# Nonlinear Dynamic Model Evaluation From Disturbance Measurements

Ian A. Hiskens, *Senior Member, IEEE*

***Abstract*—The nonlinear nonsmooth nature of power system dynamics complicates the process of validating system models from disturbance measurements. This paper uses a Gauss–Newton method to compute a set of model parameters that provide the best fit between measurements and model response. Trajectory sensitivities are used to identify parameters that can be reliably estimated from available measurements. An overview of trajectory sensitivity analysis is provided. An example based on the Nordel system is used throughout the paper to illustrate the various concepts. This example exhibits both soft and hard nonlinearities.**

***Index Terms*—Identifiability, nonlinear least-squares, parameter estimation, power system dynamics, trajectory sensitivity.**

## I. Introduction

SYSTEM-WIDE measurements of power system disturbances are frequently used in event reconstruction to gain a better understanding of system behavior [1]–[3]. In undertaking such studies, measurements are compared with the behavior predicted by a model. Differences are used to tune the model, i.e., adjust parameters, to obtain the best match between the model and the measurements.

An example of the model tuning procedure, and the importance of correct modeling, is provided in [3]. In that case, the power system lost stability following a large disturbance. Event reconstruction was undertaken to explore the nature of the instability. It was found that by using the "standard" set of parameters, the model did not replicate the unstable behavior. However, an exhaustive investigation showed that correct behavior could be predicted by altering a load parameter by a small amount.

This example illustrates the need for a systematic approach to,

1) identifying which parameters can be estimated reliably from the available measurements, and
2) obtaining a best estimate for those parameters.

The difficulty is that power system behavior is nonlinear. Models must therefore also be nonlinear, and in fact frequently contain hard nonlinearities, i.e., discontinuities. Parameter estimation techniques are well established for linear models [4]. However, parameter estimation for nonlinear systems is a relatively open field.

This paper initially explores the usefulness of trajectory sensitivities for identifying parameters that make good candidates for reliable estimation. The results are compared with the approach proposed in [5] for determining *well-conditioned* parameters. Exact agreement is obtained. The link with the concept of *identifiability* [4] is also presented. The paper proceeds to develop a Gauss–Newton iterative technique for determining parameter values that provide the best match between measured large disturbance system behavior and the model response.

The general parameter estimation concepts presented in this paper are not new [6]. In the power systems context, similar ideas have been used for estimating parameters of generators and AVRs/exciters [5], [7]–[9]. However, the number of parameters that could be estimated using those earlier ideas was limited, because trajectory sensitivities were generated numerically [8], [9]. A computationally efficient method of calculating trajectory sensitivities for nonlinear nonsmooth systems has recently been presented in [10]. This paper exploits that extension to consider many system-wide parameters, whilst accounting for switching action in an analytically correct manner.

A brief review of modeling is given in Section II. Trajectory sensitivity concepts are presented in Section III, along with their use in determining well-conditioned parameters. Those ideas are used in Section IV to develop the desired parameter estimation algorithm. Identifiability is considered. Conclusions are presented in Section V. Concepts are illustrated throughout using actual measurements from a disturbance on the Nordel system [11].

## II. Model and Example

### A. Model

Power systems frequently exhibit a mix of continuous time dynamics, discrete-time and discrete-event dynamics, switching action and jump phenomena. It is shown in [12], [13] that such systems, known generically as *hybrid systems*, can be modeled by a set of differential-algebraic equations, adapted to incorporate impulsive (state reset) action and switching of the algebraic equations. This *DA Impulsive Switched* (DAIS) model can be written in the form,

$$\dot{x} = f(x, y) \tag{1}$$

$$0 = g^{(0)}(x, y) \tag{2}$$

$$0 = \begin{cases} g^{(i-)}(x, y) & y_{d,i} < 0 \\ g^{(i+)}(x, y) & y_{d,i} > 0 \end{cases} \qquad i = 1, \ldots, d \tag{3}$$

$$x^+ = h_j(x^-, y^-) \qquad y_{e,j} = 0 \qquad j \in \{1, \ldots, e\} \tag{4}$$

where

- $x$ includes continuous dynamic states, for example generator angles, velocities and fluxes; discrete dynamic

Manuscript received February 17, 2000; revised Arpil 5, 2001. This research was supported by the Australian Research Council through the Grant "Estimation of Power System Parameters," the EPRI/DoD Complex Interactive Networks/Systems Initiative, and the Grainger Foundation.

