Information Geometry for Model Verification in Energy Systems

Mark K. Transtrum
Brigham Young University
Dept. of Physics and Astronomy
Provo, UT, USA
mktranstrum@byu.edu

Andrija T. Sarić
Faculty of Technical Sciences
Dept. of Power, Electr. and Comm. Eng.
Novi Sad, Serbia
asaric@uns.ac.rs

Alex M. Stanković
Tufts University
Dept. of Electrical and Computer Eng.
Medford, MA, USA
astankov@ece.tufts.edu

***Abstract*—The paper describes a new class of system identification procedures that are tailored to electric power systems with renewable sources. Our procedure builds on computational advances in differential geometry, and offers a new, global, intrinsic characterization of challenges frequently encountered in system identification of electric power systems. The approach also benefits from increased availability of high-quality measurements. While the proposed procedure is illustrated on renewable source IEEE 14-bus based example in a multi-machine benchmark power system, it is equally applicable to identification of other system components (for example, dynamic loads).**

***Index Terms*--Renewable Energy Systems, System Identification, Power System Stability, Information Geometry, Computational Differential Geometry.**

I. Introduction

Dynamic models of power systems (for example, electro-mechanical models used in transient analysis) have grown in size to thousands of generators and tens of thousands of nodes. However, this growth in quantitative terms has largely not been accompanied with improvements in fidelity. Specifically, models have been largely unable to replicate major events like the 2003 blackout in the Eastern interconnection [1] and several such events in the 1990's in the Western interconnection [2]. This is even more disquieting, given the relatively widespread presence of sensors that have made detailed recordings during transients.

System identification is particularly needed in medium voltage (MV) networks, where much of renewable energy integration is occurring. These "active distribution" networks are evolving fast because of: 1) changes within (MV lines are easier to build, with potentially more renewables in spatial proximity), 2) changes above (in transmission - e.g., topology control), and 3) changes below (more storage, power electronic loads). The increased presence of renewable sources interfaced through power electronic converters has recently led to some qualitatively new stability problems. While the technical causes of these "harmonic resonances" are still under investigation, it is clear that these new phenomena transcend the traditional time decomposition between electromechanical transient stability and network transients.

Attempts to use on-line data to improve dynamical models of key components, e.g. synchronous generators, have a long history in power systems. First efforts to employ general dynamical systems concepts like trajectory sensitivity go back more than a quarter century [3, 4]. That particular approach has been extended to hybrid systems in [5]. Another influential approach that is based on local information extracted from the measurement Jacobian is described in [6]. To deal with ill conditioning of the parameter estimation problem, the reference proposes that a subset of parameters of the generator model be fixed to prior values while estimating the remaining parameters from the available data (denoted as the subset selection method). In the sequel we only list references of immediate relevance to our development here. For example, [7] considers parameter estimation for a single generator and re-casts synchronous generator parameter identification in a differential-algebraic equation (DAE) framework. Reference [8] considers the same overall setup involving Phasor Measurement Unit (PMU) derived measurements. The authors estimate parameters one-by-one (later up to 3) and use the Extended Kalman filter to validate the parameter values.

Measurement-derived parameter estimation is equally relevant for load modeling, as it often introduces the largest uncertainty in the overall dynamic model. Optimization-based approaches to parameter identification often face the so-called plateau phenomenon, when the criterion function becomes insensitive over large portions of the parameter space and the problem has multiple local solutions. We outline a novel differential geometric explanation using model manifolds later; the problem itself is well documented in [9]. Finally, we mention that our dynamic system identification procedure is relevant for steady-state problems in power systems. One such example comes from conventional state estimation, which is essentially a (steady-state) parameter identification whose practical implementations have long been plagued by convergence and uniqueness problems [10]. Our approach is relevant for microgrids, for virtual entities (virtual utilities, energy hubs) that are often considered essential in the long-term evolution of smart grids, and for future electricity markets that

will operate on shorter time-scales and thus depend on model fidelity of system dynamics.

The outline of the paper is as follows: *Section II* describes the power system models used for identification; *Section III* discuss necessary conditions for well-conditioning of parameters estimation, while in *Section IV* the information geometry, semi-global and global sensitivity based approach to model identification is described; *Section V* shows obtained results in a multi-machine benchmark example and *Section VI* contains our recommendations and conclusions.

II. Power System Model for Identification

Models of power systems are typically written in DAEs form [11]:

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

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

where $\boldsymbol{x}$ is the vector of (differential) state variables, z are the algebraic variables, $\boldsymbol{p}$ are parameters, $\boldsymbol{u}$ are inputs (typically assumed to be known in estimation studies) and t is the (scalar) time variable. System measurement vector is assumed to be of the form [8]:

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

The parameters ($\boldsymbol{p}$) are to be estimated from measurements ($\boldsymbol{y}$), and there typically exists some prior information about parameters, often in the form of plausible ranges for each. The least-square optimization formulation of the identification problem is by far the most prevalent in the literature.

