* **3**, e189 (2007).
- [20] B. C. Daniels, Y.-J. Chen, J. P. Sethna, R. N. Gutenkunst, and C. R. Myers, *Curr. Opin. Biotechnol.* **19**, 389 (2008).
- [21] M. K. Transtrum, B. B. Machta, and J. P. Sethna, *Phys. Rev. Lett.* **104**, 060201 (2010).
- [22] M. K. Transtrum, B. B. Machta, and J. P. Sethna, *Phys. Rev. E* **83**, 036701 (2011).
- [23] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, *Science* **342**, 604 (2013).
- [24] M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna, *J. Chem. Phys.* **143**, 010901 (2015).
- [25] C. H. LaMont and P. A. Wiggins, [arXiv:1506.05855v3](https://arxiv.org/abs/1506.05855v3).
- [26] C. H. LaMont and P. A. Wiggins, [arXiv:1506.05855v4](https://arxiv.org/abs/1506.05855v4).
- [27] P. Mendes and D. Kell, *Bioinformatics* **14**, 869 (1998).
- [28] C. G. Moles, P. Mendes, and J. R. Banga, *Genome Res.* **13**, 2467 (2003).
- [29] J. O. Ramsay, G. Hooker, D. Campbell, and J. Cao, *J. R. Stat. Soc.: Series B* **69**, 741 (2007).
- [30] L. Yao and W. A. Sethares, *IEEE Trans. Signal Process.* **42**, 927 (1994).
- [31] Q. He, L. Wang, and B. Liu, *Chaos, Solitons Fractals* **34**, 654 (2007).
- [32] M. Rodriguez-Fernandez, P. Mendes, and J. R. Banga, *Biosystems* **83**, 248 (2006).
- [33] E. Baake, M. Baake, H. G. Bock, and K. M. Briggs, *Phys. Rev. A* **45**, 5524 (1992).
- [34] B. Boashash, *Proc. IEEE* **80**, 520 (1992).
- [35] A. S. Pikovsky, M. G. Rosenblum, G. V. Osipov, and J. Kurths, *Phys. D (Amsterdam, Neth.)* **104**, 219 (1997).
- [36] I. Z. Kiss and J. L. Hudson, *Phys. Rev. E* **64**, 046215 (2001).
- [37] D. Gabor, *J. Inst. Electr. Eng., Part 3* **93**, 429 (1946).
- [38] M. Rosenblatt, *Ann. Math. Stat.* **27**, 832 (1956).
- [39] E. Parzen, *Ann. Math. Stat.* **33**, 1065 (1962).
- [40] P. W. Anderson, *Science* **177**, 393 (1972).
- [41] C. Theodoropoulos, Y.-H. Qian, and I. G. Kevrekidis, *Proc. Natl. Acad. Sci. USA* **97**, 9840 (2000).
- [42] M. K. Transtrum and P. Qiu, *PLoS Comput. Biol.* **12**, e1004915 (2016).
- [43] M. K. Transtrum and P. Qiu, *Phys. Rev. Lett.* **113**, 098701 (2014).
- [44] C. W. Gear, J. M. Hyman, P. G. Kevrekidis, I. G. Kevrekidis, O. Runborg, and C. Theodoropoulos, *Commun. Math. Sci.* **1**, 715 (2003).
- [45] M. Desroches, J. Guckenheimer, B. Krauskopf, C. Kuehn, H. M. Osinga, and M. Wechselberger, *SIAM Rev.* **54**, 211 (2012).
- [46] H. H. Mattingly, M. K. Transtrum, M. C. Abbott, and B. B. Machta, *Proc. Natl. Acad. Sci. USA* **115**, 1760 (2018).
- [47] R. Chachra, M. K. Transtrum, and J. P. Sethna, *Phys. Rev. E* **86**, 026712 (2012).
- [48] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang, *Cell* **138**, 760 (2009).
- [49] R. FitzHugh, *Biophys. J.* **1**, 445 (1961).
- [50] J. Nagumo, S. Arimoto, and S. Yoshizawa, *Proc. IREE* **50**, 2061 (1962).
- [51] A. L. Hodgkin and A. F. Huxley, *J. Physiol.* **117**, 500 (1952).
- [52] P. B. Jensen, L. Pedersen, S. Krishna, and M. H. Jensen, *Biophys. J.* **98**, 943 (2010).
- [53] E. N. Lorenz, *J. Atmos. Sci.* **20**, 130 (1963).
- [54] J. Hindmarsh and R. Rose, *Proc. R. Soc. London, Ser. B* **221**, 87 (1984).
- [55] X.-J. Wang, *Phys. D (Amsterdam, Neth.)* **62**, 263 (1993).
- [56] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, *Proc. R. Soc. London, Ser. A* **454**, 903 (1998).
- [57] T. Yalçinkaya and Y.-C. Lai, *Phys. Rev. Lett.* **79**, 3885 (1997).

Chapter 4

Data-Driven Model Reduction for Network Simplification and Effective Equivalents

Networks play an important role in epidemiology, ecology, systems biology, the Internet, and power systems, among others. Most real-world networks involve huge numbers of interacting components, agents, or nodes. A detailed, “mechanistic” mathematical description of such a network, which might be used to study and predict its behavior, quickly becomes very complicated. Often the available observational data are insufficient to identify all of the parameter values needed to complete the model and validate it. In other cases, detail in certain parts of the system is simply irrelevant and unnecessary to accomplish the purpose for which the model was created.

In this chapter, I demonstrate that MBAM can be used as a tool for reducing models of networked systems in a data-driven way. I apply the method to a small network from the field of power systems and show that the amount of reduction is tailored to the observations being made on the system. This results in several “effective” networks which capture the primary features of each particular choice of observations being made on the system. I also discover several important types of reductions (i.e., parameter limits) that are applicable to power systems and, more generally, to certain types

of networks. Information about the types of reductions which can be performed on models with similar mathematical structures is an important step toward scaling up MBAM model reduction to very large models.

This work is an extension of previous work done by Mark Transtrum, Andrija Sarić, and Aleksandar Stanković which lead to a short conference paper [45] ©2018 IEEE. My contributions to this work include performing most of the model reduction calculations discussed herein, which are an improvement on the previous work in several ways, as explained in the article; expanding sections which were carried over from the previous article; and writing much new material, including all of the results. Jacob Nuttall assisted in running some of the calculations during the reduction process. Sarić and Stanković provided expertise in power systems, as well as the power systems model used and initial parameter values.

This article has been prepared for submission to a peer-reviewed journal but has not yet been submitted. Some preliminary results that are repeated here were published in the conference proceedings of the 2019 North American Power Symposium [46] ©2019 IEEE. I hereby confirm that the use of this material is compliant with all publishing agreements.

Data-Driven Model Reduction for Network Simplification and Effective Equivalents

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—Models of dynamical phenomena in large power networks have grown significantly over time. Although available measurements have also increased, data are generally insufficient to constrain all of the parameters. This is because the models are often sloppy, that is, system behavior is determined by a few parameter combinations. Because some parameter combinations are neither identifiable from data nor useful for making new predictions, they can be eliminated without significantly affecting the reliability of the model. We use the Manifold Boundary Approximation Method (MBAM) to remove unidentifiable parameters from a model of a small power network (the IEEE 14-bus test system) in a data-driven way. We consider five different possible measurement structures and show that the reduction is tailored to the available measurements. The result in each case is both a reduced dynamical model and an equivalent network that together serve as a “gray-box” approximation or effective theory appropriate for the available measurement data. Parsimonious models give insights into the relationships among observed system components and latent variables that can guide developing dynamic equivalents and other effective descriptions of large power systems.

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

I. INTRODUCTION

As power systems have evolved in response to changes in markets, source and load types, and other pressures, models used to study dynamic phenomena in power systems have also grown in size and detail. Advances in computing capabilities have helped alleviate the computational strain resulting from the increases in size. However, these more-detailed models still have trouble replicating real events recorded in power systems [1].

This has resulted in many efforts to improve models through system identification. The use of trajectory sensitivities in identifying parameters of various system components, for example, was demonstrated in [2], [3] and extended to hybrid systems in [4]. But detailed models often have many more parameters than can be effectively identified from the available measurements. One proposal for dealing with parameter insensitivity is subset selection [5], in which the subset of

parameters that are most insensitive or ill-conditioned are held fixed during parameter estimation.

Network dynamics, where the discrete structure of a network is blended with the continuous dynamics of system components [6], presents additional challenges. In many cases, measurements are not available for certain parts of the network, or analysts may only be interested in a particular portion of the network. Identifiability of linear dynamical networks where the topology is known *a priori* but measurements may not be available for every node was previously explored in [7]. A joint parameter and topology estimation method was demonstrated on radial networks in [8].

Parameter insensitivity, lack of available measurements, and lack of a need for detail in some parts of a model all motivate the use of reduced models. One approach to model reduction is to identify dominant modes in the dynamics and then to neglect the others. Several such modal techniques are reviewed in [9], but these have the drawback that they can only be used for linearizable systems. The same limitation holds for the Krylov subspace methods and many of the SVD-based methods (i.e., described in [10], [11]). Another approach that has been used in nonlinear power systems models is singular perturbation theory [12], in which small time constants are found in a model and used to distinguish slow- and fast-timescale dynamics. These can then be taken to limiting values (i.e., zero) in order to reduce model complexity.

Static and dynamic equivalents are often used when a detailed model is only desired in part of a network (see [13] for a more extensive literature review of static and dynamic equivalents). Static equivalents, such as Ward and REI equivalents (originally proposed in [14] and [15], respectively), replace part of the network with a simplified representation. In some cases this is accomplished via a Kron reduction of the network, where the nodal admittance matrix of the reduced network is the Schur complement, with respect to the nodes being eliminated, of the unreduced nodal admittance matrix [16]. Recently some static equivalencing methods were compared in [17]. Dynamic equivalents complement static equivalents by finding groups of dynamic components that can be replaced by single equivalents (a couple of reviews of dynamic equivalencing techniques can be found at [18],

[19]).

Our approach to improving power systems models focuses on nonlinear systems and combines system identification with model reduction. We use tools of information geometry (information theory combined with differential geometry) to first find, then remove both dynamic and network parameters that are not identifiable from operational data. The parameters that remain not only have reduced uncertainty, which improves overall model predictivity, but also retain physical interpretability, because they can be related back to the parameters of the original model. Accordingly, the resulting reduced models not only are more manageable computationally but can serve as “gray-box” approximations or effective theories that are more amenable to insight, illuminating key phenomenological relationships between parts of the system and eliminating mechanistic detail where it is not needed [20].

In addition, our approach does not impose any particular final structure on the reduced network (as opposed to, e.g., REI equivalents). Rather, the network reduction that occurs depends on the measurement structure imposed at the outset of the reduction process. At the core of our method is the interpretation of a model as a manifold embedded in the space of measurements. This manifold retains all information about model predictions, in contrast to the cost surface in parameter space which condenses the prediction and measurement vectors into a single number [21].

This paper improves on and extends the work presented in [22] and [23]. We have previously applied techniques of information geometry to some power systems components in [24], [25], [26]. Here we explore global identifiability properties of a networked system with 58 unknown component and network parameters. Our previous results showed that the level of reduction achieved depends on the observations available or chosen in the system. In this paper we illustrate the data-driven nature of our approach by considering several possible sets of partial measurements and comparing with prior results.

In Sec. II, we formulate the parameter identifiability problem, and Sec. III outlines our model reduction method. We describe the test system and measurement schemes that we used for this study in Sec. IV. Results for each of the measurement schemes are given in Sec. V, and Sec. VI presents our conclusions.

II. PROBLEM FORMULATION

Equation (1) represents a generic transient stability model in differential-algebraic form [27]:

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

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 observables. In system identification, the parameters \mathbf{p} are unknown and are 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 [21]. In particular, the model’s predictions may be insensitive to coordinated changes in certain combinations of parameters, making them *practically* unidentifiable [28]. This insensitivity is measured by the eigenvalues of the Fisher Information Matrix (FIM). In order to construct the FIM, it is necessary to calculate the parametric sensitivities using the sensitivity equations,

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

To illustrate, the eigenvalues of the FIM for the model introduced in Sec. IV-A are shown in Fig. 1a. Small eigenvalues indicate that the associated parameter combinations will be unidentifiable, but more importantly that they have little effect on model predictions. When these parameter combinations are removed from the model, parameters that are retained are able to explicitly control model behavior. Our method for removing parameters combination, discussed in Sec. III, is known as the Manifold Boundary Approximation Method (MBAM). It was introduced to the physics community in [29] and adapted to power systems applications in [24], [25], [26].

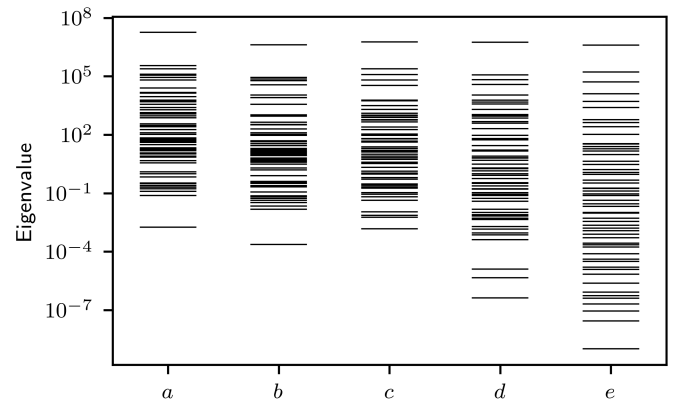


Fig. 1. Eigenvalues of the FIM, prior to the model reduction process, for the following sets of observations (see Sec. IV-B for details): a “full”; b “half”; c “Area A”; d “Area B”; e “supplier-consumer”. When variables are eliminated from observation (going from a to any of b, c, d, or e), the eigenvalues decrease because there is less information available about the parameters.

Sensitivity of model parameters depends strongly on what is being measured in the system. Changing the measurement structure (and, accordingly, the observation function \mathbf{h}) changes the eigenvalues of the FIM. This is illustrated in Fig. 1b through e, where eigenvalues of the FIM are shown for various partial measurement structures (described in full in Secs. IV-B). In every case, the partial set of measurements carries less information than the full set, so the eigenvalues

decrease (compare a to each of the others). This allows greater reduction to be achieved while maintaining model performance in those parts of the system that are of interest to analysts.

III. MANIFOLD BOUNDARY APPROXIMATION METHOD

We remove unidentifiable parameter combinations using the Manifold Boundary Approximation Method (MBAM) [29]. What follows is a summary of the method; for a detailed explanation in the context of power systems, see [24]. Equation (1) can be interpreted as a mapping from the parameter vector \mathbf{p} in parameter space to a second space known as *data space*. This mapping defines a manifold in data space, where the parameters act as coordinates on the manifold (see Fig. 2 for an illustration). Model manifolds are usually bounded, with a hierarchical structure like a polygon (faces, edges, etc.). Each boundary cell corresponds to a reduced model with fewer parameters. We identify these boundary cells by constructing geodesics (the analogs of straight lines on curved surfaces) on the model manifold.