The author is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

Publisher Item Identifier S 0885-8950(01)09462-7.

states, such as transformer tap positions and protection relay logic states; and parameters $\lambda$ such as generator reactances, controller gains and switching times.[1]

- $y$ are algebraic states, e.g., load bus voltage magnitudes and angles.

The model can capture complex behavior such as hysteresis, nonwindup limits and rule-based systems [12].

Away from events, system dynamics evolve smoothly according to the familiar differential-algebraic model

$$\dot{x} = f(x, y) \tag{5}$$
$$0 = g(x, y) \tag{6}$$

where $g$ is composed of $g^{(0)}$ together with appropriate choices of $g^{(i-)}$ or $g^{(i+)}$, depending on the signs of the corresponding elements of $y_d$. At switching events (3), some component equations of $g$ change. To satisfy the new $g = 0$ equation, algebraic variables $y$ may undergo a step change. Reset events (4) force a discrete change in elements of $x$ that correspond to discrete states. Algebraic variables may again step to ensure $g = 0$ is always satisfied.

The *flows* of $x$ and $y$ are defined as

$$x(t) = \phi_x(x_0, t) \tag{7}$$
$$y(t) = \phi_y(x_0, t) \tag{8}$$

where $x(t)$ and $y(t)$ satisfy (1)–(4), along with initial conditions,

$$\phi_x(x_0, t_0) = x_0 \tag{9}$$
$$g(x_0, \phi_y(x_0, t_0)) = 0. \tag{10}$$

### B. Example

Initial testing of the model evaluation process used "measurements" obtained by adding noise to simulation results [14]. However credibility of the process is best demonstrated using real measurements of an actual disturbance. Such measurements are difficult to obtain. Limited recordings were available though of a disturbance on the Nordel system [11]. That case was therefore chosen for illustrating the various concepts presented throughout the paper.[2] A simplified model of the Nordel system is shown in Fig. 1.

A full description of the disturbance sequence is given in [11]. To summarize, the major influences on system behavior were:

- a three-phase fault occurred on bus 27, with all connected lines tripped at 60 ms;
- line 3–16 tripped around 4 s later due to overloading;
- 150MVAr reactor switched in at bus 12 approximately 6 s after the initial fault was cleared; and
- the units comprising generator G2 tripped after about 8 s.

[1]Incorporating parameters $\lambda$ into the state $x$ allows a convenient development of trajectory sensitivities. To ensure that parameters remain fixed at their initial values, the corresponding differential equations (1) are defined as $\dot{\lambda} = 0$.

[2]The actual results (parameter estimates) are of little interest *per se*. The value of the example lies in the application of the model evaluation process to a real system with real (though limited) measurements.