It turns out that the key quantities in the case of least square identification are parametric sensitivities whose dynamics is described by the following equations:

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

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

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

These equations are linear in terms of sensitivities, but the matrices involved do vary along a system trajectory. Please note that in the case of multivariable states and parameters, the overall problem dimensionality grows quickly as, for example, $\partial \boldsymbol{x}/\partial \boldsymbol{p} = \left[\partial \boldsymbol{x}/\partial p_1 \;\; \partial \boldsymbol{x}/\partial p_2 \;\; \cdots \;\; \partial \boldsymbol{x}/\partial p_p\right]^{\mathrm{T}}$ and $\partial \boldsymbol{x}/\partial p_1$ is an n-dimensional vector (p is total number of uncertain parameters).

Details about DAEs modeling of a Direct-Drive Synchronous Generator (DDSG), a Double-Fed Induction Generator (DFIG) and a transmission network are provided in *Appendix*.

III. Parameter Estimation

Equation (6) determines the $(m \times p)$-dimensional (m is total number of available measurements) Jacobian matrix $\boldsymbol{J}_p(t) = \partial \boldsymbol{h}(t)/\partial \boldsymbol{p}$, or the matrix of first partial derivatives of system measurement vector (3) with respect to the parameter vector ($\boldsymbol{p}$) at each time point. In the neighborhood of true parameter values, the full Hessian matrix of second derivatives is well approximated with $\boldsymbol{H}_p(t) = \boldsymbol{J}_p^{\mathrm{T}}(t)\boldsymbol{J}_p(t)$, which is symmetric and positive semidefinite (so all its eigenvalues are real and non-negative). Suppose that $\boldsymbol{H}_p(t)$ is singular with only one eigenvalue at 0. A direct implication of the singularity is that variation of parameters along the corresponding eigenvector cannot be detected from measurements. Typically $\boldsymbol{H}_p(t)$ is not exactly singular and nearness to singularity is measured by the *condition number* $\kappa(\boldsymbol{H}_p)$, which for a symmetric and positive matrix is the ratio of the largest ($\lambda_{\max}$) to smallest eigenvalue ($\lambda_{\min}$) [6, 11].

There are several consequences of the Hessian near-singularity. First is that the solution of DAEs (1), (2) varies much more slowly with parameter vector ($\boldsymbol{p}$) in some directions than in others. Second is that the vector $\boldsymbol{p}$ is poorly estimated in directions where the curvature is small (relative to directions with high curvature). Numerous references [3, 6, 7, 8] state that all parameters cannot be estimated together in a typical case. In our work, the ill-conditioned parameters are detected locally using participation factors of $\boldsymbol{H}_p(t)$ [11].

IV. Information Geometry, Semi-Global and Global Sensitivities

Recent advances focusing on data space rather than parameter space have proven beneficial for understanding the global properties of models and for advancing numerical techniques for exploring them [12, 13]. This approach, usually known as *information geometry*, since it combines *information theory* with *differential geometry*, is a natural mathematical language for exploring parameterized models [14, 15]. The essence of the approach is the interpretation of a model as a manifold embedded in the space of data, known as the model manifold. The key aspects are:

1. There is no information loss in the model manifold, i.e. the manifold retains information about all model predictions [12, 13]. In contrast, the cost surface in parameter space condenses the high-dimensional quantities such as the prediction and data vectors into a single number, i.e., the cost.
2. Information geometry separates the model, i.e. the manifold embedded in data space, from the data to which it is being fit, i.e. a point in the data space [12, 13]. This is a useful abstraction which allows one to study the intrinsic properties of the model irrespective of a particular experimental observation. In contrast, the cost surface in parameter space is a function of and often very sensitive to the observation.
3. The set of points that constitute the model manifold are the same regardless of how the model is parameterized [12-15]. The parameters are not ignored completely, but act as coordinates on the manifold, i.e., labels for specific predictions.
4. The Riemannian metric on the model manifold (which describes differences in predictions of two models that are infinitesimally apart) is the Fisher Information Matrix, i.e., the Hessian matrix described above [12-15]. Information geometry therefore serves as a natural bridge between the local analysis and the global analysis.
5. The language of differential geometry naturally accommodates the potentially large dimensionality of both the parameter and data spaces.

Our paper complements the commonly used local parameter sensitivity analyses with semi-global and global techniques. **Semi-global** methods of sensitivity analysis address the shortcomings of local approaches by sampling parameter space in a finite neighborhood around the best fit, using tools such as scanning and Bayesian methods. Information Geometry has proven beneficial for understanding the **global** properties of models and for advancing numerical techniques for exploring them. This recognition leads to the interpretation of a model as a manifold embedded in the space of data.

V. Application

We have developed a Matlab-derived simulation environment built around the IEEE 14-bus test system shown in Fig. 1 (see Appendix D in [16] for detailed input data), modified to include direct-drive synchronous generators (DDSG – used by industry to model solar plants and a new generation of wind), doubly-fed induction generators (DFIG – capturing prevalent type of wind plants today) and synchronous generators (SG – describing conventional units and interconnections). We consider stability-related models in the DAEs based form (1), (2).