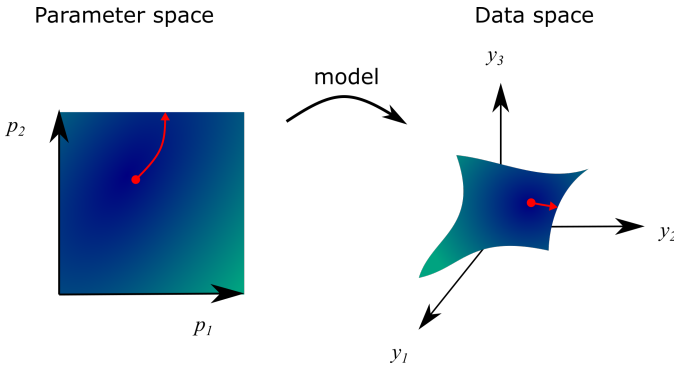


Fig. 2. Illustration of parameter space (on the left) and the model manifold in data space (on the right), both colored by the objective function used for system identification. The mapping from parameter space to data space is defined by the model [e.g., (1)]. A geodesic on the model manifold corresponds to some path through parameter space (red curves). When the geodesic reaches the manifold boundary, some combination of parameters goes to extreme values (see Fig. 3).

When a geodesic encounters the boundary of the manifold, one or more parameters become extreme (go to zero or infinity; see Fig. 3 for an example), corresponding to a limit that can be formally evaluated in the model. The result is a simplified model with one less parameter. This model is then tuned to match the predictions of the original model. We repeat this process of constructing geodesics, evaluating limits, and tuning the reduced model until the predictions of the reduced model no longer faithfully reproduce those of the original.

The procedure is data-driven in the sense that the least sensitive parameter combinations are removed first, so all behavior predicted by the model (e.g., participating modes of the system) that would be measurable at a given level of uncertainty is preserved. This approach unifies concepts from other methods of model reduction (such as singular perturbation approximation and dynamic equivalents) in a single framework while at the same time allowing us to find approximations that have not previously been considered.

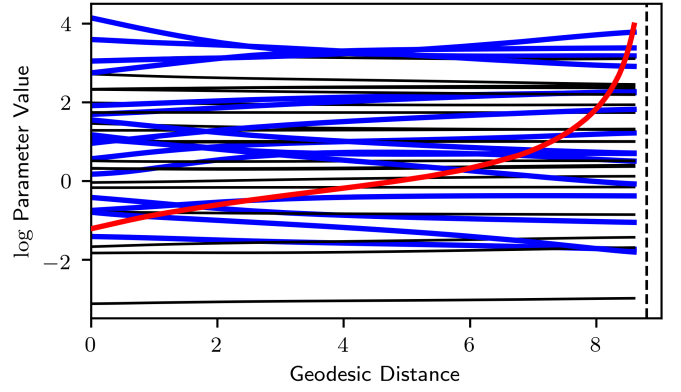


Fig. 3. Parameter values along a geodesic from the “area A” model reduction (see Sec. V-C). Some parameters adjust (blue) while others remain mostly unchanged (black). One goes to infinity (red), encountering a singularity (dashed line, approximate location) when the geodesic reaches a manifold boundary. This indicates the limit $x_q \rightarrow \infty$ (Step 26 in Table II).

IV. APPLICATION

A. Test system

We apply MBAM to the IEEE 14-bus test system (Fig. 4) with synchronous generators (SG) on Buses 1, 2, 3, 6, and 8. These are modeled as follows:

- Bus 1: a fourth-order model (rotor angle, speed, and transient electromotive forces in the d - and q -axes).
- Buses 2 and 3: a classical, second-order model (rotor angle and speed only).
- Buses 6 and 8: a sixth-order model (rotor angle, speed, and both transient and subtransient electromotive forces in the d - and q -axes).

Some parameters (such as rotor moments of inertia) are held fixed to predetermined values, so that we have 38 tunable parameters for generator and controller elements.

20 additional parameters come from the network, where we take the susceptance B_{ik} of each network edge as a tunable parameter and model the conductance G_{ik} as being proportional to the susceptance:

$$G_{ik} = c_{ik} B_{ik}. \quad (3)$$

We discuss the significance of this choice in Sec. V-A. The result is a total of 58 tunable parameters, for both network elements and dynamic components.

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 a more realistic scenario than the one considered in [22], where the mechanical power seen by each generator was simultaneously increased at $t = 0$ s. In order to include the full decay of transients back to the steady state, we also consider longer observation times than in [22], going all the way to $t = 100$ s. 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.

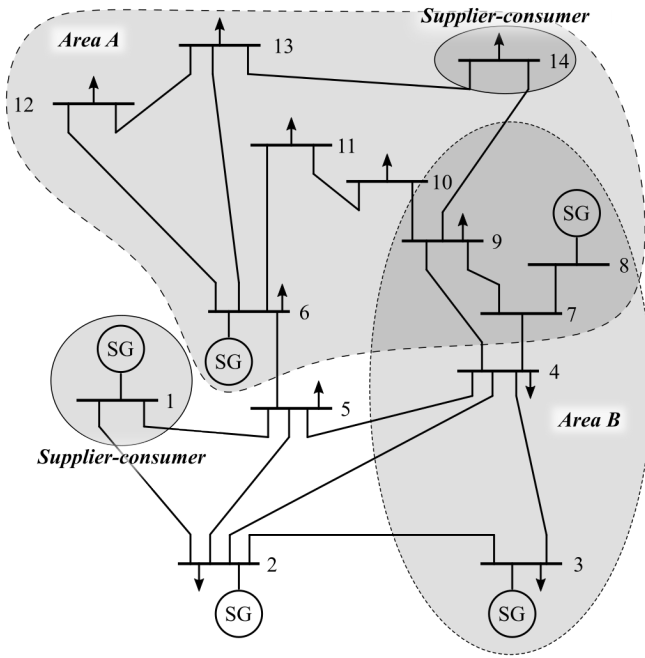


Fig. 4. IEEE 14-bus test system, with observation areas marked.

Transients at Bus 8 from the original model as well as from the “Area A” reduction (see Sec. IV-B) are shown in Fig. 5. Notice that, aside from a few visible deviations, the match is quite good. The deviations indicate that all of the parameters of the model are now identifiable and the reduction process is complete.

B. Partial response matching

We consider several different measurement schemes for the test system described in Sec. IV-A.

The first is a “full” set of observables, which includes the rotor angle and speed and both real and reactive powers of all generators, as well as voltage magnitude and angle on all buses. Preliminary results of applying MBAM for this set of observations were given in [22], [23]. We highlight the results from [23] in Sec. V-B for comparison with the other measurement schemes described below.

In the second, we still include observations from the entire network, but only half of the variables from the full set (specifically, only real and reactive powers of all generators and voltage magnitude on all buses). Our reduction results for this “half” set of observations are discussed in Sec. V-B.

In the remaining measurement schemes, only part of the network is observed, as indicated in Fig. 4. The third and fourth include all variables from the full set that are in Areas A (Buses 7-14 and associated generators) and B (Buses 3, 4, and 7-9 and associated generators), respectively. We discuss results for these reductions in Sec. V-C.

For the final measurement scheme we chose a very sparse set of observations which might be reminiscent of a supplier-consumer relationship, consisting of only Buses 1 and 14

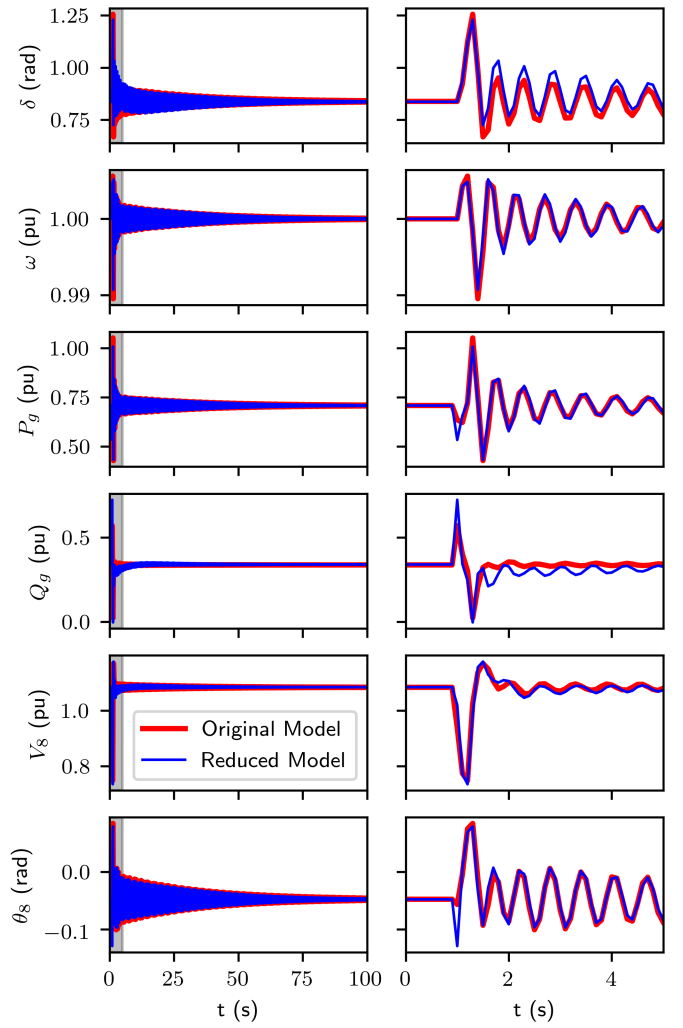


Fig. 5. Transients in generator and bus variables at Bus 8, following a short circuit in Bus 14 at $t = 1$ s (cleared at $t = 1.25$ s), for both the original model and the reduced model obtained when observing only Area A (see Sec. IV-B). The shaded regions from the left-hand plots are plotted on the right for detail. For most of the transients (especially those not pictured), there is still good agreement between the original and reduced models. The few deviations that are visible here indicate that all parameters are identifiable and the reduction process is complete.

(including the generator on Bus 1). Results for this set were previously given in [23]; they are repeated in Sec. V-D for comparison with the previous choices of observations.

V. RESULTS AND DISCUSSION

A. Circuit Theory Analysis of Network Reduction

In our reduction results for all five measurement structures, network reductions occurred due to the following two parameter limits involving the line susceptances: $B_{ik} \rightarrow 0$ and $B_{ik} \rightarrow \infty$. In this section, we present a formal analysis of these limits and show their relation to other reduction techniques (such as Kron reduction).

Recall our choice to let the conductances be proportional to the susceptances, $G_{ik} = c_{ik} B_{ik}$. This means that varying

B_{ik} changes the magnitude of the complex admittance $Y_{ik} = G_{ik} + jB_{ik}$, but not its argument:

$$|Y_{ik}| = B_{ik} \sqrt{1 + c_{ik}^2} \quad (4)$$

$$\arg(Y_{ik}) = \arctan\left(\frac{1}{c_{ik}}\right) \quad (5)$$

Thus the two limits for B_{ik} are equivalent to $|Y_{ik}| \rightarrow 0$ and $|Y_{ik}| \rightarrow \infty$. This is effectively the same as taking the length of the line either to infinity (disconnecting the line) or to zero (inserting a short between the two nodes), respectively.

Formal application of the limit $|Y_{ik}| \rightarrow 0$ to the nodal admittance matrix \mathbf{Y}_{bus} is straightforward. The two off-diagonal elements Y_{ik} and Y_{ki} become zero (assuming \mathbf{Y}_{bus} is symmetric) and the corresponding term in the two diagonal elements $Y_{ii} = -\sum_{n \neq i} Y_{ni}$ and $Y_{kk} = -\sum_{n \neq k} Y_{nk}$ drops out of the sum. As mentioned in [22], the new nodal admittance matrix \mathbf{Y}'_{bus} is

$$\mathbf{Y}'_{\text{bus}} = \mathbf{Y}_{\text{bus}} + Y_{ik} \mathbf{e}_{ik} \mathbf{e}_{ik}^T, \quad (6)$$

where \mathbf{e}_{ik} is the N_{bus} -dimensional branch-bus incidence vector whose i th element is +1, k th element is -1, and other elements are 0.

The case where $|Y_{ik}| \rightarrow \infty$ is not as straightforward, because the four elements Y_{ik} , Y_{ki} , Y_{ii} , and Y_{kk} all diverge in this limit. However, we can treat this case using a Kron reduction as follows.

First, we take the current balance equations $\mathbf{I} = \mathbf{Y}_{\text{bus}} \mathbf{V}$ and reorder them so that the injected current I_k and voltage V_k on node k are the last elements in the vectors \mathbf{I} and \mathbf{V} and Y_{kk} is in the last row and column of \mathbf{Y}_{bus} . We can then partition $\mathbf{I} = \mathbf{Y}_{\text{bus}} \mathbf{V}$ as

$$\begin{bmatrix} \mathbf{I}_A \\ I_k \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{AA} & \mathbf{Y}_{Ak} \\ \mathbf{Y}_{Ak}^T & Y_{kk} \end{bmatrix} \begin{bmatrix} \mathbf{V}_A \\ V_k \end{bmatrix}, \quad (7)$$

where

$$\mathbf{Y}_{AA} = \begin{bmatrix} Y_{11} & \cdots & Y_{1i} & \cdots \\ \vdots & \ddots & \vdots & \ddots \\ Y_{1i} & \cdots & Y_{ii} & \cdots \\ \vdots & \ddots & \vdots & \ddots \end{bmatrix} \quad \text{and} \quad \mathbf{Y}_{Ak} = \begin{bmatrix} Y_{1k} \\ \vdots \\ Y_{ik} \\ \vdots \end{bmatrix} \quad (8)$$

and we have assumed \mathbf{Y}_{bus} is symmetric. Equation (7) provides of system of two matrix equations:

$$\mathbf{I}_A = \mathbf{Y}_{AA} \mathbf{V}_A + \mathbf{Y}_{Ak} V_k \quad (9)$$

$$I_k = \mathbf{Y}_{Ak}^T \mathbf{V}_A + Y_{kk} V_k. \quad (10)$$

Solving (10) for V_k , we have

$$\begin{aligned} V_k &= Y_{kk}^{-1} (I_k - \mathbf{Y}_{Ak}^T \mathbf{V}_A) \\ &= \frac{I_k - Y_{1k} V_1 - \cdots - Y_{ik} V_i - \cdots}{-Y_{1k} - \cdots - Y_{ik} - \cdots} \end{aligned} \quad (11)$$

In the limit that $|Y_{ik}| \rightarrow \infty$, L'Hôpital's rule gives

$$V_k = V_i, \quad (12)$$

which shows that the voltages in the two nodes become identical in this limit. This is the same result obtained in [23] by a different method.