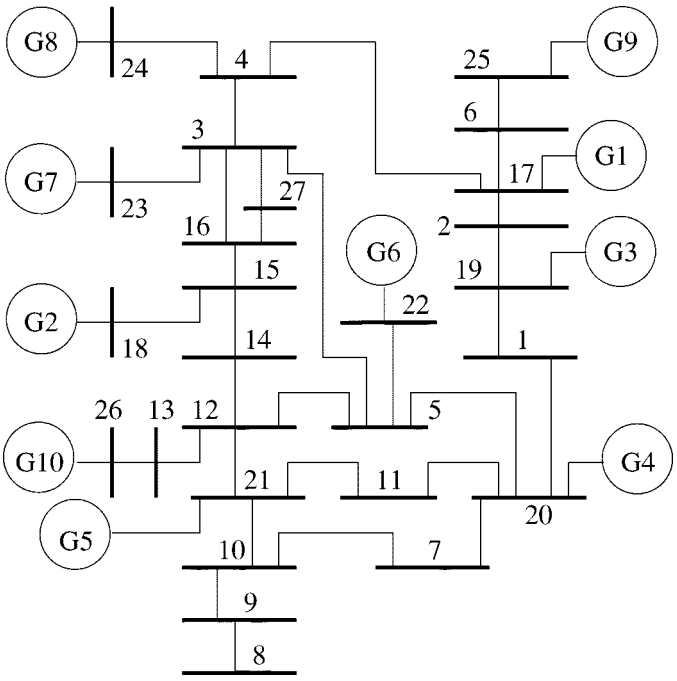
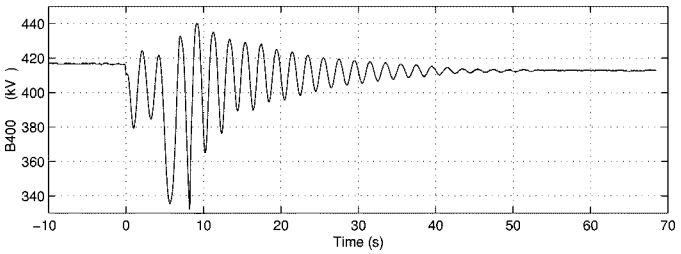
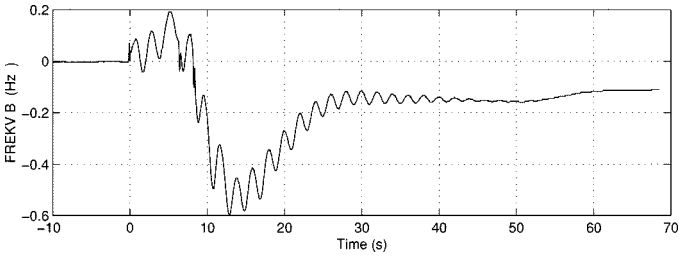



Fig. 1. Nordel system model.




(a)




(b)

Fig. 2. Disturbance recordings. (a) Voltage recording. (b) Frequency recording.

The available disturbance recordings are provided in Fig. 2. They show bus voltage magnitude and frequency deviation at bus 12. The measurements can be separated into two phases, 1) up to 8 s, during which time the system was subjected to the above event sequence, and 2) the recovery phase beyond 8 s. The system underwent a significant frequency deviation during the latter phase, in response to the loss of generator G2. Poorly damped oscillations are clearly evident during that period.

The 27 bus, 10 generator system of Fig. 1 provides a reasonably accurate (though certainly approximate) model of the Nordel 400 kV system. The initial power flow was developed using data recorded by the SCADA system just prior to the disturbance. Generators were represented by a fourth order

machine model [15], a simplified excitation system consisting of a first-order regulator and a transient gain reduction block, and a simple first-order governor/turbine model. No PSSs were modeled. This model consisted of 148 dynamics states and 494 algebraic states. Model parameters were based on [16], though the level of aggregation and simplification introduced approximations. No load model information was available, so all loads were modeled as constant admittance. That assumption is considered further in Section III-C.

This example provides a useful test case, for the following reasons:

- the disturbance was complex,
- the system was on the verge of instability,
- limited measurements were available, and
- model parameters were not known accurately.

It will be shown that even under these adverse conditions, it was possible to obtain a very good match between measurements and model response.

The example will focus on the first 10 s of the disturbance recording, i.e., the period during which the event sequence occurred, where behavior was highly nonlinear. Beyond that time, the simple model is no longer appropriate. To adequately capture the influence of the sizeable frequency deviation, more extensive governor/turbine and load models are required. Also, the poorly damped oscillations are indicative of PSSs that are not well tuned for the unusual post-disturbance power flow pattern. PSSs would be required in the model in order to match the observed damping ratio. (Note that the oscillations beyond 10 s could be analyzed using standard linear estimation techniques; see [17] for example.)