Our environment is based on PSAT, which is a suite of freely available Matlab routines well documented in [16], to which we have added our code for evaluation of measurement sensitivities (in Matlab) and for computational differential geometry (in Python). Our Matlab code is fully general in the sense that it allows for a variety of measurements (rotor angle and speed, nodal active and reactive power injections, nodal voltage magnitude and angle, branch active and reactive flows, and branch current magnitude).

A. Local Sensitivity

From (A1), (A2) for DDSG (assumed wind-driven) state variables of interest are: two mechanical variables [speed (ω) and pitch angle (θ_p)] and electrical states (i_{sq} and i_{cd}), while three algebraic variables are: i_{sd}, i_{cq} and $P_w^*(\omega_m)$. Parameters (electrical) of interest are: T_p, T_{sq}, T_{cd}, x_d and x_q. Mechanical parameters H_m correspond to the much slower dynamics and are assumed known [estimated from motion equation in (A1)] for simplicity of presentation.

In the case of DFIG from (A3,4) each unit (assumed double-fed induction generator driven) has 4 states (ω_m, θ_p, i_{rd} and i_{rq}) and 2 algebraic variables ($P_w^*(\omega_m)$ and V_{ref}); assumed unknown electrical parameters are: T_p, T_e, x_s and x_μ.

Available measurements for the distributed sources are the rotor angle (δ)[1], speed (ω), and the real and reactive powers (P_g and Q_g, respectively), as well as the terminal voltage magnitude and angle (V and θ, respectively).

In the case of parameter estimation for a single machine from local measurements, we can remove the algebraic equations (denoted with $\boldsymbol{g}$ in *Appendix*) altogether; the algebraic variables $\boldsymbol{z}$ (1)-(3) still remain (V, θ for network buses and other

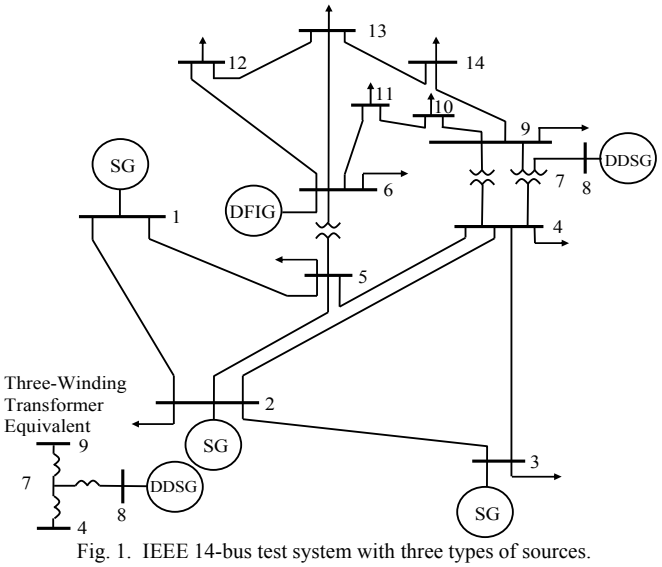


Fig. 1. IEEE 14-bus test system with three types of sources.

for generator units as described above). In order to demonstrate salient features of our method on a model that is relevant, but straightforward enough for tracking key relationships, we focus on the four differential equations on DDSG and DFIG examples (denoted with $\boldsymbol{f}$ in (A1), (A3) in *Appendix*).

However, in actual power systems (also in the analyzed test IEEE 14-bus power system), there exist additional dynamical components [exciters, automatic voltage regulators (AVR), steam/wind turbines etc.], as well as multiple generators and loads in buses. A single DDSG unit described by $\boldsymbol{f}$ in (A1) would see these other algebraic components through variation in the current components (i_{sd} and i_{cq}), mechanical (wind) power [P_w or $P_w^*(\omega_m)$] and in the complex voltage in the point of connection (represented by V and θ). For simplicity we assume that interface variables for DDSG in bus 8 (that is i_{sd}, i_{cq}, $P_w^*(\omega_m)$, V_8 and θ_8) are known functions of time, but independent of the parameters considered. Similarly, the interface variables for DFIG in bus 6 (that is V_{ref}, $P_w^*(\omega_m)$, V_6 and θ_6) are known functions of time. This, of course, is an approximation for a multi-generator system, but it allows direct compareson with numerous references that focus on a single generator.

We start by displaying transients in sensitivities following inadvertent opening of the line 2–4 (see Fig. 1), which is reclosed after 200 ms. For example, the transient variations of the voltage magnitude and angle variations at bus 8 (where DDSG is connected) are approximately 4 % and 20 degree, respectively.

As an illustrative example, Fig. 2 considers sensitivities of DFIG in bus 6 to two parameters (x_s and x_μ), where the measurement is the active power. Note that the "Basic" case refers to the power transient with fixed parameters. From (4)–(6) we can conclude that for "Basic" case the responses are equal with DAEs responses (1), (2).

In Table I the eigenvalues, condition numbers and participation factors for different uncertain parameter sets are presented.

[1] Note that the rotor angle cannot be measured directly. However, the indirect methods where the rotor angle is calculated from on-line measurements of active/ reactive powers and voltage in connection point can be applied (for example, see ref. [17, eq. (8)].