Returning to (9), we eliminate the dependence on V_k by using (11) to obtain

$$\mathbf{I}_A - \mathbf{Y}_{Ak} Y_{kk}^{-1} I_k = (\mathbf{Y}_{AA} - \mathbf{Y}_{Ak} Y_{kk}^{-1} \mathbf{Y}_{Ak}^T) \mathbf{V}_A, \quad (13)$$

where we now see the Schur complement of Y_{kk} in \mathbf{Y}_{bus} appearing on the right. After taking the limit $|Y_{ik}| \rightarrow \infty$, (13) will give us a new set of current balance equations for the network with node k eliminated:

$$\mathbf{I}' = \mathbf{Y}'_{\text{bus}} \mathbf{V}_A \quad (14)$$

$$\mathbf{I}' \equiv \lim_{|Y_{ik}| \rightarrow \infty} \mathbf{I}_A - \mathbf{Y}_{Ak} Y_{kk}^{-1} I_k \quad (14a)$$

$$\mathbf{Y}'_{\text{bus}} \equiv \lim_{|Y_{ik}| \rightarrow \infty} (\mathbf{Y}_{AA} - \mathbf{Y}_{Ak} Y_{kk}^{-1} \mathbf{Y}_{Ak}^T). \quad (14b)$$

To show this, we first examine $\mathbf{Y}_{Ak} Y_{kk}^{-1}$:

$$\lim_{|Y_{ik}| \rightarrow \infty} \mathbf{Y}_{Ak} Y_{kk}^{-1} = \lim_{|Y_{ik}| \rightarrow \infty} \begin{bmatrix} \frac{Y_{1k}}{-Y_{1k} - \cdots - Y_{ik} - \cdots} \\ \vdots \\ \frac{Y_{ik}}{-Y_{1k} - \cdots - Y_{ik} - \cdots} \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ -1 \\ \vdots \end{bmatrix}. \quad (15)$$

Defining \mathbf{e}_i as an $(N_{\text{bus}} - 1)$ -dimensional vector whose i th element is +1 and other elements are 0 (note the difference from the \mathbf{e}_{ik} vector defined earlier), \mathbf{I}' becomes

$$\mathbf{I}' = \mathbf{I}_A + \mathbf{e}_i I_k = \begin{bmatrix} I_1 \\ \vdots \\ I_i + I_k \\ \vdots \end{bmatrix}, \quad (16)$$

indicating that the current injections from node k get moved to node i .

Next, we examine $\mathbf{Y}_{AA} - \mathbf{Y}_{Ak} Y_{kk}^{-1} \mathbf{Y}_{Ak}^T$:

$$\begin{aligned} \mathbf{Y}_{AA} - \mathbf{Y}_{Ak} Y_{kk}^{-1} \mathbf{Y}_{Ak}^T &= \\ \begin{bmatrix} Y_{11} + \frac{Y_{1k}^2}{Y_{1k} + \cdots + Y_{ik} + \cdots} & \cdots & Y_{1i} + \frac{Y_{1k} Y_{ik}}{Y_{1k} + \cdots + Y_{ik} + \cdots} & \cdots \\ \vdots & \ddots & \vdots & \ddots \\ Y_{1i} + \frac{Y_{1k} Y_{ik}}{Y_{1k} + \cdots + Y_{ik} + \cdots} & \cdots & Y_{ii} + \frac{Y_{ik}^2}{Y_{1k} + \cdots + Y_{ik} + \cdots} & \cdots \\ \vdots & \ddots & \vdots & \ddots \end{bmatrix}. \end{aligned} \quad (17)$$

The off-diagonal elements become

$$\lim_{|Y_{ik}| \rightarrow \infty} Y_{nm} + \frac{Y_{nk} Y_{mk}}{Y_{1k} + \cdots + Y_{ik} + \cdots} = Y_{nm}, \quad (18)$$

for $n \neq m$, $n \neq i$, and $m \neq i$ and

$$\lim_{|Y_{ik}| \rightarrow \infty} Y_{ni} + \frac{Y_{nk} Y_{ik}}{Y_{1k} + \cdots + Y_{ik} + \cdots} = Y_{ni} + Y_{nk}, \quad (19)$$

for $n \neq i$; and the diagonal elements become

$$\lim_{|Y_{ik}| \rightarrow \infty} Y_{nn} + \frac{Y_{nk}^2}{Y_{1k} + \dots + Y_{ik} + \dots} = Y_{nn} \quad (20)$$

for $n \neq i$ and

$$\begin{aligned} Y_{ii} + \frac{Y_{ik}^2}{Y_{1k} + \dots + Y_{ik} + \dots} \\ &= - \sum_{n \neq i, k} Y_{ni} - Y_{ik} + \frac{Y_{ik}^2}{Y_{1k} + \dots + Y_{ik} + \dots} \\ &= - \sum_{n \neq i, k} Y_{ni} + \frac{-Y_{1k}Y_{ik} - \dots - Y_{ik}^2 - \dots + Y_{ik}^2}{Y_{1k} + \dots + Y_{ik} + \dots} \\ &= - \sum_{n \neq i, k} Y_{ni} - \frac{Y_{ik}}{Y_{1k} + \dots + Y_{ik} + \dots} \sum_{n \neq i, k} Y_{nk} \quad (21) \\ \lim_{|Y_{ik}| \rightarrow \infty} Y_{ii} + \frac{Y_{ik}^2}{Y_{1k} + \dots + Y_{ik} + \dots} &= - \sum_{n \neq i, k} Y_{ni} + Y_{nk}; \quad (22) \end{aligned}$$

so as with the current injections, all of the branches connected to node k also get moved to node i . The new nodal admittance matrix is

$$\mathbf{Y}'_{\text{bus}} = \mathbf{Y}_{AA} + \sum_{l=1}^3 \mathbf{Y}_l \quad (23)$$

where the \mathbf{Y}_l are three $(N_{\text{bus}} - 1) \times (N_{\text{bus}} - 1)$ matrices with 1) \mathbf{Y}_{Ak} as the i th column and zeros elsewhere, 2) \mathbf{Y}_{Ak}^T as the i th row and zeros elsewhere, and 3) Y_{kk} as the (i, i) th element and zeros elsewhere.

For the power balance equations $\mathbf{S} = \text{diag}(\mathbf{V})\mathbf{I}^*$, partitioning as before leads to

$$\mathbf{S}_A = \text{diag}(\mathbf{V}_A)\mathbf{I}_A^* \quad (24)$$

$$\mathbf{S}_k = \mathbf{V}_k \mathbf{I}_k^*. \quad (25)$$

The new power injection vector is

$$\begin{aligned} \mathbf{S}' &= \text{diag}(\mathbf{V}_A)\mathbf{I}'^* \\ &= \text{diag}(\mathbf{V}_A)(\mathbf{I}_A + \mathbf{e}_i \mathbf{I}_k)^* \\ &= \mathbf{S}_A + \mathbf{e}_i \mathbf{V}_i \mathbf{I}_k^* \\ &= \mathbf{S}_A + \mathbf{e}_i \mathbf{S}_k \quad (26) \end{aligned}$$

where we used (24), (12), and (25). The power injections on node k are simply moved to node i .

B. “Full” and “half” observations

Using MBAM, we were able to reduce the model down to 37 parameters with “full” observations [23]. Switching to “half” observations allowed a couple more reductions, for a total of 35 parameters, which is equal to what was predicted in [22]. Table I shows the reductions that occurred when using “half” observations; those that also occurred with “full” observations are marked with an asterisk. The most notable difference between these two reductions is the limit $B_{9,10} \rightarrow \infty$, which indicates merging Buses 9 and 10 (see [23] for more on merging buses). This is interesting because the voltage magnitudes of both buses are being observed, yet we

are still able to combine them without introducing noticeable deviations from the original model. The reduced network for “half” observations is shown in Fig. 6.

In contrast with [22], we found that parameter limits involving time constants were not always paired with removal of reactances, as would be the case in, e.g., a typical singular perturbation reduction. This is because we initialized the system in steady state, making the initial conditions (ICs) parameter-dependent. Specifically, while the ICs did depend on the reactances, they did not depend on time constants. To illustrate, consider the following equation [30] for the transient emf e'_q :

$$T'_{d0} \dot{e}'_q = -e'_q - (x_d - x'_d)i_d + v_f. \quad (27)$$

Steady state implies $\dot{e}'_q = 0$, so the ICs are (partially, along with the other equations of the model) determined by the constraint

$$0 = -e'_q - (x_d - x'_d)i_d + v_f, \quad (28)$$

which does not depend on T'_{d0} but does depend on x_d and x'_d . MBAM naturally picks up on this subtlety because it is reflected in the structure of the model manifold which MBAM navigates. A more in-depth exploration of various power systems models using MBAM would likely reveal additional such subtleties in the parameter reductions of the model.

TABLE I
REDUCTIONS WITH “HALF” OBSERVATIONS.

Step	Reduction	Location	Step	Reduction	Location
1	* $T'_{q0} \rightarrow \infty$	Bus 8	13	* $T''_{q0} \rightarrow 0$	Bus 8
2	* $T'_{q0} \rightarrow \infty$	Bus 6	14	* $x'_q \rightarrow x'_q$	Bus 8
3	* $x_q \rightarrow \infty$	Bus 8	15	* $x_d \rightarrow x'_d$	Bus 6
4	* $T'_{d0} \rightarrow \infty$	Bus 8	16	* $T'_{d0} \rightarrow \infty$	Bus 6
5	* $x_q \rightarrow \infty$	Bus 6	17	* $K_e \rightarrow 0$	Bus 6
6	* $x_d \rightarrow \infty$	Bus 8	18	* $K_a \rightarrow 0$	Bus 8
7	* $K_e \rightarrow 0$	Bus 8	19	* $K_a \rightarrow 0$	Bus 6
8	* $B_{12,13} \rightarrow 0$	Line 12-13	20	* $x_d \rightarrow x'_d$	Bus 1
9	$B_{9,10} \rightarrow \infty$	Line 9-10	21	$T''_{q0}, x'_q, e'_d \rightarrow \infty$	Bus 6
10	* $x''_d \rightarrow 0$	Bus 6	22	$x'_q \rightarrow 0$	Bus 1
11	* $B_{2,5} \rightarrow 0$	Line 2-5	23	* $T''_{d0} \rightarrow 0$	Bus 6
12	* $x''_d \rightarrow 0$	Bus 8			

*Also occurred with “full” observations.

We conducted a sensitivity analysis of the reduced models with “full” and “half” observations and found that our reduction method has effectively removed only the smallest eigenvalues of the FIM (see Fig. 7). The parameters that remain are the ones that are most significant in determining the behavior of the model.

C. Areas A and B

In large systems, engineers are usually most interested in the behavior of only part of the system, called the *internal* or *study* system. The *external* system, which is of less interest, can be reduced using static and dynamic equivalents, as long as its effects on the study system are retained [13]. By

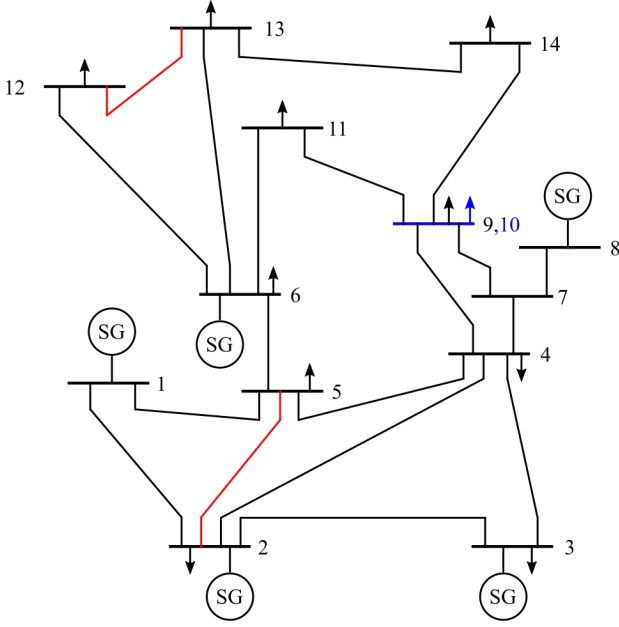


Fig. 6. Reduced network for “half” observations. Branches marked in red were removed during the reduction process for both “full” and “half” observations. Buses 9 and 10 were merged (blue) only in the “half” reduction.

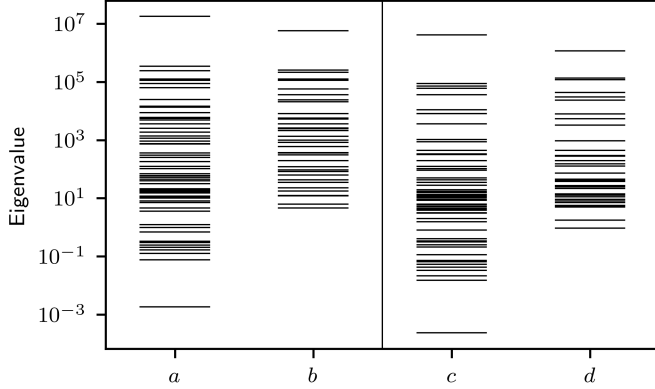


Fig. 7. Eigenvalues of the FIM for “full” (a and b) and “half” (c and d) observations; the reduced models are b and d. The smallest eigenvalues have been effectively “erased,” so the most identifiable parameters have been retained.

including observables only from the study system, MBAM can be used to simultaneously eliminate unidentifiable parameters in the study system and obtain equivalents for the external system. We demonstrate this using the “Area A” and “Area B” measurement regions shown in Fig. 4.

With “Area A” measurements, we were able to reduce the model down to 31 parameters - significantly fewer than predicted in [22]. The “Area B” measurements were not considered in [22]; in this case the model was reduced down to 25 parameters. Reduction steps are shown in Tables II and III, respectively. Notably, compared to the “full” or “half” measurement structures considered in Sec. V-B, much more

reduction occurred in both of these reductions in the generator on Bus 1, which is not being observed in either area. In addition, Area B had more reduction in the generator on Bus 6 than did Area A. Both of these points are consistent with the expected lack of need to retain detail outside the observed area.

TABLE II
REDUCTIONS WITH “AREA A” OBSERVATIONS.

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

TABLE III
REDUCTIONS WITH “AREA B” OBSERVATIONS.

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

Particularly striking are the network reductions that occurred in each of the “Area A” and “Area B” reductions. These can be seen in Figs. 8 and 9. Although significant external network reduction occurred, in both cases buses were retained that could be considered “far” from the observed region because they are not directly connected to any of the observed buses (notably, Buses 1-3 in the “Area A” reduced network and Bus

1 in the “Area B” reduced network). This indicates that some external network structure can be “seen” through its effects on the study system, even when it is not being directly observed. We note that this may be related to the presence of generators on these external buses.

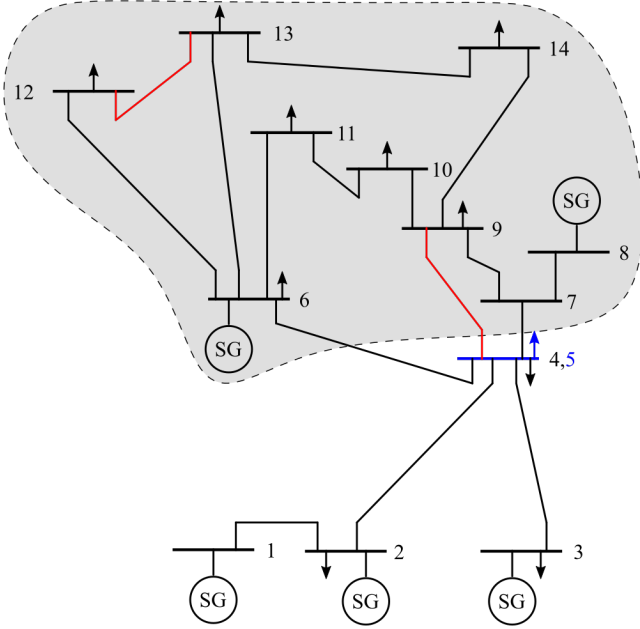


Fig. 8. Reduced network for “Area A” observations (shaded region). Branches that were removed during the reduction from inside the observed area are marked in red; those removed from outside have been omitted for clarity. Buses (and associated components) that were merged during the reduction are marked in blue.