The simplified system model does not provide bus frequency, so it is not possible to directly compare model response with the frequency recording. Consequently only the voltage recording of Fig. 2(a) will be used for parameter estimation. However, the frequency measured at bus 12, shown in Fig. 2(b), will later be compared with the simulated frequency of generator G5 to ensure model reasonability.

A comparison of the measured voltage response with that predicted by the model, using the initial guess for parameter values, is given in Fig. 3. There is clearly scope for improving the propinquity.

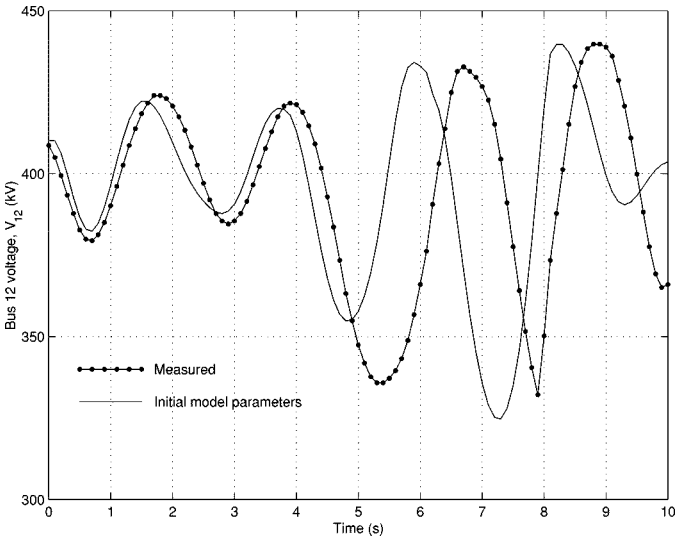



Fig. 3. Comparison of voltage recording with initial model behavior.

## III. Trajectory Sensitivities

### A. Background

In event reconstruction, an initial mismatch between measurements and model response, such as in Fig. 3, is common. To reduce the mismatch, some parameters should be altered. Unfortunately the choice of which parameters is seldom obvious. Trajectory sensitivities can be used to guide that choice.[3]

Trajectory sensitivities provide a way of quantifying the variation of a trajectory resulting from (small) changes to parameters and/or initial conditions [6]. To obtain the sensitivity of the flows $\phi_x$ and $\phi_y$ to initial conditions $x_0$, and hence to parameter variations, the Taylor series expansions of (7) and (8) are formed. Neglecting higher order terms gives

$$\Delta x(t) = \frac{\partial x(t)}{\partial x_0}\,\Delta x_0 \equiv x_{x_0}(t)\Delta x_0 \tag{11}$$

$$\Delta y(t) = \frac{\partial y(t)}{\partial x_0}\,\Delta x_0 \equiv y_{x_0}(t)\Delta x_0. \tag{12}$$

It is important to keep in mind that $x_0$ incorporates parameters $\lambda$, so sensitivity to $x_0$ includes sensitivity to $\lambda$. Equations (11) and (12) describe the changes $\Delta x(t)$ and $\Delta y(t)$ in a trajectory, at time $t$ along the trajectory, for a given (small) change in initial conditions $\Delta x_0$. The time-varying partial derivatives $x_{x_0}$ and $y_{x_0}$ are known as *trajectory sensitivities*. An overview of the variational equations describing the evolution of these sensitivities is provided in Appendix A.

Along smooth sections of the trajectory, the trajectory sensitivities evolve according to a linear time-varying differential-algebraic system (19) and (20). For large power systems, these equations have high dimension. However the computational burden is minimal when an implicit numerical integration technique such as trapezoidal integration [15] is used to generate the trajectory.[4] An overview of this result is provided in Appendix B.

More complete details of both Appendices A and B can be found in [10].