A summary for the analyzed disturbance:

- DDSG's and DFIG's inertia (H_m) is well-conditioned and can be estimated from the motion equation – see **Case 1** in Table I.
- Time constants (T_p , T_{sq} and T_{cd} for DDSG, or T_p and T_e for DFIG) are ill-conditioned (two zero eigenvalues dominantly influence their participation factors) and cannot be estimated simultaneously from electrical equations – see **Cases 2–4** in Table I.
- Machine reactances (x_d and x_q for DDSG, or x_s and x_μ for DFIG) are well-conditioned (condition numbers are 1028.55 and 2571.18, respectively) and can be estimated simultaneously from electrical equations and available measurements – see **Case 5** in Table I.
- Eigenvalue plots are slowly changing with different structures of uncertain parameter sets (see λ_i columns in Table I).
- Transient sensitivity of eigenvalues, shown in Fig. 3, indicates that the above conclusions are quite general for analyzed time period (note that eigenvalues in Table I are calculated for the last analyzed transient point ($t = 10$ s), while the condition number is calculated as an average value over the transient).

B. Semi-Global and Global Analysis

We have generated artificial data for a set of "true" parameter values and performed a Markov Chain Monte Carlo (MCMC) sampling [18] of the posterior distribution for the model fit to the data.

For analyzed types of distributed sources (DDSG and DFIG) projecting the cloud of points onto each pair of parameter axes typically results in a cloud that is not elliptical. Deviations from an elliptical cloud indicate that a simple local analysis will not capture many important structural features of the model.

In the case of DDSG, our results suggest that parameters x_q (first row in Fig. 4a) and T_{sq} (second column in Fig. 4a) are likely to be difficult to estimate. Similarly the parameter x_s (first column and second row in Fig 4b) and T_e (second column in Fig 4b) will be difficult to estimate.

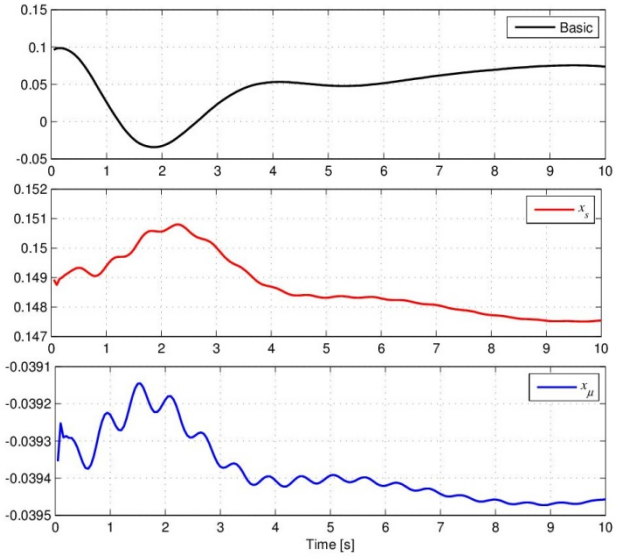


Fig. 2. DFIG Example: real power transient (top panel) and real power sensitivities to variations in x_s and x_μ (bottom two panels).

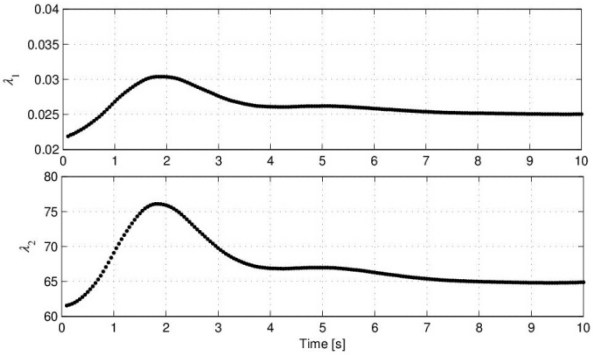


Fig. 3. Example of transient sensitivity of eigenvalues for uncertain parameter set shown by **Case 5** (DFIG) in Table I.

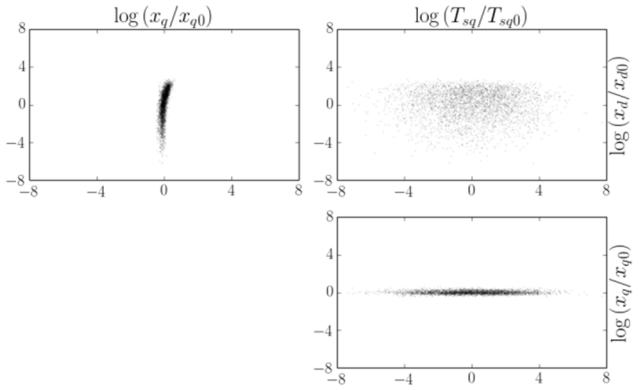
TABLE I

EIGENVALUES (λ_i), CONDITION NUMBERS [$\kappa(\boldsymbol{H_p})$] AND PARTICIPATION FACTORS (p_{ki}) FOR CHARACTERISTIC SETS OF UNCERTAIN PARAMETERS.