As in the case of “half” observations, in “Area B” we find that some external buses have merged into those being observed (Buses 9 and 4). We also note that Bus 14, where the short occurs, has now been merged with other external buses in “Area B,” indicating that in general, the exact location of a short in the external system may not be discernible from the available measurements and so may be irrelevant.

We conducted a sensitivity analysis of the “Area A” and “Area B” reduced models and confirmed that the small eigenvalues of the FIM have been eliminated (see Fig. 10).

D. Supplier-consumer observations

The final measurement scheme includes observations on only two buses, 1 and 14, as well as the generator on Bus 1. This choice allows for maximal network reduction to occur in the system. We expect the results obtained for this set of observations to be particularly indicative of the level of network reduction that could be achieved for large power systems where only portions of the system are being directly observed.

We saw the most reduction for this set of observations, with the final reduced model consisting of 20 parameters (see Table IV for reduction steps). This matches the prediction made in [22].

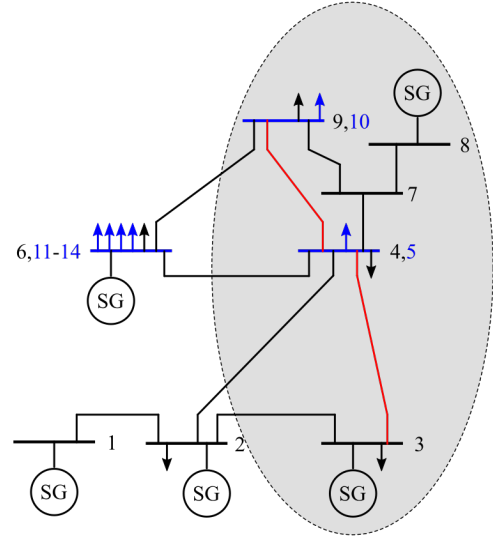


Fig. 9. Reduced network for “Area B” observations (shaded region). Branches that were removed during the reduction from inside the observed area are marked in red; those removed from outside have been omitted for clarity. Buses (and associated components) that were merged during the reduction are marked in blue.

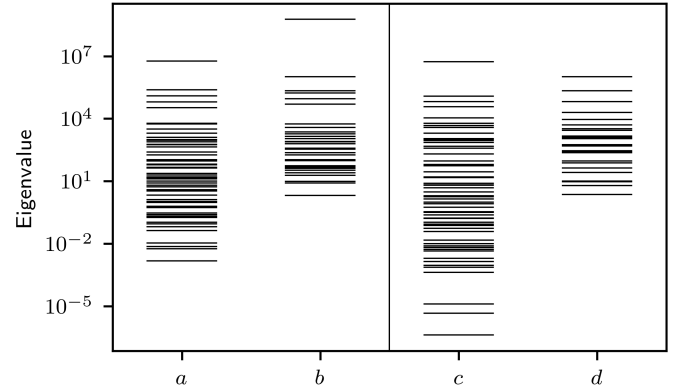


Fig. 10. Eigenvalues of the FIM for “Area A” (a and b) and “Area B” (c and d) observations; the reduced models are b and d.

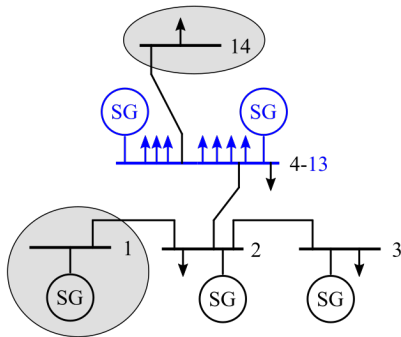
As with Areas A and B (see Sec. V-C), we did not see a complete reduction of the unobserved parts of the network (see Fig. 11. A “backbone” consisting of Buses 1-3, each with its own generator, remains essentially unchanged in all three reductions; but it is also noteworthy that in this reduction, the other two buses with generators have merged.

A sensitivity analysis on the 20-parameter reduced model reveals that many of the small eigenvalues have again been eliminated (see Fig. 12).

VI. CONCLUSION

In this paper, we have demonstrated capabilities of the Manifold Boundary Approximation Method (MBAM) to simultaneously reduce dynamical components and network structure and perform parameter estimation in the IEEE 14-bus test

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



system. These efforts improve on previous studies [22], [23] by using a more realistic fault scenario and longer observation times and by applying MBAM in several new measurement scenarios.

more limited sets of observations.

Accordingly, MBAM is a potentially useful way of deriving equivalents, in models of large power networks, that strike a better balance between model simplicity and accuracy. These equivalents represent an effective network of interactions that is sufficient to predict the observed dynamical behavior while eliminating unnecessary detail. Such effective networks are useful not only because of reduced computational effort and model uncertainty but as a tool for developing insight about the collective system-level behavior.

Finally, we identified several types of parameter limits that can be used in power systems models with similar mathematical structures. Identifying these limits will be crucial in our ongoing efforts to scale up these methods to large systems.

- [1] G. Andersson, P. Donalek, R. Farmer, N. Hatziaargyriou, 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] J. Sanchez-Gasca, C. Bridenbaugh, C. Bowler, and J. Edmonds, "Trajectory sensitivity based identification of synchronous generator and excitation system parameters," *IEEE Transactions on Power Systems*, vol. 3, no. 4, pp. 1814–1822, 1988.
- [3] S. M. Benchluch and J. H. Chow, "A trajectory sensitivity method for the identification of nonlinear excitation system models," *IEEE Transactions on Energy Conversion*, vol. 8, no. 2, pp. 159–164, 1993.
- [4] I. A. Hiskens, "Nonlinear dynamic model evaluation from disturbance measurements," *IEEE Transactions on Power Systems*, vol. 16, no. 4, pp. 702–710, 2001.

- [5] 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.
- [6] A. Van Der Schaft and B. Maschke, "Port-hamiltonian systems on graphs," *SIAM Journal on Control and Optimization*, vol. 51, no. 2, pp. 906–937, 2013.
- [7] 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.
- [8] J. Yu, Y. Weng, and R. Rajagopal, "Patopa: A data-driven parameter and topology joint estimation framework in distribution grids," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4335–4347, 2017.
- [9] 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.
- [10] 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.
- [11] 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>
- [12] 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.
- [13] F. Milano and K. Srivastava, "Dynamic rei equivalents for short circuit and transient stability analyses," *Electric Power Systems Research*, vol. 79, no. 6, pp. 878–887, 2009.
- [14] J. B. Ward, "Equivalent circuits for power-flow studies," *Electrical Engineering*, vol. 68, no. 9, pp. 794–794, 1949.
- [15] P. Dimeo, *Nodal analysis of power systems*. Kent, UK: Abacus Press, 1975.
- [16] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2012.
- [17] S. M. Ashraf, B. Rathore, and S. Chakrabarti, "Performance analysis of static network reduction methods commonly used in power systems," in *2014 Eighteenth National Power Systems Conference (NPSC)*. IEEE, 2014, pp. 1–6.
- [18] U. Annakkage, N.-K. C. Nair, Y. Liang, A. Gole, V. Dinavahi, B. Gustavsen, T. Noda, H. Ghasemi, A. Monti, M. Matar *et al.*, "Dynamic system equivalents: A survey of available techniques," *IEEE Transactions on Power Delivery*, vol. 27, no. 1, pp. 411–420, 2011.
- [19] X. Lei, D. Povh, and O. Ruhle, "Industrial approaches for dynamic equivalents of large power systems," in *2002 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No. 02CH37309)*, vol. 2. IEEE, 2002, pp. 1036–1042.
- [20] M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna, "Perspective: Sloppiness and emergent theories in physics, biology, and beyond," *The Journal of chemical physics*, vol. 143, no. 1, p. 07B201_1, 2015.
- [21] 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.
- [22] 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.
- [23] B. L. Francis, J. R. Nuttall, M. K. Transtrum, A. T. Sarić, and A. M. Stanković, "Network reduction in transient stability models using partial response matching," in *2019 North American Power Symposium (NAPS)*. IEEE, 2019, pp. 1–6.
- [24] 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.
- [25] —, "Information geometry approach to verification of dynamic models in power systems," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 440–450, 2017.
- [26] A. T. Sarić, M. K. Transtrum, and A. M. Stanković, "Information geometry for model identification and parameter estimation in renewable energy—dfig plant case," *IET Generation, Transmission & Distribution*, vol. 12, no. 6, pp. 1294–1302, 2017.
- [27] P. Kundur, *Power system stability and control*. McGraw-Hill New York, 1994, vol. 7.
- [28] 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.
- [29] M. K. Transtrum and P. Qiu, "Model reduction by manifold boundaries," *Physical Review Letters*, vol. 113, no. 9, p. 098701, 2014.
- [30] F. Milano, *Power system modelling and scripting*. Springer Science & Business Media, 2010.

Chapter 5

Piecemeal Reduction of Large Systems

5.1 Introduction

Mathematical models from many fields of study involve large, interconnected network structures. Often these models stretch and go beyond the limits of current computing capabilities, available measurement data, and intuitive comprehension. For these reasons many efforts have been put toward developing model reduction techniques. In particular, there is a need for reduction methods that preserve physical interpretability of the reduced model, especially its interpretation as a network of interconnected components. In some cases a model of a system may be so large that it is impractical or even impossible to reduce as a whole with current methods. In these cases it is necessary (and potentially still beneficial with more manageably-sized models) to have methods for reducing pieces of the model separately and then reassembling them.

Most model reduction methods that have been developed involving interconnected subsystems or network structure have been applied to whole systems simultaneously and not to pieces or subsystems separately (i.e., in open loop). This is because, even when such methods preserve stability or other characteristics in the open-loop subsystem, they do not guarantee preservation

of underlying physical conservation laws, potentially leading to instabilities when the reduced subsystem is reconnected. Still, there are methods that at least preserve all or part of the interconnection structure between subsystems during the reduction process so that subsystems are effectively reduced separately.

Traditional application of such methods as balanced truncation or Krylov subspace projection does not preserve interconnection structure, because state variables of the reduced model are usually linear combinations of the underlying “physical” states. However, modifications of these methods that preserve interconnection structure by structuring the coordinate transformations or projections involved have been explored in [47–52] and demonstrated in industrial advanced process control [53] and power systems [54]. A variation on balanced truncation that reduces identical subsystems in a network without disturbing the network structure and while maintaining homogeneity of the subsystems is presented in [55]. An alternative approach for a one-dimensional arrangement of identical interconnected subsystems that allows for uncertainties (represented as a Δ -block) is taken in [56].

For networks of identical subsystems, the idea of *clustering* has been proposed for preserving partial network structure, in which a system is decomposed into areas (clusters) whose subsystems (nodes) are then aggregated (rather than being individually reduced). Criteria for determining clusters have included cluster reducibility [57–59], almost equitable partitions [60–62], edge importance [63, 64], and vertex/cluster dissimilarity [65, 66]. Most of these methods are only applicable to first- or second-order subsystems or specific network structures (although in [66] both of these restrictions are lifted), and all of them are suitable only for linear systems.

In power systems, the related concept of *area aggregation* has been explored in [67–71] for identical first- and second-order (linear) subsystems and extended to detailed generator models in [72]. This approach is based on *slow coherency*, in which a separation of timescales occurs due to areas being much more densely connected internally than they are to other areas. Area

aggregation and slow coherency fall under the broader scope of *dynamic equivalencing* in power systems (see [73] for a review of dynamic equivalencing methods), which splits the system into a detailed *study area* to be kept and an *external area* to be reduced. How to choose the study area usually depends on the application and on the particular method of reduction.

Although similar to slow coherency, the notion of *synchronic modal equivalencing* (SME) [74] was developed to address situations where there is no clear distinction between slow and fast modes of the system or when modes other than the slowest may be of interest. The equivalents are formed by organizing the system into groups whose motion can be well-approximated by a linear combination of the motions of a set of *basis generators*. For the most part, little reduction is performed on the network itself; although generators which are neither in the study group nor in the chosen basis are replaced with current sources, their buses (nodes) are retained in-place. Both slow coherency and SME require a multi-step procedure for determining which generators to group and how to aggregate them.

One of the challenges of dynamic equivalencing is determining the parameters of the generator equivalents so as to retain the characteristics of the generators being replaced. In [72], explicit formulas relating the parameters of aggregated machines to those of the equivalent are derived assuming the aggregated machines are identical, based on maintaining the mathematical structure of the individual machine models. Alternatively, [75] proposes two methods, one using a fitting algorithm to determine the parameters using aggregated values only as a starting point, and the other based on aggregating impedances determined for certain oscillation frequencies. In the case of SMEs, no aggregation is performed; rather a chosen set of generators is simply retained as-is.

Recent developments in information geometry have led to an alternative approach to model reduction that maintains physical interpretability of the model and is amenable to fully nonlinear systems, known as the Manifold Boundary Approximation Method (MBAM) [29]. By focusing on the tunable parameters of the model, rather than the states, MBAM is able to directly consider

physically-meaningful limiting behaviors of the model and preserve all relevant underlying conservation laws and symmetries of the system. In addition, the MBAM reduction process is data-driven, tailored specifically to maintain fidelity of the chosen set of observations, so detail is naturally kept where it is needed and discarded where it is not.

Conventional application of MBAM proceeds by finding and then applying various parameter limits to the model, reducing the number of parameters by one at each step. Although initially this step-by-step procedure is necessary in order to discover which parameter limits are appropriate and physically meaningful for the particular mathematical structure of the model, for large models with hundreds or thousands of parameters it is impractical. It is also very computationally intensive, requiring not only repeated evaluation of the model but repeated calculation of the Jacobian of the model with respect to the parameters (i.e., parameter sensitivities) for each step of the reduction. A variation of MBAM, known as LinearizedMBAM [76], has been developed that utilizes prior knowledge about the types of viable parameter limits, obtained from applying conventional MBAM to small or moderate systems, to find hundreds of limits for a large model in a single calculation, making reduction of large models much more efficient and tractable. But unfortunately even this is insufficient for reducing models of very large systems having more than tens of thousands of parameters.

Model reduction methods that can be performed on pieces of the system separately while maintaining physical interpretability can be useful, even for moderately-sized systems. In this chapter, we propose a procedure for splitting a model of a large system into pieces, reducing the pieces using MBAM or LinearizedMBAM, and reassembling them to construct a reduced model of the whole system. The outline of the chapter is as follows. All possible sequences of parameter limits that can be applied to a model are captured in a directed graph called a *Hasse diagram*; Sec. 5.2 discusses how the Hasse diagram may be navigated piecewise. Section 5.3 outlines key characteristics a successful piecemeal reduction strategy should have, discusses how these may

be addressed in the context of MBAM, and ends with our proposed strategy. In Sec. 5.4, we demonstrate the proposed strategy on an example network. This is followed by some discussion of our results and concluding remarks in Sec. 5.5.