### B. Selecting Well-Conditioned Parameters

As indicated earlier, trajectory sensitivities can be used to guide the search for well-conditioned parameters, i.e., parameters that are good candidates for reliable estimation. If a parameter exerts a large influence on the trajectory of a particular state, then the corresponding trajectory sensitivity will be large. For example, referring to (11) and (12), if the parameter corresponding to the $j$th element of $x_0$ has a sizeable influence on the trajectory of state $y_i$, then the trajectory sensitivity $y_{x_0,(i,j)}(t)$

[3]Applications of trajectory sensitivities extend well beyond parameter estimation, and include optimal control [18] and boundary value problems [19].

[4]Commercial simulation packages should use implicit numerical integration, as numerical stability of explicit methods cannot be guaranteed. Such packages are therefore amenable to efficient computation of trajectory sensitivities.

will take on sizeable values over time. Note that the same parameter may exert negligible influence on some other state, say $x_k$, in which case the trajectory sensitivity $x_{x_0,(k,j)}(t)$ would be small for all time. It is common for parameters to influence some states but not others. The corresponding trajectory sensitivities would be large for the former states and small for the latter.

Large trajectory sensitivities are important because they imply the corresponding parameters have leverage in altering the model trajectory to better match the measured response. Small trajectory sensitivities, on the other hand, imply that large changes in parameter values would be required to significantly alter the trajectory. Parameters in the former category are well-conditioned, whereas the latter parameters are ill-conditioned. These intuitive concepts are formalized in Section IV-C.

Note that only parameters which influence measured states can be identified. A parameter may have a significant influence on system behavior, but if that influence is not observable in the measured states, then the parameter is not identifiable.

### C. Example (Continued)

As indicated in Section II-B, the recording of the voltage at bus 12 is to be used to evaluate model parameters. Therefore, we are primarily interested in the sensitivity of that voltage state $V_{12}$ to parameter variations. All parameters could be considered. However for the sake of the example, we shall restrict attention to the following diverse groups of parameters:[5]

- generator inertia and damping constants, Fig. 4(a),
- generator AVR setpoints, Fig. 4(b),
- real and reactive load indices, Fig. 4(c), and
- tripping times[6] for line 3–16 and generator G2, Fig. 4(d).

Considering the different scales used for sensitivity in Fig. 4(a)–(d), it is clear that some parameters exert quite an influence on the $V_{12}$ trajectory, and hence are well-conditioned, whilst other parameters have negligible influence and so are ill-conditioned.

Fig. 4(a) shows that of all the generator inertia and damping constants, only those of G2 have any significant influence on $V_{12}$. This influence of G2 is no doubt due to the major role it plays in the disturbance.

It can be seen from Fig. 4(b) that the $V_{12}$ trajectory is very sensitive to AVR setpoint values. This is due to the direct influence of AVR's on system voltages and hence on $V_{12}$. The size of the sensitivities is surprising though, with gains up to 10, i.e., a setpoint change would be amplified in $V_{12}$ by up to a factor of 10. Interestingly, the G2 sensitivity is out of phase with the others, again reflecting the unique involvement of G2 in the disturbance. The AVR setpoints were initialized from reliable SCADA information. Therefore even though the setpoints are