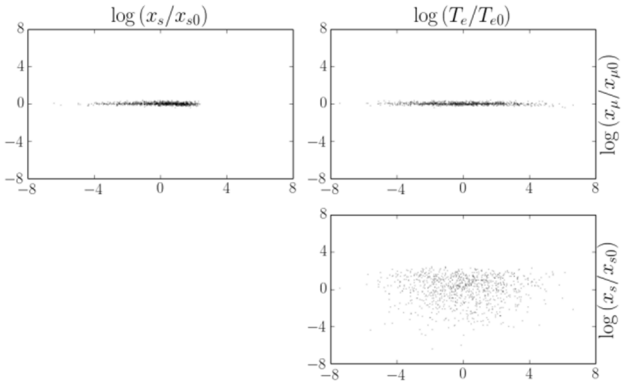
	DDSG				DFIG			
	Uncertain parameters	λ_i	$\kappa(\boldsymbol{H_p})$	p_{ki}	Uncertain parameters	λ_i	$\kappa(\boldsymbol{H_p})$	p_{ki}
Case 1	H_m	5.0443	1	1.0000	H_m	33.2627	1	1.0000
Case 2	T_{sq} , T_{cd} , x_d , x_q *	0.0003 0.5541 12.803 3061.9	58292952	0.452; 0.536; 0.002; 0.010 0.277; 0.196; 0.239; 0.289 0.243; 0.267; 0.093; 0.396 0.029; 0.001; 0.665; 0.305	T_p , T_e , x_s , x_μ	0.0000 0.0010 0.0304 134.16	∞	0.500; 0.500; 0.000; 0.000 0.212; 0.212; 0.031; 0.545 0.037; 0.037; 0.731; 0.195 0.251; 0.251; 0.238; 0.260
Case 3	T_{cd} , x_d , x_q	0.5011 9.1376 4758.3	33919	0.650; 0.137; 0.213 0.349; 0.215; 0.436 0.001; 0.648; 0.351	T_e , x_s , x_μ	0.0007 0.0289 100.20	142064	0.597; 0.022; 0.381 0.068; 0.660; 0.272 0.335; 0.318; 0.347
Case 4	T_{sq} , x_d , x_q	0.3931 9.2188 3334.9	230320	0.653; 0.182; 0.165 0.326; 0.159; 0.515 0.021; 0.659; 0.320	T_p , x_s , x_μ	0.0007 0.0280 97.243	141473	0.593; 0.019; 0.388 0.073; 0.663; 0.264 0.335; 0.318; 0.347
Case 5	x_d , x_q	6.1160 3182.6	1028	0.339; 0.661 0.661; 0.339	x_s , x_μ	0.0250 64.615	2571**	0.522; 0.478 0.478; 0.522

* Analysis for pitch time constant (T_p) is provided for DFIG with the same derived conclusions.

** For example, the condition number varies during the transient (0–10 s) from minimal value of 1813 to maximal 510814.



a. DDSG in bus 8



b. DFIG in bus 6

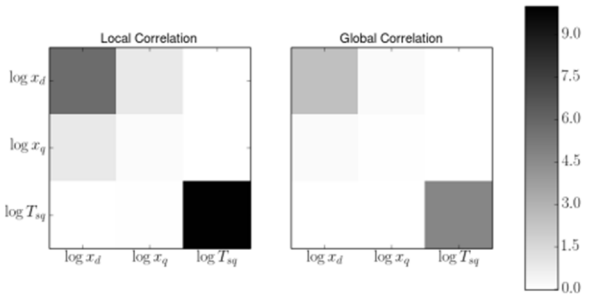
Fig. 4. Projections of the Bayesian posterior sampling for each pair of electrical log-parameters. "True" parameter values are denoted by the 0 subscript.

Observe how few of the clouds are elliptical (for example, the cloud in the first column, first row of Fig 4a (x_d vs x_q) is both curved and asymmetric with a long tail of points corresponding to small values of x_d), indicating that the local sensitivity analysis will be inadequate to quantify details of the parameter correlations.

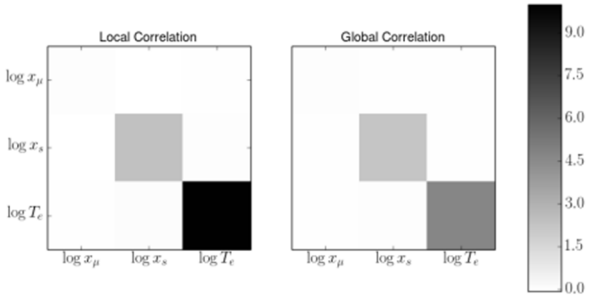
Notice that the local analysis captures many qualitative features (e.g., the relative uncertainty of each parameter), but is not quantitatively accurate when compared with the more rigorous Bayesian sampling.

The inadequacy of the local analysis is further demonstrated by considering covariance matrices for both the local and semi-global analysis. Correlation matrices estimated from a local sensitivity analysis and a semi-global Bayesian sampling in log-parameters are shown in left and right panels in Fig. 5, respectively. Notice that the local correlation matrix gives a qualitatively correct picture of the correlations, but is not quantitatively accurate. For example, the local correlation matrix (i.e., the inverse Fisher Information Matrix) correctly predicts the relative difficulty of inferring each parameter. However, the local analysis is quantitatively inaccurate when compared with the covariance matrix of the Bayesian sampling. Furthermore, from the Bayesian sampling, higher-order moments can be estimated that describe the non-quadratic features of the sampling cloud (Fig. 4).