5.2 Navigating the Hasse diagram

Section 2.1 of Ch. 2 described the model manifold, whose boundary structure may be identified as a *graded poset*¹. In most cases of interest, this poset has the additional structure necessary to identify it as an *abstract polytope*, in which case the poset elements are called *faces*. Abstract polytopes capture the connectivity properties of traditional polytopes (which are the n -dimensional generalizations of polygons and polyhedrons). When a model manifold has a polytopal boundary structure, we will refer to the boundary structure as the *model polytope*; otherwise, we will refer the boundary structure as the *model boundary poset*.

All discrete posets² may be represented visually with a *Hasse diagram* (see Fig. 5.1). At the top of the Hasse diagram for a model boundary poset is a vertex representing the full (unreduced) model with N parameters.³ The other vertices of the Hasse diagram represent reduced models that may be obtained by taking various combinations of parameter limits. Each edge represents a

¹ A poset (partially ordered set) is a set \mathcal{P} together with an ordering relation $\leq_{\mathcal{P}}$. A graded poset is a poset that also has a rank function that maps from the poset to the natural numbers. For graded posets arising from the topological connectivity properties of geometric figures, the rank of an element is the dimension of the corresponding structural component of the figure, e.g., vertices and apices are rank 0, edges are rank 1, etc.

² A discrete poset is a poset whose elements are countable.

³ This is the number of *structurally identifiable* [77] parameters in the model (see also [78]), rather than the number of nominal parameters. A complete characterization of parameter identifiability requires calculation of the full Jacobian J of the model predictions with respect to the parameters. Structurally unidentifiable parameter combinations appear as eigenvectors of $J^T J$ with zero eigenvalue. The total number of structural parameters is independent of how the parameters are defined, so we will refer to individual parameters as well as to parameter combinations that act or can be defined as single parameters simply as parameters.

parameter limit that may be evaluated in the model at that edge's upper vertex, typically reducing the number of parameters by one. The number of parameters in each reduced model determines its rank; all vertices at the same level in the Hasse diagram represent reduced models with the same number of parameters. By finding and implementing parameter limits in the model, MBAM effectively navigates down the Hasse diagram, one edge at a time. Formally, the model reduction problem which MBAM solves can be stated as follows: *Given the Hasse diagram for a model boundary poset, construct a directed path from the top vertex to the lowest possible vertex such that the bias introduced in the model (in some chosen norm) is no greater than a predetermined bound.* The process of constructing a directed path through the Hasse diagram is what we will refer to as *navigating* the Hasse diagram. Typically the norm and bound chosen depend on the particular application and which types of bias are more important to the user.

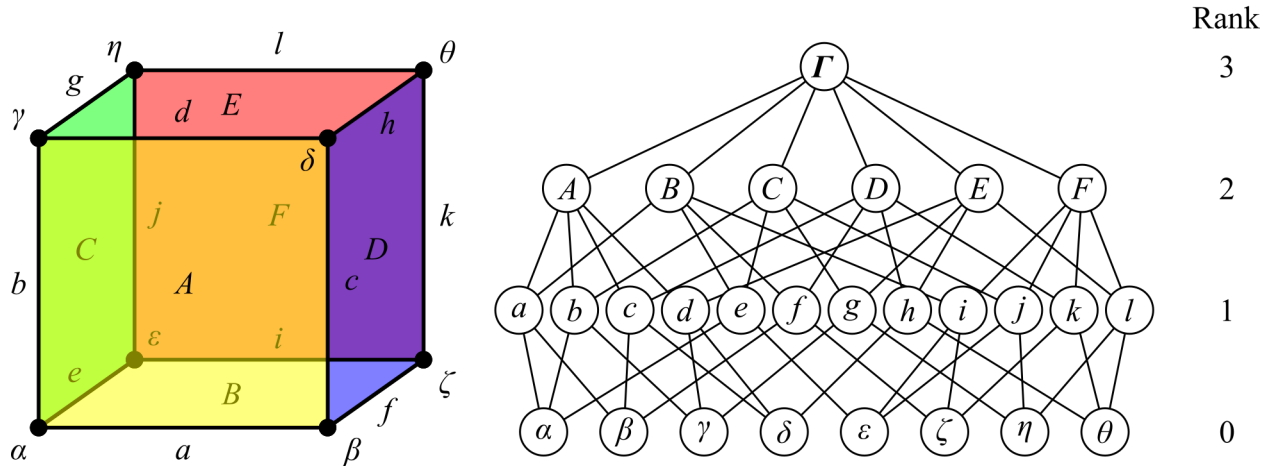


Figure 5.1 *Left:* A cube. Faces are labeled with capital Latin letters, edges with lowercase Latin, and corners with lowercase Greek. The connectivity between these structural elements is captured by a partially ordered set called an abstract polytope. *Center:* Hasse diagram for the cube. Γ represents the whole cube, and the succeeding three levels represent faces, edges, and corners, respectively. *Right:* The rank indicates the dimensionality of each structural element.

As an example, consider the network shown in Fig. 5.2. A model based on this network has a nonnegative parameter for each edge and predicts observations made at nodes (for details, see

Sec. 5.4). The cube in Fig. 5.1 represents the model polytope for such a model; we have redrawn the Hasse diagram from Fig. 5.1 in Fig. 5.2 to illuminate the relationship between Hasse diagram edges and the parameters of the model. Already the diagram is not simple (compared to the corresponding network); generally speaking, the order of the Hasse diagram (its number of vertices) grows much faster than the number of parameters in the model. Large models have incredibly large, complicated Hasse diagrams. Reducing a large system by splitting it into pieces and reducing them separately can be viewed as navigating parts of the Hasse diagram separately and then combining the results.

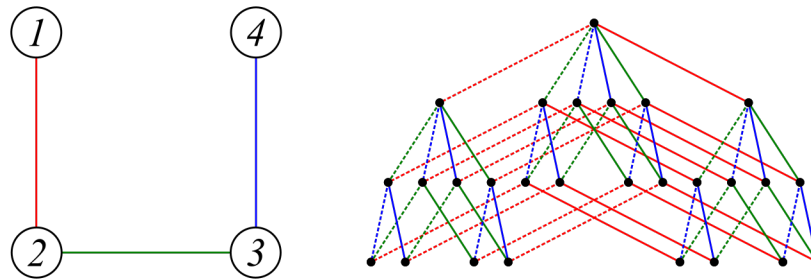


Figure 5.2 Network with four nodes and three edges. A model based on this network that has nonnegative parameters for each edge and makes predictions at nodes would have the cube in Fig. 5.1 as its model polytope. The Hasse diagram from Fig. 5.1 is repeated here on the right, with edges colored to indicate which parameter (i.e., which edge in the network) they correspond to. These edge parameters each have two possible limits (see Sec. 5.3.3): dashed lines represent limits where a network edge is cut or removed, and solid lines represent limits where a network edge is contracted and its two vertices are merged.

This raises the question of how to identify parts of the Hasse diagram that can be navigated independently. Because piecewise navigation involves the construction of multiple paths, the key criterion is that each path be compatible with every other path constructed, i.e., each sequence of parameter limits must be interchangeable with every other sequence of parameter limits. An example of this is shown in Fig. 5.3, which shows both a network associated with a four-parameter model (one parameter for each network edge) and the Hasse diagram for this model. Clearly, any path from the top vertex of the Hasse diagram along only black edges could be replicated from

either of the two orange edges at the top, and likewise both choices for orange are available at every point in the central set of black edges.

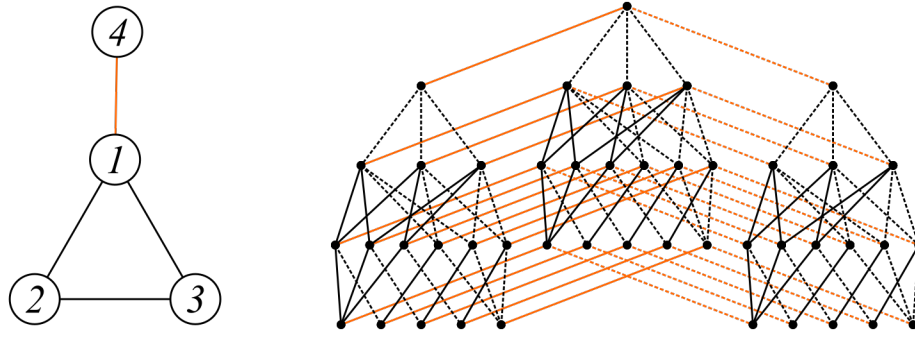


Figure 5.3 *Left:* Network with 4 nodes and 4 edges. Colors distinguish two groups of parameters. *Right:* Hasse diagram for a model based on this network. Starting from the top, any path involving only black edges may be replicated from either of the two orange edges, and vice versa.

This requirement that Hasse diagram paths be compatible has implications for how the parameters in the model should be grouped for piecemeal reduction. For example, it is necessary that these groups of parameters be disjoint to avoid getting conflicting limits for the same parameter. In some cases, simply partitioning the parameters is sufficient, provided that the partition meets certain additional requirements. We discuss such partitions in Sec. 5.2.1. In other cases, it may be necessary to (temporarily) leave some parameters out of the piecemeal reduction process in order to ensure that paths will be compatible. As a result, the combined path will be limited as to how far down the Hasse diagram it can extend, and further reduction may be possible after the initial piecemeal reduction is performed. We discuss partial navigation of the Hasse diagram in Sec. 5.2.2.

5.2.1 Prime posets

As a special case, compatibility of Hasse diagram paths is guaranteed if the model boundary poset is *factorable*, i.e., decomposable as a Cartesian product of (nontrivial) posets. The Cartesian product

of two posets⁴ \mathcal{P} and \mathcal{Q} is defined as

$$\mathcal{P} \times \mathcal{Q} \equiv \{(F, G) | F \in \mathcal{P}, G \in \mathcal{Q}\}, \quad (5.1)$$

where the ordering is given by

$$(F, G) \leq_{\mathcal{P} \times \mathcal{Q}} (F', G') \text{ if and only if } F \leq_{\mathcal{P}} F' \text{ and } G \leq_{\mathcal{Q}} G'. \quad (5.2)$$

Any connected⁵ poset has a unique prime factorization with respect to the Cartesian product [80].⁶

Accordingly, given a poset \mathcal{P} , we may write its prime factorization as

$$\mathcal{P} = \prod_i \mathcal{P}_i. \quad (5.3)$$

It has been shown [79, 81] that if $\mathcal{P} = \prod_i \mathcal{P}_i$, then $H(\mathcal{P}) = \prod_i H(\mathcal{P}_i)$; that is, the Hasse diagram for \mathcal{P} is factorable as the Cartesian product of the Hasse diagrams of the factors \mathcal{P}_i . This means that the Hasse diagrams of the factors $H(\mathcal{P}_i)$ may be navigated separately and the resulting directed paths combined (in any order) to obtain a single directed path through the Hasse diagram $H(\mathcal{P})$ of the full poset. It also means that any parameter limits associated with one prime \mathcal{P}_j are independent of (and compatible with) those associated with all other primes $\mathcal{P}_{i \neq j}$. The network in Fig. 5.2 has a factorable model boundary poset, as indicated by the repeated path structure throughout its Hasse

⁴ For abstract polytopes, several variations of this product, which are discussed in [79], are possible depending on the treatment of the null and greatest faces, including 1) the join product, which includes pairings of all faces of the two polytopes; 2) the Cartesian product, which excludes pairings of the null face of either polytope with the non-null faces of the other, 3) the direct sum, which excludes pairings of the greatest face of either polytope with the non-greatest faces of the other, and 4) the topological product, which excludes pairings of both the greatest and null faces with the others. In the context of models with polytopal boundary structures, the null face has no meaning and is of no use, so pairings of faces in one model polytope with the null face of the other are likewise meaningless. All other pairings are meaningful, so the Cartesian product is the appropriate one for model polytopes.

⁵ A poset \mathcal{P} is *connected* if, for any two elements $F, G \in \mathcal{P}$, there exists a sequence $\{F = F_0, F_1, \dots, F_k = G\}$ such that $F_i \leq F_{i-1}$ or $F_i \geq F_{i-1}$ for every $i \in \{1, \dots, k\}$.

⁶ [79] showed that this also holds for polytopes, which are connected by definition.

diagram. For comparison, an example network whose model boundary poset is prime is shown in Fig. 5.4.

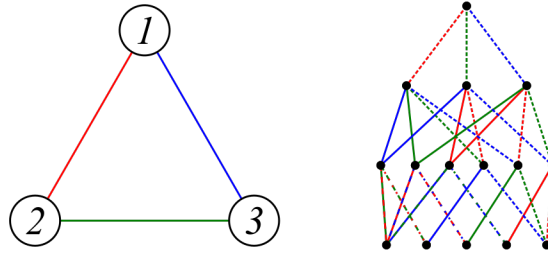


Figure 5.4 Network (left) with the same number of parameters as the one in Fig. 5.2 but slightly different topology, and associated Hasse diagram (right). In contrast with Fig. 5.2, the repeating structure of paths in the Hasse diagram is broken; the corresponding model polytope is not factorable. Some edges near the bottom of the diagram have two colors because it is ambiguous which (bare) parameter remains in the model at that point in the reduction.

We expect that groups of parameters associated with various prime factors in the poset decomposition will be disjoint and lead to a partition of the parameters. Although identifying these groups may, in general, be nontrivial, we can specify a couple of characteristics they possess, which should aid in their discovery. Because the limits of a prime poset's parameters are necessarily compatible with all other parameter limits in the model, any parameter whose limits result in loss of the *structural identifiability*⁷ of any other parameter in the model must belong to the same group as that parameter. In addition, a prime poset's parameters won't participate in any coordinated limits (e.g., two parameters going to infinity simultaneously while their ratio remains constant, or the like) with the parameters of any other prime poset.

⁷ See footnote 3.

5.2.2 Partial navigation of the Hasse diagram

If the model boundary poset is factorable, then so is the Hasse diagram, which may be completely navigated (top to bottom) by parts without any conflicts. It is also possible to partially navigate the Hasse diagram by parts even when it is not factorable or when the factorable pieces are too large to be manageable. As previously stated, the necessary requirement is that each partial path constructed be compatible with every other such path. A counterexample is shown in Fig. 5.5, which shows a Hasse diagram where the set of choices at the second reduction step (going from the top vertex down) for the remaining parameters depends on which path is taken on the first reduction step for the first parameter. If the network in Fig. 5.5 were expanded from a 4-cycle⁸ to a 6-cycle (see Fig. 5.6), then the first two parameters could be reduced independently, even though as a whole the model boundary poset is not factorable. This has implications for dividing up a large network and reducing parts of it separately (see Sec. 5.3.3).

5.3 Piecemeal Reduction Strategy

In general, there are several properties we have identified that are desirable in a piecemeal reduction strategy. These include:

1. Accounting for the effects of the rest of the system on each subsystem during its reduction. Typically a subsystem will exhibit different behavior when disconnected from the rest of the system (in open loop) than when connected (in closed loop). Fidelity to the closed-loop, rather than open-loop, behavior should be maintained.
2. Preserving connectivity with the rest of the system. Reduced subsystems should be compatible at points of connection so they can be reassembled after reduction.