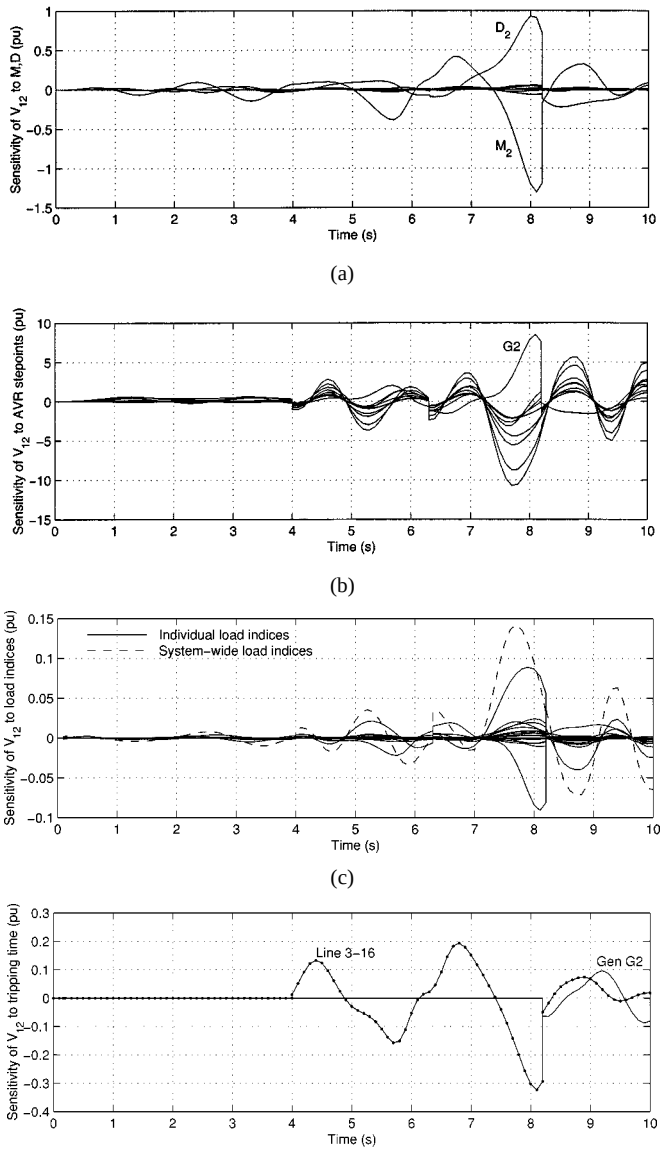



Fig. 4. Trajectory sensitivities for various parameters. (a) Generator inertia and damping constants. (b) Generator AVR setpoints. (c) Load indices. (d) Tripping times.

*potentially extremely influential, they will not be considered in the parameter estimation process.

Two load index cases were considered:

1) each individual load was modeled by a pair of real and reactive power indices, and
2) all loads were modeled by a single common pair of indices.

From Fig. 4(c), it is clear that individual load indices have very little influence on the $V_{12}$ trajectory. Even in the second case, where the parameters affect all loads jointly, their influence is quite small. (In fact the effect of the reactive power index cannot even be seen.) This provides justification for initially assuming constant admittance loads; it seems load modeling is quite unimportant in this case. That's certainly not a generic result though. Load modeling was crucial in the disturbances reported in [3], [20], [21].

[5]These 49 trajectory sensitivities were generated from a single simulation, using the approach described in Appendix B.

[6]Tripping times are not exogenous inputs, but rather outputs of protection devices. The example does not model protection, so tripping times revert to unknown parameters. In a more complete system representation that included protection, parameters of the protection model could be considered as candidates for estimation.

It is interesting to consider the influence of line and generator switching times on the $V_{12}$ trajectory. Clearly a switching time parameter cannot influence the trajectory before the actual switching event. It can be seen from Fig. 4(d) that the time at which line 3–16 trips has a reasonable influence on the $V_{12}$ trajectory. Because generator G2 doesn't trip until 8.2 s, it's tripping time has limited influence over the 10 s period of interest.

To summarize, it appears from the trajectory sensitivities that the well-conditioned parameters (apart from the AVR setpoints) are: generator G2 inertia and damping constants ($M_2$, $D_2$), line 3–16 tripping time ($t_{3\text{–}16}$), and possibly the system-wide real power load index ($np$). These parameters will be considered further in Section IV-D.

## IV. Parameter Estimation

### A. Background

Measurements of power system dynamic behavior are typically obtained using data acquisition systems (DASs) [2] that provide sequences $m_0, m_1, m_2, \ldots, m_q$ of sampled system quantities. The aim of parameter estimation is to find the set of model parameters that results in the best fit between the measurement samples and model predictions.

Usually a DAS provides measurement sequences for many different quantities. However for clarity the parameter estimation algorithm will initially be developed assuming a single measurement sequence. The more general case is presented in Section IV-E.

The development of the parameter estimation algorithm assumes that measurements correspond to algebraic states. This does not restrict the application of the algorithm though, as it is always possible to add extra algebraic equations to generate "outputs" $y_i$ that match the measurements. These new functions augment the original algebraic constraints.