Next, we consider a frequentist (non-Bayesian) alternative to MCMC sampling that has better scaling properties in the case of large systems for identifiability analysis of individual parameters. We fix one parameter and fit the model by varying the remaining parameters; we identify the confidence intervals for each parameter by considering what range still yields a good fit (measured by log likelihood). For example, in the case of DDSG and focusing on x_d (Fig. 6, top), notice that it has an upper cut-off, but can be made very small. This is corroborated by MCMC sampling (Fig. 4a, top left plot) in which the cloud of points is asymmetric with a long tail of points at small values of x_d . Unlike MCMC, the likelihood profile reveals that there is no minimum cut-off for x_d – it can be made arbitrarily small and still fit the data. Because of this, x_d is known as a practically non-identifiable parameter. The lower-bound on x_d from the MCMC sampling is a consequence of the weak Bayesian prior used to regularize the model. Without this prior, the MCMC sampling would not have converged. We also show the likelihood profile for T_e (Fig. 6b) from the DFIG model. The likelihood profile suggests that T_e can be made both very large and small as indicated by the MCMC sampling.



a. DDSG in bus 8



b. DFIG in bus 6

Fig. 5. Correlation matrices estimated from a local sensitivity analysis, i.e., the inverse Fisher Information matrix (left), and a semi-global Bayesian sampling (right) in log-parameters. Darker colors correspond to larger values in the covariance matrices.

In Fig. 7, for DDSG we project the cloud of MCMC points onto a two-dimensional plane in parameter space defined by the parameters x_d and x_q (cf. Fig. 4a, top left panel). The contours correspond to level sets of the negative log-likelihood (i.e., a two-parameter joint likelihood profile), so that the background greyscale represents the cost function for the two dimensional cross section. We then consider geodesic curves (i.e., straight lines on the model manifold in data space) constrained to move through this two-dimensional cross section. Notice how the

geodesics are highly nonlinear (reflecting the incompleteness of the local analysis) but align with the long-narrow canyon of the sloppy direction. In this way the geodesics systematically explore the non-local structure of the parameter space.

The geodesics extend the parameter identifiability analysis of both the MCMC and likelihood profiles. The geodesic curves are parameterized by the proper distance on the model manifold, i.e., by changes in model behavior. When geodesic curves extend parameter values to zero or infinity in a finite distance on the model manifold, the corresponding parameter is susceptible to identifiability problems. To illustrate, we have seen that the parameter x_d (DDSG in bus 8) is practically unidentifiable from below (i.e., can be made arbitrarily small). In the geometric analysis, this is manifest by the geodesic paths in which x_d approach zeros as in Fig. 7.

The differential geometric approach is a global identifiability analysis (as opposed to the semi-global MCMC and likelihood methods) that is intrinsic to the model, i.e., independent of the actually observed data. To illustrate the utility of this approach consider the DDSG example illustrated in Fig. 7. For the data set that gave rise to Fig. 7, the parameter x_q (DDSG in bus 8) is identifiable and well-constrained by the data. However, we see that there is a flat canyon in the cost surface characterized by x_q becoming zero (near $\log(x_d/x_{d0}) = 2$). Imagine an alternative data set with fewer data or data of less quality or in which the true value of x_q was a bit smaller. In any of these scenarios, the parameter x_q would have been unidentifiable from below. We therefore say that x_q is *susceptible* to being unidentifiable.

By calculating geodesics in many directions, we identify all the ways in which parameters or parameter combinations are *susceptible* to being non-identifiable. This analysis is *global* and intrinsic to the model, i.e., it completely characterizes all potential identifiability problems that could arise from any data set and for any true parameter values. It is useful for selecting the number and type of measurements necessary to avoid these potential identifiability problems. It can also guide model selection. More fundamentally, it characterizes the information structure of the model's parameter space. We illustrate for our two test cases.

For the case of DDSG in bus 8, our analysis shows that each of the three parameters is susceptible to non-identifiability from below. In addition, the parameter x_q is susceptible to being non-identifiable from above (i.e., can become infinite).

For the case of DFIG in bus 6, our analysis reveals that x_s and T_e are both non-identifiable from below. In addition, the parameters can all be non-identifiable due to non-trivial correlations. Parameters x_s and T_e can be non-identifiable due to a correlation in which x_s becomes zero and T_e becomes infinite. Similarly, there is also a correlation in which x_μ and x_s become zero and T_e becomes infinite.