⁸ A k -cycle is a (sub)network with k nodes and k edges connected in a loop.

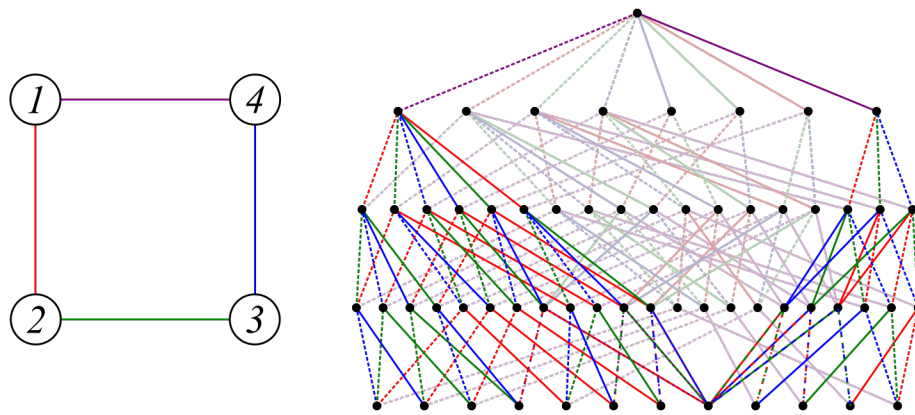


Figure 5.5 Network and Hasse diagram for a four parameter model. Some edges in the Hasse diagram have been grayed out to help illuminate certain structures that exist within the diagram. Going from the top vertex of the Hasse diagram to the next level down (the first reduction step), there are two choices for each parameter, but at the second reduction step some of these choices are lost. For example, if the left purple edge is followed, all six choices for red, green, and blue are preserved, and we obtain the network and Hasse diagram shown in Fig. 5.2; but if the right purple edge is taken, there are only three paths available and we obtain the network and Hasse diagram shown in Fig. 5.4 instead. If this Hasse diagram were a subset of a larger diagram (see Fig. 5.6), it could be partially navigated as long as the navigation stopped at what is the first level down from the top of this diagram.

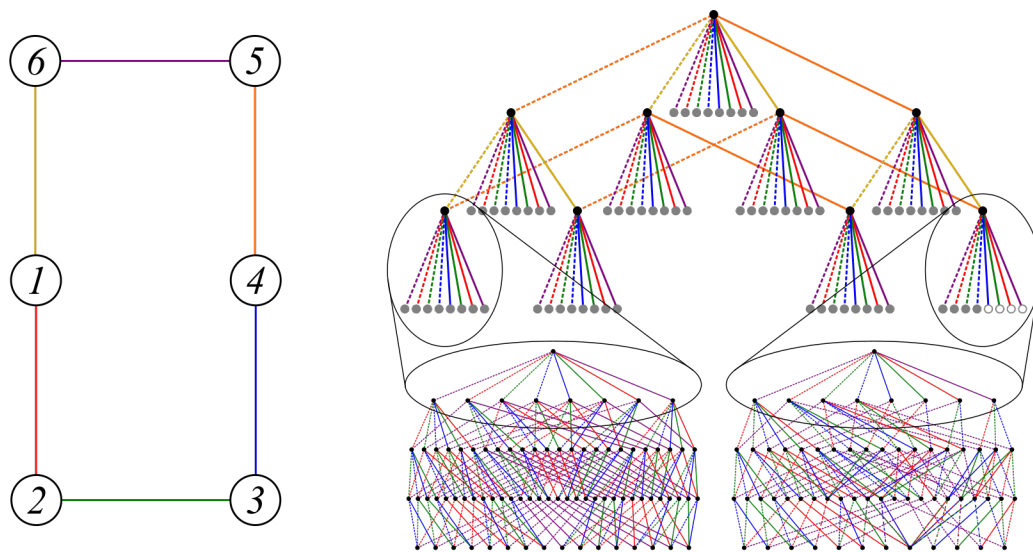


Figure 5.6 Network and partial Hasse diagram for a six parameter model. Edges (and their succeeding nodes) descending from gray and white nodes have been omitted for clarity. At gray nodes, two options are available for each remaining color; at white nodes only one option is available. As shown (bottom right), the Hasse diagram in Fig. 5.5 is a subset of this one. Although the repeating structure does not carry through to the bottom of the diagram, as it would if the corresponding model polytope were factorable, it does carry through the first few levels, making it possible to navigate the upper part of the diagram piecemeal.

3. Ensuring that reconnecting the reduced subsystems does not introduce instabilities or other unwanted artifacts. That is, the reduction process should respect underlying conservation laws of the system.
4. Maintaining physical interpretability. One of the key reasons for performing model reduction is to develop insight into the relationships between different components of the system by eliminating unnecessary detail. This insight is lost if the reduced model is not physically interpretable.

In this section, we discuss each of these properties in the context of reduction via MBAM, concluding with an outline of our proposed reduction strategy. Most of our attention is on the second property; we address the others only briefly.

5.3.1 Fidelity of closed-loop behavior

There are a couple of ways the first property can be attained. One is by retaining the entire model in simulation but only calculating parameter sensitivities of observations within the subsystem to be reduced and only with respect to parameters in that subsystem. This does not constitute a “true” piecemeal reduction, since the entire model must still be evaluated when calculating predictions and parameter sensitivities, but it at least reduces the size of the parameter Jacobian that must be calculated. Alternatively, there will be external state variables that appear in the subsystem equations in places where the subsystem being reduced connects to the rest of the system. These state variables can be replaced with inputs that simulate or replicate their behavior (perhaps obtained by recording their behavior during simulation of the entire system before reduction). There may be other ways to maintain fidelity of the closed-loop behavior of the subsystem; however, careful attention to this property is not the focus of this work.

5.3.2 Respecting conservation laws and maintaining physical interpretability

The third and fourth properties are both naturally achieved by MBAM. No truncations or projections are performed which could potentially violate underlying conservation laws, and no state transformations occur that could obscure the physical interpretation of system states. Instead, reduction proceeds by implementing a series of physically-meaningful parameter limits which simply restrict model predictions to an appropriate submanifold of the model manifold. Unnecessary state variables are naturally eliminated by, for example, steady-state approximations and singular limits in the parameters. As a result, the final reduced model has a direct interpretation in terms of the parameters and states of the original detailed one. In addition, as long as only proper reductions are performed (as discussed in Sec. 5.3.3), no other instabilities or artifacts will be introduced by performing the reduction piecewise.

5.3.3 Preserving connectivity

The second property is directly connected with the discussion in Sec. 5.2 and has particular implications for how subsystems are chosen. For example, if two subsystems are chosen to overlap each other, the two reductions may be incompatible in the overlapping region (i.e., may have conflicting parameter limits), preventing straightforward reassembly. Even if the system is partitioned so that no overlaps occur, there can be subtleties that arise along the boundaries that distinguish subsystems which need to be addressed.

To illustrate, consider the network shown on the left in Fig. 5.7.⁹ Associated with each edge is a

⁹ In the context of power systems, this could represent a small power network, with nodes representing buses, edges representing transmission lines, and w_{ik} corresponding to the magnitude of the line impedance. It was established in Ch. 4 that letting $w_{ik} \rightarrow \infty$ disconnects Buses i and k whereas letting $w_{ik} \rightarrow 0$ shorts the two buses (effectively merging them into a single bus). In this example, it is assumed that each bus has both a power source (generator) and a power

tunable parameter w_{ik} representing the effective length of the edge. Letting $w_{ik} \rightarrow \infty$ corresponds to removing the edge between nodes i and k , while $w_{ik} \rightarrow 0$ corresponds to contracting that edge and merging nodes i and k . Assume that observations are made at nodes and not along edges.

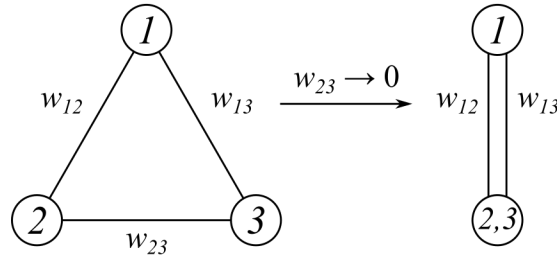


Figure 5.7 3-node network, connected in a cycle. Edge weights w_{ik} represent the effective length of the edge, such that $w_{ik} \rightarrow \infty$ represents edge removal and $w_{ik} \rightarrow 0$ represents edge contraction. Letting $w_{13} \rightarrow 0$ before either $w_{12} \rightarrow \infty$ or $w_{23} \rightarrow \infty$ results in a structurally unidentifiable parameter because any change in either w_{12} or w_{23} can then be compensated by the other.

Now consider what happens if the limit $w_{13} \rightarrow 0$ is naively applied (right-hand network in Fig. 5.7). Any change in w_{12} or w_{23} can now be compensated by the other, so there is a structurally unidentifiable parameter combination. This means the number of (identifiable) parameters has been reduced by two, not one. Such a reduction is forbidden; it does not correspond to any transition in the Hasse diagram which represents the model boundary poset.¹⁰ Either of the limits $w_{12} \rightarrow \infty$ and $w_{23} \rightarrow \infty$ comes first. When a 3-cycle (or cycle of any size) occurs within a subsystem, conventional MBAM will naturally find the appropriate limits at each reduction step because all of the cycle's parameters will be included in the parameter Jacobian, but not when a cycle occurs at the boundary between subsystems (Fig. 5.8). To avoid the situation in Fig. 5.7, the following restrictions would need to be placed on reduction of each subsystem:

1. No edge connecting two nodes on the subsystem boundary should be contracted.

sink (load), so if any bus becomes completely disconnected from the others, its voltage will remain finite, ensuring a bounded model manifold for any set of voltage observations made on the system; see Sec. 5.4.

¹⁰ This is the origin of the difference between Figs. 5.2 and 5.4 in Sec. 5.2.

2. An edge connecting a boundary node with a interior node should not be contracted if the interior node is also connected to another boundary node.

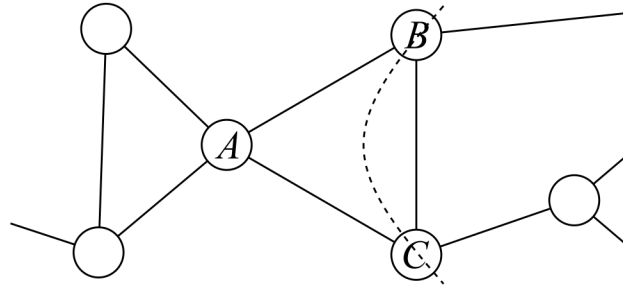


Figure 5.8 Part of a hypothetical network showing a proposed subsystem boundary (dashed line) that intersects a 3-cycle. Separately reducing the left and right subsystems could lead to complications for reassembly. On the left, naively contracting either edge AB or AC (or both) would make BC structurally unidentifiable. On the right, naively contracting BC would make AB or AC structurally unidentifiable.

These restrictions are specific to this type of network structure (i.e., undirected edges with positive weights) and measurement structure (i.e., observations on nodes but not on edges). Other types (e.g. directed edges, as in biochemical networks or other compartment models; edges with either positive or negative weights, as in spin interactions in a lattice; etc.) may require different restrictions. In general, the restrictions that are implemented in each case should ensure that the only reductions performed on subsystems are those which could have been performed on the complete system, i.e., which correspond to edges in the Hasse diagram for the full model boundary poset. We term such reductions *proper*. This is an important additional property of reduction via MBAM: *by performing only proper reductions, one can guarantee that the model obtained by reducing the pieces is something that could have been obtained by reducing the whole* (if it were computationally tractable to do so).

In many cases such restrictions are conservative and there will be additional reduction left undone along subsystem boundaries that (in principle) could have been achieved by reducing the whole all at once. In terms of the Hasse diagram, only part of the diagram will be navigated (see

Fig. 5.6 in Sec. 5.2). This may seem to indicate that the boundaries should be chosen carefully in order to minimize these restrictions. This is only partly true. If the piecemeal reduction process is performed iteratively, then subsystem boundaries can be redefined or removed at each iteration so that reductions that were restricted during one iteration become unrestricted at the next. This can be done in a way that allows all reductions that could have been performed on the whole to be explored by MBAM during reduction of the pieces. We address this further in the proposed reduction strategy which follows.

5.3.4 Proposed strategy

We propose the following strategy for piecemeal reduction via MBAM (or LinearizedMBAM). This strategy ensures subsystems are compatible after reduction and allows all reductions that could have been performed on the whole system to be explored during reduction of the subsystems.

1. Define the subsystems to be reduced.
2. Implement restrictions along boundaries where each subsystem connects with the others to ensure subsystems will be compatible after reduction.
3. Use some method to replicate or simulate the input-output behavior of each subsystem at connection points with other subsystems so the behavior of each subsystem reflects what it would have been if still connected.
4. Reduce the subsystems using MBAM or LinearizedMBAM.
5. Reassemble the reduced subsystems and define new subsystems that cover previous boundary regions.
6. Repeat 2-5 until no more restrictions remain or all restrictions have been lifted in at least one iteration of the process.

5.4 Resistor Network Example

In this section we demonstrate our strategy on a network of resistors. Consider the network shown in Fig. 5.9. Each branch has an associated resistance R_{ik} which is taken as a model parameter. Currents I_i^{in} are injected at each of the junctions (nodes) i , flow through the network, and are returned via resistors to ground R_i^{out} at every junction. These resistances are fixed and assumed to be known. Each of the outgoing currents I_i^{out} is measured, as are voltages V_i at junctions.

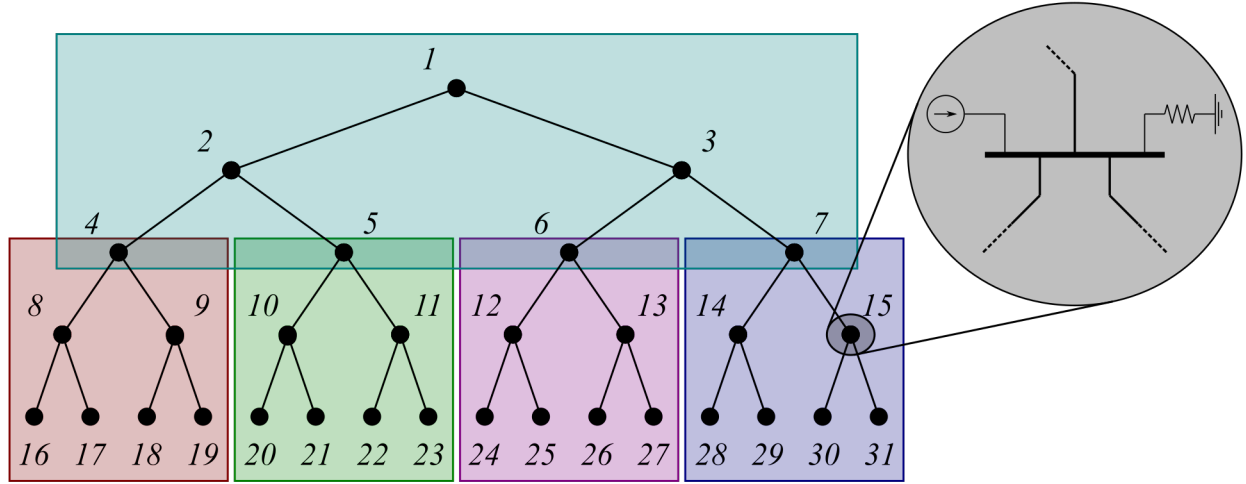


Figure 5.9 Diagram for a test case resistor network. Each branch/edge has an associated resistance. Voltage is measured at junctions (nodes; numbered). Current is injected at every junction, flows through the network, and is returned through a resistor to ground at each junction (see detail at right). Each shaded box is a subnetwork to be reduced separately in the piecemeal reduction process.