### B. Parameter Estimation From a Single Measurement

The algebraic state corresponding to measurement sequence $m$ will be denoted $\breve{y}$. The estimation process involves varying a subset of parameters $\theta$ to obtain the best match between the sequence $m$ and the flow $\breve{y}(t)$ predicted by the model (8). A Gauss–Newton method is used.

The model provides the flow $\breve{y}(\theta, t)$ for all $t \geq t_0$. But the samples in the sequence $m$ are measured at certain time instants. Therefore the model is sampled at each time instant $t_k$, $k = 0, 1, \ldots, q$, resulting in the sequence $\breve{y}_0(\theta), \breve{y}_1(\theta), \ldots, \breve{y}_q(\theta)$, where $\breve{y}_k(\theta) = \breve{y}(\theta, t_k)$. The aim is to determine the value of $\theta$ that minimizes the discrepancy between the model response $\breve{y}_k(\theta)$ and the measured samples $m_k$ over all $k$.

Let the mismatch between the measured value and the model output at each sample time be

$$e_k(\theta) = \breve{y}_k(\theta) - m_k \qquad k = 0, 1, \ldots, q$$

or in vector form

$$e(\theta) = \breve{y}(\theta) - m \tag{13}$$

where

$$e(\theta) = \begin{bmatrix} e_0(\theta) \\ e_1(\theta) \\ \vdots \\ e_q(\theta) \end{bmatrix}, \qquad \breve{y}(\theta) = \begin{bmatrix} \breve{y}_0(\theta) \\ \breve{y}_1(\theta) \\ \vdots \\ \breve{y}_q(\theta) \end{bmatrix}, \qquad m = \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_q \end{bmatrix}.$$

The desired value of $\theta$ minimizes the least-squares cost

$$\mathcal{V}(\theta) = \tfrac{1}{2} \sum_{k=0}^{q} |e_k(\theta)|^2 = \tfrac{1}{2} \, ||e(\theta)||_2^2. \tag{14}$$

The problem has been reduced to a nonlinear least-squares formulation that can be solved using the Gauss–Newton method [22]. This is an iterative approach which is based on linearizing $e(\theta)$ around the point $\theta^j$,

$$\tilde{e}(\theta, \theta^j) = e(\theta^j) + \frac{\partial e(\theta^j)}{\partial \theta} (\theta - \theta^j). \tag{15}$$

From (13) it follows that

$$\frac{\partial e(\theta^j)}{\partial \theta} = \frac{\partial \breve{y}(\theta^j)}{\partial \theta} = \begin{bmatrix} \breve{y}_\theta(\theta^j, t_0) \\ \breve{y}_\theta(\theta^j, t_1) \\ \vdots \\ \breve{y}_\theta(\theta^j, t_q) \end{bmatrix} \equiv S(\theta^j) \tag{16}$$

where $\breve{y}_\theta$ is composed of the columns of trajectory sensitivities $\breve{y}_{x_0}$ that correspond to the subset of parameters $\theta$. Because $S(\theta^j)$ is formed from $\breve{y}_\theta$, evaluated at time steps $t_0, t_1, \ldots, t_q$, it shall be referred to as the *sensitivity matrix*. Substitution of $S(\theta^j)$ into (15) gives

$$\tilde{e}(\theta, \theta^j) = e(\theta^j) + S(\theta^j)(\theta - \theta^j).$$

The value of $\theta$ obtained at the $(j+1)$th iteration minimizes $\frac{1}{2}||\tilde{e}(\theta, \theta^j)||_2^2$, i.e.,

$$\theta^{j+1} = \arg\min_{\theta} \{ \tfrac{1}{2} \, ||\tilde{e}(\theta, \theta^j)||_2^2 \}.$$

Assuming $S(\theta^j)^t S(\theta^j)$ is invertible, this minimization results in the iterative scheme

$$\begin{aligned} S(\theta^j)^t S(\theta^j) \Delta\theta^{j+1} &= S(\theta^j)^t e(\theta^j) \\ &= S(\theta^j)^t \left( \breve{y}(\theta^j) - m \right) \qquad (17) \\ \theta^{j+1} &= \theta^j - \alpha^{j+1} \Delta\theta^{j+1} \qquad (18) \end{aligned}$$

where $\alpha^{j+1}$ is a scalar that determines the step size.[7] The invertibility of $S^t S$ relates directly to identifiability, and is considered further in Section IV-C.