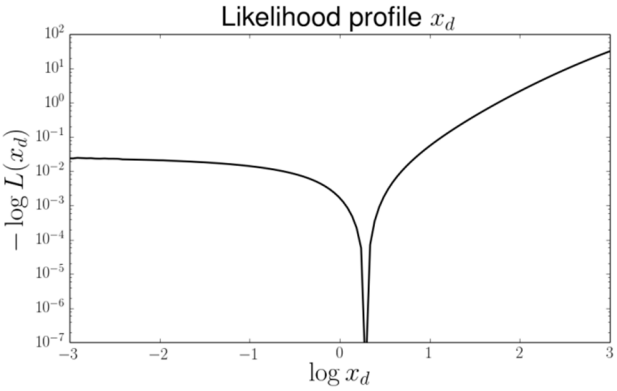


Fig. 6a. Likelihood profiles for x_d (DDSG in bus 8)

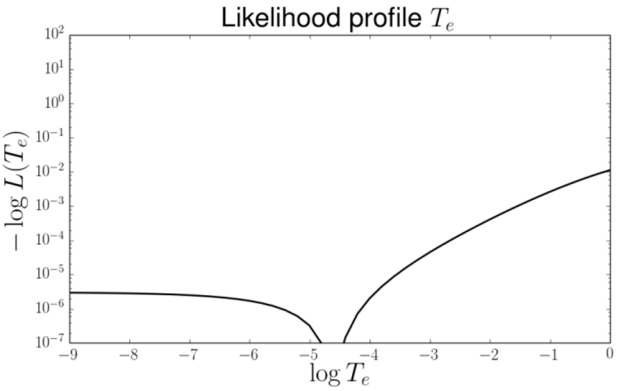


Fig. 6b. Likelihood profiles for T_e (DFIG in bus 6)

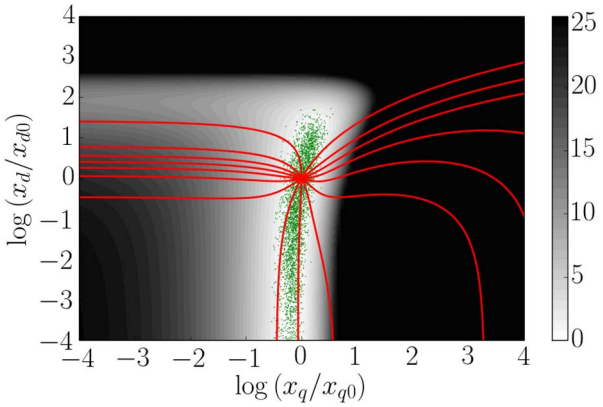


Fig. 7. Global exploration of the manifold (DDSG in bus 8)

VI. Conclusions

This paper outlines a new class of system identification procedures that are specifically tailored to electric power systems, in particular to distributed sources. Our procedure builds on computational advances in differential geometry and offers a new, global characterization of challenges frequently encountered in system identification of electric power systems. In particular, we use information geometry to develop global sensitivity analysis of DAE models in power systems.

References

[1] G. Andersson et al., "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," *IEEE Trans. Power Systems*, vol. 20, no. 4, pp. 1922-1928, Nov. 2005.

[2] D.N. Kosterev, C.W. Taylor, and W.A. Mittelstadt, "Model validation the August 10, 1996 WSCC system outage," *IEEE Trans. Power Systems*, vol. 14, no. 3, pp. 967-979, Aug. 1999.
[3] J.J. Sanchez-Gasca et al., "Trajectory sensitivity based identification of synchronous generator and excitation system parameters," *IEEE Trans. Power Systems*, vol. 3, no. 4, pp. 1814-1822, Nov. 1988.
[4] S.M. Benchluch and J.H. Chow, "A trajectory sensitivity method for the identification of nonlinear excitation system models," *IEEE Trans. Energy Conversion*, vol. 8, no. 2, pp. 159-164, June 1993.
[5] I.A. Hiskens, "Nonlinear dynamic model evaluation from disturbance measurement," *IEEE Trans. Power Systems*, vol. 16, no. 4, pp. 702-710, Nov. 2001.
[6] M. Burth, G.C. Verghese, and M. Velez-Reyes, "Subset selection for improved parameter estimation in on-line identification of a synchronous generator," *IEEE Trans. Power Systems*, vol. 14, no. 1, pp.218-225, Feb. 1999.
[7] E. Cari and L.F.C. Alberto, "Parameter estimation of synchronous generators from different types of disturbances," *IEEE PES General Meeting*, 2011.
[8] Z. Huang et al., "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Trans. Power Systems*, vol. 28, no. 2, pp. 1939-1949, May 2013.
[9] B.K. Choi and H.-D. Chiang, "Multiple solutions and plateau phenomenon in measurement-based load model development: issues and suggestions," *IEEE Trans. Power Systems*, vol. 24, no. 2, pp. 824-831, May 2009.
[10] A. Abur and A.G. Exposito, "Detecting multiple solutions in state estimation in the presence of current magnitude measurements," *IEEE Trans. Power Systems*, vol. 12, no. 1, pp. 370-375, Feb. 1997.
[11] P. Kundur, *Power System Stability and Control*, New York, USA: McGraw-Hill, 1994.
[12] M.K. Transtrum, B.B. Machta, and J.P. Sethna, "Why are nonlinear fits to data so challenging?," *Physical Review Letters*, 104(6):060201, 2010.
[13] 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*, vol. 83, 036701, 2011.
[14] S.-I. Amari, *Differential-Geometrical Methods in Statistics,* Springer, 1985.
[15] S.-I. Amari et al. *Differential Geometry in Statistical Inference*, Springer Lecture Notes-Monograph Series, 1987.
[16] F. Milano, *Power System Modelling and Scripting*, London, UK: Springer, 2010.
[17] E. Ghahremani and I. Kamwa, "Local and wide-area PMU-based decentralized dynamic state estimation in multi-machine power systems," *IEEE Trans. Power Systems*, to be published.
[18] S. Chib and E. Greenberg, "Understanding the Metropolis Hastings algorithm," *The American Statistician*, vol. 49, no. 4, pp. 327-335, Nov. 1995.