We model the network as follows. Current conservation at each of the nodes gives

$$I_i^{\text{in}} - I_i^{\text{out}} = \sum_{k=1}^{N_{\text{node}}} I_{ik}, \quad (5.4)$$

where I_{ik} is the current flowing out of node i to node k . The voltage difference across each branch is given by

$$V_i - V_k = I_{ik} R_{ik}. \quad (5.5)$$

Finally, the outgoing current I_i^{out} from each node is determined by

$$V_i = I_i^{\text{out}} R_i^{\text{out}}. \quad (5.6)$$

Assuming the input currents I_i^{in} and resistances to ground R_i^{out} are known, these three sets of equations can be solved for the node voltages V_i , branch currents I_{ik} (not observed), and outgoing currents I_i^{out} as a function of the branch resistances R_{ik} .

This network structure was chosen to lead to a model manifold with specific characteristics. The acyclic topology of the network allows each of the resistances R_{ik} to be varied independently of the others, leading to a manifold whose boundary poset is a hypercube. In addition, each of the branch resistances can be varied to either extreme (0 or ∞) without any of the observable variables becoming unphysical. These two extremes are, in fact, the DC analogues of the two parameter limits observed and presented in Ch. 4, where two nodes either become disconnected ($R_{ik} \rightarrow \infty$) or merged ($R_{ik} \rightarrow 0$).¹¹ This can be seen when the limits are applied to Eq. (5.5):

$$I_{ik} = \lim_{R_{ik} \rightarrow \infty} \frac{V_i - V_k}{R_{ik}} = 0, \quad (5.7)$$

$$V_i - V_k = \lim_{R_{ik} \rightarrow 0} I_{ik} R_{ik} = 0. \quad (5.8)$$

In particular, even if a node or subset of nodes becomes completely disconnected from the rest of the network, there is always a path to ground for the current to follow and the voltage on the node will remain finite, resulting in a bounded model manifold. In more realistic scenarios (e.g., in power systems where not every bus has both a generator and a load), care must be taken to appropriately identify the viable parameter limits and hence the structure of the model boundary poset.

We divided the network into five subnetworks (as indicated in Fig. 5.9) and applied LinearizedMBAM to each. In this case there are no cycles in the network, so there are no restrictions necessary at subsystem boundaries. Since this is a synthetic network and there is no measurement

¹¹ Because resistance is inversely related to admittance, $Y = 1/Z = 1/(R + jX)$, the limits for the resistance given here are the opposite of those for the admittance discussed in Ch. 4.

uncertainty to provide a scale for the bias introduced during the reduction, we arbitrarily ended the reduction of each subnetwork when 1/3 of the parameters remained (i.e., 2 out of 6). Table 5.1 lists the parameter reductions found, while Fig. 5.10 shows the resulting network. For comparison, we also continued the reduction until no parameters remained; the additional parameter limits are shown in Table 5.1 in parentheses. The total bias (sum of squared error between the predictions of the unreduced and reduced models) for each of these cases is shown in Table 5.2, along with the bias for two other cases: letting all parameters go to zero or to infinity. It is clear that, using MBAM, we have found a set of parameter limits that introduces comparatively little bias.

Table 5.1 Parameter reductions, listed in the order in which they occurred in the reduction of each subnetwork. Reduction of each proceeded until 2/3 of the parameters had been reduced; reductions which would have occurred after this point are listed in parentheses. Numerals indicate the order in which these reductions occurred when the entire network was reduced simultaneously instead of piecemeal.

Subnetwork 1	Subnetwork 2	Subnetwork 3	Subnetwork 4	Subnetwork 5
2) $R_{3,7} \rightarrow \infty$	11) $R_{8,16} \rightarrow 0$	4) $R_{10,21} \rightarrow 0$	1) $R_{13,26} \rightarrow 0$	3) $R_{7,15} \rightarrow \infty$
10) $R_{2,5} \rightarrow 0$	15) $R_{4,8} \rightarrow \infty$	6) $R_{11,23} \rightarrow 0$	5) $R_{12,24} \rightarrow \infty$	20) $R_{14,29} \rightarrow \infty$
9) $R_{1,3} \rightarrow 0$	19) $R_{9,18} \rightarrow \infty$	8) $R_{11,22} \rightarrow \infty$	7) $R_{12,25} \rightarrow 0$	21) $R_{15,31} \rightarrow 0$
12) $R_{1,2} \rightarrow \infty$	17) $R_{4,9} \rightarrow 0$	14) $R_{5,10} \rightarrow \infty$	22) $R_{6,12} \rightarrow 0$	26) $R_{15,30} \rightarrow 0$
13) $(R_{3,6} \rightarrow \infty)$	25) $(R_{8,17} \rightarrow 0)$	18) $(R_{5,11} \rightarrow 0)$	23) $(R_{6,13} \rightarrow \infty)$	28) $(R_{7,14} \rightarrow 0)$
16) $(R_{2,4} \rightarrow 0)$	30) $(R_{9,19} \rightarrow 0)$	27) $(R_{10,20} \rightarrow 0)$	24) $(R_{13,27} \rightarrow 0)$	29) $(R_{14,28} \rightarrow \infty)$

Table 5.2 Total bias (sum of squared error) introduced for different reduced models of the network in Fig. 5.9. The first is LinearizedMBAM piecemeal reduction where 2/3 of the parameters are reduced. The second is the same but reducing all of the parameters. The third and fourth are other possible (but less optimal) reductions.

Piecemeal, 2/3 reduced	Piecemeal, fully reduced	All parameters to 0	All parameters to ∞
70	190	740	3800

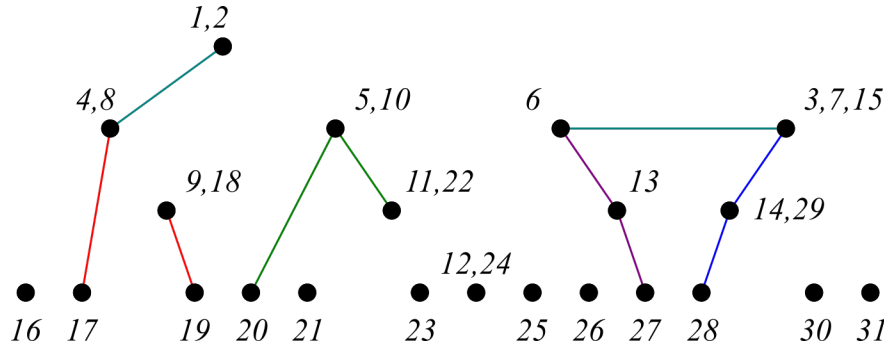


Figure 5.10 Resistor network from Fig. 5.9, after reducing 2/3 of the parameters. The remaining branches are colored according to which subnetwork they were in prior to reduction. Merged junctions are indicated. At a certain level of measurement precision, this network would be indistinguishable from the one in Fig. 5.9. It is also computationally simpler and serves as an abstraction of the unreduced network’s primary (measurable) features, e.g., which nodes are electrically “close together” (merged), “far apart” (disconnected), or in between.

To show that our results are robust to the particular choice of subnetworks, we reduced the entire network simultaneously (i.e., without splitting it into subnetworks) and found the same parameter reductions, with only minor differences in the order (see Table 5.1).

5.5 Discussion and conclusions

In this chapter, we have proposed a strategy for splitting a system into subsystems (i.e., splitting model parameters and predictions into groups), reducing the subsystems using MBAM (or LinearizedMBAM), and then reassembling them. In some cases, to avoid conflicts, not all of the parameters can be included in a single reduction, in which case the reduction can be performed iteratively and the parameters can be regrouped at each iteration. To the extent that various parameters in the model can be reduced independently, the choice of how to group parameters for piecemeal reduction is quite flexible and arbitrary. As long as the parameter limits of the various groups are interchangeable (as discussed in Sec. 5.2), the final reduced model will not depend significantly on how the parameters are grouped.

We demonstrated our piecemeal reduction strategy using LinearizedMBAM on a self-similar network with identical, first-order subsystems (i.e., nodes with a single state). In contrast to other methods, our strategy may also be applied to arbitrary networks with nonidentical, nontrivial (i.e., possibly dynamic and nonlinear) subsystems at each node. No assumptions about the structure of the network or the nodes are required, and both the network and the nodes may be reduced simultaneously. In addition, no particular network structure needs to be imposed on the final reduced model. Rather, detail is maintained (for example, in a study area) by specifying which parts of the system are going to be observed and hence which predictions' fidelity should be maintained in the reduced model. The reduction then proceeds by finding the physically-meaningful limiting approximations which affect these model predictions the least. The resulting network topology and subsystem internal structure reflect their salient features while eliminating unnecessary detail; structure preservation occurs automatically.

Although beyond the scope of this research, there are many additional questions that could be addressed or considered in future work. One of the assumptions implicit in our strategy is that the effect of parameters in one subsystem on model predictions that are “distant” (i.e., in a different subsystem) can be neglected when reducing a given subsystem, either because it is negligible or because it is correlated with the effect of parameters on “nearby” model predictions. The primary way this effect would manifest is that, after the subsystems have been reduced and reassembled, the total bias in the whole will be significantly greater than the combined biases of the reduced subsystems. Some reductions that are considered good approximations for a given subsystem may not be good approximations for the whole because they introduce too much bias in other subsystems, but this would not be discovered until the reduced subsystems are recombined. One possible way to test for this would be to examine the Jacobian of parameter sensitivities. Dividing a system up into subsystems also divides parameters and predictions into groups, which in turn divides the Jacobian into different blocks. Some blocks will consist of sensitivities of predictions to parameters from the

same subsystem, while others will consist of sensitivities of predictions to parameters from other subsystems. Analysis of these blocks (such as a singular value decomposition) may reveal whether parameters can significantly affect distant predictions without affecting those nearby.

Other questions regard the relationship between the model mapping (i.e., from parameters to predictions) and the model boundary poset. For example, under what conditions does the model mapping lead to a polytope boundary structure, and when will it fail to do so? Can prime factors in the model boundary poset decomposition be determined directly from an examination of the mathematical structure of the model (especially the parameters)? or how can they be discovered most effectively?

This work has established the foundation for piecemeal reduction via MBAM. It is an important stepping stone in ongoing efforts to scale up the MBAM model reduction technique to very large systems.

Chapter 6

Conclusion

The Manifold Boundary Approximation Method (MBAM) is an important new tool for simplifying (esp. nonlinear) models of large, complex systems, not only reducing the computational burden of using them but also illuminating key relationships among components by abstracting away unnecessary detail. In this dissertation I have made significant contributions to the development of MBAM, extending its uses to oscillatory, networked, and very large systems.

In Ch. 3, I analyzed the properties of models of oscillatory systems and showed that their model manifolds characteristically have a high effective dimensionality not suitable for model reduction via MBAM. I used the concept of analytic signal (AS) as well as kernel density estimation (KDE) to transform the predictions being made by these models in a way that leads to smooth model manifolds with low effective dimensionality, i.e., the hyperribbon structure characteristic of other sloppy models. This also alleviates the local minima problem that arises when inferring parameters in oscillatory systems.

These techniques have potential application in large, complex oscillatory systems, such as networks of neurons studied in neuroscience, where model reduction may help illuminate aggregate or other “macroscopic” behavior. In addition, using the AS methods I discussed in Ch. 3 also provides a means of decoupling the effects of parameters on limit cycle behavior from their effects

on the oscillation frequency of the system, which could be useful for analysis and control in systems where it is desirable to adjust aspects of the limit cycle without changing the system frequency, or vice versa.

There is a notable difference between the methods proposed for periodic oscillatory systems (using AS) and chaotic oscillatory systems (using KDE) in that information about the temporal sequence of model predictions is kept in the first case and lost in the second. Although the exponential divergence of nearby trajectories that is characteristic of chaotic systems makes long-term prediction problematic, there still exists a short-term predictability that distinguishes chaotic systems from stochastic ones [82, 83] which is washed out by KDE. It may prove useful to develop an alternative to or an adaptation of KDE that preserves this short-term temporal sequencing information.

One idea is to consider not just the distribution but the flow of model predictions in phase space. For example, in a multiple-shooting approach, rather than calculating an entire trajectory from a single initial point in phase space, a set of initial points distributed in some way throughout phase space could each be used to predict just one succeeding point. This short-term trajectory would be determined by the flow in phase space that is defined by the differential equations of the model and so would not suffer from the exponential divergence of long-term trajectories. Kernel density estimates of these short-term predictions could then be compared using a metric much like the one suggested in Ch. 3.

Another approach is to calculate the flow of the distribution in phase space directly, rather than estimating the distribution from the trajectories of points in the space. It is shown in [84] that, given an ODE system $dx_i/dt = F_i(x)$, the evolution of the distribution $u(t, x)$ of model predictions in phase space satisfies the following continuity equation:

$$\frac{\partial u}{\partial t} + \sum_i \frac{\partial (u F_i)}{\partial x_i} = 0. \quad (6.1)$$

Solving this equation for $u(t, x)$, one could again use the metric suggested in Ch. 3 for comparing

distributions. Then again, solving a PDE such as this one may be too computationally intensive to be useful in many cases, especially if the phase space has more than a handful of dimensions (which most do, especially for large complex systems). Still, the concept is intriguing and could lead to something useful.

In Ch. 4, I performed MBAM reduction on a small network from the field of power systems, showing how the reduction is dependent on the choice of system observations. This process also resulted in the discovery of some of the different types of parameter limits which are appropriate for the mathematical structures that arise in power systems models.

In the reduction I performed, some component (i.e., machine) parameters were kept constant due to computational limitations. Without expert intuition, the appropriate reduction limits for these parameters remain unclear due to the complexity of the governing equations of the model. A more detailed analysis and exploration of the model manifold for a single high-order machine would not only validate limiting approximations already found by experts but uncover other as-yet-unsuspected but valid parameter limits. Such limits are easily found by MBAM without the need of expert intuition.

I discussed piecemeal reduction of large systems in Ch. 5 in terms of navigating the Hasse diagram for the boundary structure of the model manifold. I also outlined and demonstrated a piecemeal reduction strategy on a small network. Some interesting additional questions that could prompt future work have already been posed there.

One obstacle that remains in order to be able to scale up MBAM for very large systems is that prior knowledge about the appropriate parameter limits for the mathematical structures in the model is needed, so for example, MBAM reduction of one or more small or moderate systems is usually necessary before LinearizedMBAM can be applied to a large one. This highlights the question whether it is possible to determine the boundary structure of the model manifold by direct examination or analysis of the governing equations of the model. Then again, if the small

or moderate systems to which MBAM is first applied are carefully chosen, it may be possible to explore the entire boundary structure and find the appropriate types of parameter limits without much effort and without the need of much analysis of the equations.

Model reduction can be a tool not only for reducing computational burden but for abstraction, which is at the core of scientific modeling. My hope is that the tools, methods, and concepts discussed in this dissertation prove useful in solving many important problems in science and help provide insight into many different kinds of systems.

Bibliography

- [1] L. P. Kadanoff, “Scaling laws for Ising models near T_c ,” *Physics Physique Fizika* **2**, 263 (1966).
- [2] S. Izvekov and G. A. Voth, “Multiscale coarse graining of liquid-state systems,” *The Journal of chemical physics* **123**, 134105 (2005).
- [3] W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, “The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models,” *The Journal of chemical physics* **128**, 244114 (2008).
- [4] W. G. Noid, P. Liu, Y. Wang, J.-W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen, and G. A. Voth, “The multiscale coarse-graining method. II. Numerical implementation for coarse-grained molecular models,” *The Journal of chemical physics* **128**, 244115 (2008).
- [5] W. G. Noid, “Perspective: Coarse-grained models for biomolecular systems,” *The Journal of chemical physics* **139**, 09B201_1 (2013).
- [6] L. Pernebo and L. Silverman, “Model reduction via balanced state space representations,” *IEEE Transactions on Automatic Control* **27**, 382–387 (1982).
- [7] P. V. Kokotovic, R. E. O’Malley Jr, and P. Sannuti, “Singular perturbations and order reduction in control theory—An overview,” *Automatica* **12**, 123–132 (1976).

-
- [8] 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*, pp. 1401–1407 (1985).
- [9] R.-C. Li *et al.*, "Structure-preserving model reduction using a Krylov subspace projection formulation," *Communications in Mathematical Sciences* **3**, 179–199 (2005).
- [10] C. Beattie and S. Gugercin, "Interpolatory projection methods for structure-preserving model reduction," *Systems & Control Letters* **58**, 225–232 (2009).
- [11] D. Chaniotis and M. Pai, "Model reduction in power systems using Krylov subspace methods," *IEEE Transactions on Power Systems* **20**, 888–894 (2005).
- [12] R. V. Polyuga and A. Van der Schaft, "Structure preserving model reduction of port-Hamiltonian systems by moment matching at infinity," *Automatica* **46**, 665–672 (2010).
- [13] B. Besselink, U. Tabak, A. Lutowska, N. Van de Wouw, H. Nijmeijer, D. J. Rixen, M. Hochstenbach, and W. Schilders, "A comparison of model reduction techniques from structural dynamics, numerical mathematics and systems and control," *Journal of Sound and Vibration* **332**, 4403–4422 (2013).
- [14] D. Bonvin and D. Mellichamp, "A unified derivation and critical review of modal approaches to model reduction," *International Journal of Control* **35**, 829–848 (1982).
- [15] A. C. Antoulas and D. C. Sorensen, "Approximation of large-scale dynamical systems: An Overview," Technical report (2001) .
- [16] S. Chaturantabut and D. C. Sorensen, "Nonlinear model reduction via discrete empirical interpolation," *SIAM Journal on Scientific Computing* **32**, 2737–2764 (2010).

-
- [17] M. Rewieński and J. White, “Model order reduction for nonlinear dynamical systems based on trajectory piecewise-linear approximations,” *Linear algebra and its applications* **415**, 426–454 (2006).
- [18] J. M. A. Scherpen, “Balancing for nonlinear systems,” *Systems & Control Letters* **21**, 143–153 (1993).
- [19] S. Lall, P. Krysl, and J. E. Marsden, “Structure-preserving model reduction for mechanical systems,” *Physica D: Nonlinear Phenomena* **184**, 304–318 (2003).
- [20] K. Carlberg, R. Tuminaro, and P. Boggs, “Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics,” *SIAM Journal on Scientific Computing* **37**, B153–B184 (2015).
- [21] L. Peng and K. Mohseni, “Symplectic model reduction of Hamiltonian systems,” *SIAM Journal on Scientific Computing* **38**, A1–A27 (2016).
- [22] B. M. Afkham and J. S. Hesthaven, “Structure preserving model reduction of parametric Hamiltonian systems,” *SIAM Journal on Scientific Computing* **39**, A2616–A2644 (2017).
- [23] S. Chaturantabut, C. Beattie, and S. Gugercin, “Structure-preserving model reduction for nonlinear port-Hamiltonian systems,” *SIAM Journal on Scientific Computing* **38**, B837–B865 (2016).
- [24] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna, “Universally sloppy parameter sensitivities in systems biology models,” *PLoS computational biology* **3** (2007).
- [25] M. K. Transtrum, B. B. Machta, and J. P. Sethna, “Why are nonlinear fits to data so challenging?,” *Physical review letters* **104**, 060201 (2010).

- [26] M. K. Transtrum, B. B. Machta, and J. P. Sethna, “Geometry of nonlinear least squares with applications to sloppy models and optimization,” *Physical Review E* **83**, 036701 (2011).
- [27] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, “Parameter space compression underlies emergent theories and predictive models,” *Science* **342**, 604–607 (2013).
- [28] M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna, “Perspective: Sloppiness and emergent theories in physics, biology, and beyond,” *The Journal of chemical physics* **143**, 07B201_1 (2015).
- [29] M. K. Transtrum and P. Qiu, “Model reduction by manifold boundaries,” *Physical review letters* **113**, 098701 (2014).
- [30] M. K. Transtrum and P. Qiu, “Bridging mechanistic and phenomenological models of complex biological systems,” *PLoS computational biology* **12** (2016).
- [31] G. Bohnert and G. Venkataraman, “Identifiability, reducibility, and adaptability in allosteric macromolecules,” *Journal of General Physiology* **149**, 547–560 (2017).
- [32] C. Gaza and M. K. Transtrum, Senior thesis, Brigham Young University, 2017.
- [33] M. K. Transtrum, A. T. Sarić, and A. M. Stanković, “Measurement-directed reduction of dynamic models in power systems,” *IEEE Transactions on Power Systems* **32**, 2243–2253 (2016).
- [34] C. C. Youn, A. T. Sarić, M. K. Transtrum, and A. M. Stanković, “Information geometry for model reduction of dynamic loads in power systems,” In *2017 IEEE Manchester PowerTech*, pp. 1–6 (2017).
- [35] T. Nikšić and D. Vretenar, “Sloppy nuclear energy density functionals: Effective model reduction,” *Physical Review C* **94**, 024333 (2016).

- [36] T. A. Bahr and M. K. Transtrum, Senior thesis, Brigham Young University, 2018.
- [37] P. E. Paré, D. Grimsman, A. T. Wilson, M. K. Transtrum, and S. Warnick, “Model boundary approximation method as a unifying framework for balanced truncation and singular perturbation approximation,” *IEEE Transactions on Automatic Control* **64**, 4796–4802 (2019).
- [38] K. N. Quinn, H. Wilber, A. Townsend, and J. P. Sethna, “Chebyshev approximation and the global geometry of sloppy models,” *arXiv preprint arXiv:1809.08280* (2018).
- [39] S. W. Blume, *Electric power system basics for the nonelectrical professional* (John Wiley & Sons, 2007).
- [40] A. Von Meier, *Electric power systems: a conceptual introduction* (John Wiley & Sons, 2006).
- [41] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control* (McGraw-hill New York, 1994), Vol. 7.
- [42] F. Milano, *Power system modelling and scripting* (Springer Science & Business Media, 2010).
- [43] P. W. Sauer, M. A. Pai, and J. H. Chow, *Power system dynamics and stability: with synchrophasor measurement and power system toolbox* (John Wiley & Sons, 2018).
- [44] B. L. Francis and M. K. Transtrum, “Unwinding the model manifold: choosing similarity measures to remove local minima in sloppy dynamical systems,” *Physical Review E* **100**, 012206 (2019).
- [45] 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)*, pp. 1–5 (2018).

- [46] B. L. Francis, J. R. Nuttall, M. K. Transtrum, A. T. Sarić, and A. M. Stanković, “Network Reduction in Transient Stability Models using Partial Response Matching,” In *2019 North American Power Symposium (NAPS)*, pp. 1–6 (2019).
- [47] R. W. Freund, “SPRIM: structure-preserving reduced-order interconnect macromodeling,” In *IEEE/ACM International Conference on Computer Aided Design, 2004. ICCAD-2004.*, pp. 80–87 (2004).
- [48] H. Yu, L. He, and S. Tar, “Block structure preserving model order reduction,” In *BMAS 2005. Proceedings of the 2005 IEEE International Behavioral Modeling and Simulation Workshop, 2005.*, pp. 1–6 (2005).
- [49] L. Li and F. Paganini, “Structured coprime factor model reduction based on LMIs,” *Automatica* **41**, 145–151 (2005).
- [50] A. Vandendorpe and P. Van Dooren, “Model reduction of interconnected systems,” in *Model order reduction: theory, research aspects and applications* (Springer, 2008), pp. 305–321.
- [51] H. Sandberg and R. M. Murray, “Model reduction of interconnected linear systems,” *Optimal Control Applications and Methods* **30**, 225–245 (2009).
- [52] H. Sandberg, “An extension to balanced truncation with application to structured model reduction,” *IEEE Transactions on Automatic Control* **55**, 1038–1043 (2010).
- [53] P. Trnka, C. Sturk, H. Sandberg, V. Havlena, and J. Řehoř, “Structured model order reduction of parallel models in feedback,” *IEEE Transactions on Control Systems Technology* **21**, 739–752 (2012).
- [54] C. Sturk, L. Vanfretti, Y. Chompoobutrgool, and H. Sandberg, “Coherency-independent structured model reduction of power systems,” *IEEE Transactions on Power Systems* **29**, 2418–2426 (2014).

- [55] N. Monshizadeh, H. L. Trentelman, and M. K. Camlibel, “Stability and synchronization preserving model reduction of multi-agent systems,” *Systems & Control Letters* **62**, 1–10 (2013).
- [56] F. Al-Taie and H. Werner, “Structure-preserving model reduction for spatially interconnected systems with experimental validation on an actuated beam,” *International Journal of Control* **89**, 1248–1268 (2016).
- [57] T. Ishizaki, K. Kashima, J.-i. Imura, and K. Aihara, “Model reduction and clusterization of large-scale bidirectional networks,” *IEEE Transactions on Automatic Control* **59**, 48–63 (2013).
- [58] T. Ishizaki, K. Kashima, A. Girard, J.-i. Imura, L. Chen, and K. Aihara, “Clustered model reduction of positive directed networks,” *Automatica* **59**, 238–247 (2015).
- [59] T. Ishizaki and J.-i. Imura, “Clustered model reduction of interconnected second-order systems,” *Nonlinear Theory and Its Applications, IEICE* **6**, 26–37 (2015).
- [60] N. Monshizadeh, H. L. Trentelman, and M. K. Camlibel, “Projection-based model reduction of multi-agent systems using graph partitions,” *IEEE Transactions on Control of Network Systems* **1**, 145–154 (2014).
- [61] A. Van Der Schaft, “On model reduction of physical network systems,” In *Proceedings of 21st International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, pp. 1419–1425 (2014).
- [62] N. Monshizadeh and A. van der Schaft, “Structure-preserving model reduction of physical network systems by clustering,” In *53rd IEEE Conference on Decision and Control*, pp. 4434–4440 (2014).

- [63] B. Besselink, H. Sandberg, and K. H. Johansson, “Model reduction of networked passive systems through clustering,” In *2014 European Control Conference (ECC)*, pp. 1069–1074 (2014).
- [64] B. Besselink, H. Sandberg, and K. H. Johansson, “Clustering-based model reduction of networked passive systems,” *IEEE Transactions on Automatic Control* **61**, 2958–2973 (2015).
- [65] X. Cheng, Y. Kawano, and J. M. Scherpen, “Reduction of second-order network systems with structure preservation,” *IEEE Transactions on Automatic Control* **62**, 5026–5038 (2017).
- [66] X. Cheng, Y. Kawano, and J. M. Scherpen, “Model Reduction of Multiagent Systems Using Dissimilarity-Based Clustering,” *IEEE Transactions on Automatic Control* **64**, 1663–1670 (2018).
- [67] J. H. Chow, J. Cullum, and R. A. Willoughby, “A sparsity-based technique for identifying slow-coherent areas in large power systems,” *IEEE transactions on power apparatus and systems* pp. 463–473 (1984).
- [68] J. Chow and P. Kokotovic, “Time scale modeling of sparse dynamic networks,” *IEEE Transactions on Automatic Control* **30**, 714–722 (1985).
- [69] R. A. Date and J. H. Chow, “Aggregation properties of linearized two-time-scale power networks,” *IEEE Transactions on Circuits and Systems* **38**, 720–730 (1991).
- [70] E. Bıyık and M. Arcak, “Area aggregation and time-scale modeling for sparse nonlinear networks,” *Systems & Control Letters* **57**, 142–149 (2008).
- [71] D. Romeres, F. Dörfler, and F. Bullo, “Novel results on slow coherency in consensus and power networks,” In *2013 European Control Conference (ECC)*, pp. 742–747 (2013).

- [72] M. L. Ourari, L.-A. Dessaint, and V.-Q. Do, “Dynamic equivalent modeling of large power systems using structure preservation technique,” *IEEE Transactions on Power Systems* **21**, 1284–1295 (2006).
- [73] U. D. Annakkage *et al.*, “Dynamic system equivalents: A survey of available techniques,” *IEEE Transactions on Power Delivery* **27**, 411–420 (2011).
- [74] G. N. Ramaswamy, C. Evrard, G. C. Verghese, O. Fillatre, and B. C. Lesieutre, “Extensions, simplifications, and tests of synchronic modal equivalencing (SME),” *IEEE Transactions on Power Systems* **12**, 896–905 (1997).
- [75] X. Lei, D. Povh, and O. Ruhle, “Industrial approaches for dynamic equivalents of large power systems,” In *2002 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No. 02CH37309)*, **2**, 1036–1042 (2002).
- [76] M. K. Transtrum, B. L. Francis, A. T. Sarić, and A. M. Stanković, “Parameter reduction of large-scale models,” (In preparation).
- [77] R. Bellman and K. J. Åström, “On structural identifiability,” *Mathematical biosciences* **7**, 329–339 (1970).
- [78] 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)*, pp. 1–6 (2018).
- [79] I. Gleason and I. Hubbard, “Products of abstract polytopes,” *Journal of Combinatorial Theory, Series A* **157**, 287–320 (2018).
- [80] J. Hashimoto, “On direct product decomposition of partially ordered sets,” *Annals of Mathematics* pp. 315–318 (1951).

-
- [81] I. Rival, “The diagram,” in *Graphs and order* (Springer, 1985), pp. 103–133.
- [82] L. Lacasa and R. Toral, “Description of stochastic and chaotic series using visibility graphs,” *Physical Review E* **82**, 036120 (2010).
- [83] N. Subramaniam, J. F. Donges, and J. Hyttinen, “Signatures of chaotic and stochastic dynamics uncovered with $\tilde{\text{I}}$ -recurrence networks,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **471**, 20150349 (2015).
- [84] A. Lasota and M. C. Mackey, *Chaos, fractals, and noise: stochastic aspects of dynamics* (Springer Science & Business Media, 2013), Vol. 97.