APPENDIX

For DDSG the differential and algebraic equations respectively can be written as:

$$\boldsymbol{f} \Rightarrow \begin{cases} \dot{\omega}_m = \dfrac{\tau_m - \tau_e}{2H_m} = \dfrac{\dfrac{P_w(\omega_m)}{\omega_m} - [\psi_p + (x_q - x_d) i_{sd}] i_{sq}}{2H_m} \\ \dot{\theta}_p = \dfrac{1}{T_p}\left(K_p \phi(\omega_m - \omega_{ref}) - \theta_p\right) \\ \dot{i}_{sq} = \dfrac{1}{T_{sq}}\left(\dfrac{P_w^*(\omega_m)}{\omega_m(\psi_p - x_d i_{sd})} - i_{sq}\right) \\ \dot{i}_{cd} = \dfrac{1}{T_{cd}}\left(K_{cd}(V_{ref} - V) - i_{cd}\right) \end{cases}; \quad \text{(A1)}$$

$$\boldsymbol{g} \Rightarrow \begin{cases} i_{sd} = \dfrac{\dfrac{P_g}{\omega_m i_{sq}} - \psi_p}{x_q - x_d} \\ i_{cq} = \dfrac{Q_g - V\cos\theta i_{cd}}{V\sin\theta} \\ P_w^*(\omega_m) = \begin{cases} 0 & \text{if} \quad \omega_m < 0.5 \\ 2\omega_m - 1 & \text{if} \quad 0.5 \le \omega_m \le 1 \\ 1 & \text{if} \quad \omega_m > 1 \end{cases} \end{cases}. \quad \text{(A2)}$$

For DFIG the differential and algebraic equations respectively can be written as:

$$\boldsymbol{f} \Rightarrow \begin{cases} \dot{\omega}_m = \dfrac{\tau_m - \tau_e}{2H_m} \\ \dot{\theta}_p = \dfrac{1}{T_p}\left(K_p \phi(\omega_m - \omega_{ref}) - \theta_p\right) \\ \dot{i}_{rd} = K_V(V - V_{ref}) - \dfrac{V}{x_\mu} - i_{rd} \\ \dot{i}_{rq} = \dfrac{1}{T_e}\left(-\dfrac{x_s + x_\mu}{x_\mu V} \cdot \dfrac{P_w^*(\omega_m)}{\omega_m} - i_{rq}\right) \end{cases}; \quad \text{(A3)}$$

$$\boldsymbol{g} \Rightarrow \begin{cases} P_w^*(\omega_m) = \begin{cases} 0 & \text{if} \quad \omega_m < 0.5 \\ 2\omega_m - 1 & \text{if} \quad 0.5 \le \omega_m \le 1 \\ 1 & \text{if} \quad \omega_m > 1 \end{cases} \\ 0 = V_{ref}^0 - V_{ref} + v_s^{\text{POD}} \end{cases}, \quad \text{(A4)}$$

where:

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

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

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

Matrix and complex form of active/reactive power injection balance equations in buses is:

$$\mathbf{0} = \underline{\boldsymbol{V}}\,\underline{\boldsymbol{Y}}^*\,\underline{\boldsymbol{V}}^* - \underline{\boldsymbol{V}}\,\underline{\boldsymbol{I}}(\boldsymbol{x}, \underline{\boldsymbol{V}}), \quad \text{(A5)}$$

where:

$\underline{\boldsymbol{V}} = diag\{\underline{V}_1 \quad \underline{V}_2 \quad \cdots \quad \underline{V}_N\}$ and $\underline{\boldsymbol{Y}}$ is bus admittance matrix;

$P_{inj} = \sum_{j=1}^{N}\left[V_i V_j \left(G_{ij}\cos(\theta_i - \theta_j) + B_{ij}\sin(\theta_i - \theta_j)\right)\right]$;

$Q_{inj} = \sum_{j=1}^{N}\left[V_i V_j \left(G_{ij}\sin(\theta_i - \theta_j) - B_{ij}\cos(\theta_i - \theta_j)\right)\right]$;

DDSG: $P_g = v_{cd} i_{cd} + v_{cq} i_{cq}$; $Q_g = v_{cq} i_{cd} - v_{cd} i_{cq}$;

DFIG: $P_g = v_{sd} i_{sd} + v_{sq} i_{sq} + v_{cd} i_{cd} + v_{cq} i_{cq}$;

$Q_g = v_{sq} i_{sd} - v_{sd} i_{sq} + v_{cq} i_{cd} - v_{cd} i_{cq}$.