An estimate of $\theta$ which (locally) minimizes the cost function $\mathcal{V}(\theta)$ is obtained when $\Delta\theta^{j+1}$ is close to zero. Note that this procedure will only give local minima though, as it is based on a first-order approximation of $e(\theta)$. However if the initial

[7]Equation (17) could be solved by directly inverting $S^t S$. However faster and more numerically robust algorithms are available [23].

guess for $\theta$ is good, which is generally the case in power system studies, then a local minimum is sufficient.

### C. Identifiability

The parameter estimation process (17) and (18) breaks down if the sensitivity matrix $S$ does not have full rank, i.e., if $S^tS$ is singular. It is informative to explore the connection between this observation and the concept of identifiability [4]. This discussion summarizes a more complete presentation in [24].

Rather than formally defining the concept of a model $\mathcal{M}(\theta)$, it is sufficient to think of $\mathcal{M}(\theta)$ as having the form (1)–(4), with unknown parameters $\theta$. Associated with each model $\mathcal{M}(\theta)$ is the flow

$$\phi(\theta, t) = \begin{bmatrix} \phi_x(\theta, t) \\ \phi_y(\theta, t) \end{bmatrix}.$$

Model equality can be defined as $\mathcal{M}_1(\theta_1) = \mathcal{M}_2(\theta_2)$ if $\phi_1(\theta_1, t) = \phi_2(\theta_2, t), \quad \forall t \geq t_0$. The concept of local identifiability can then be defined.

*Definition (Local Identifiability):* A model $\mathcal{M}$ is locally identifiable at $\theta^*$ if $\mathcal{M}(\theta) = \mathcal{M}(\theta^*)$ implies that $\theta = \theta^*$.

In terms of flows, this definition states that the model $\mathcal{M}$ is locally identifiable if

$$\phi(\theta, t) - \phi(\theta^*, t) \equiv \Delta\phi(t) = 0, \qquad \forall t \geq t_0$$

implies that

$$\theta - \theta^* \equiv \Delta\theta = 0.$$

Equations (7) and (8), and (11) and (12) provide the (local) connection between $\Delta\phi$ and $\Delta\theta$ in terms of trajectory sensitivities,

$$\Delta\phi(t) = \begin{bmatrix} x_\theta(t) \\ y_\theta(t) \end{bmatrix} \Delta\theta.$$

These relationships form the basis for the following theorem.

*Theorem (Identifiability From Measurements:* A model $\mathcal{M}$ is locally identifiable at $\theta^*$ from measurements $\breve{y}$ iff $S$ has full rank.

It follows from the theorem that the trajectory sensitivities $\breve{y}_\theta$ which form $S$ must be nonzero, and linearly independent. However from a practical perspective, nonsingularity of $S^tS$ is not sufficient to guarantee reliable convergence of the parameter estimation process. Convergence will be unreliable if $S^tS$ is nearly singular, i.e., if its smallest eigenvalue is very small relative to its largest eigenvalue. In such an ill-conditioned case, inversion of $S^tS$ effectively results in dividing by a very small number, leading to large entries in $\Delta\theta^{j+1}$.

Ill-conditioning of $S^tS$ can result from a column of $S$ having very small entries, i.e., negligible sensitivity $\breve{y}_{\theta_i}$ of the trajectory to a particular parameter $\theta_i$. This corroborates the more intuitive discussion of Section III-B. Ill-conditioning can also be due to the columns of $S$ being nearly linearly dependent. This situation is more difficult to identify from inspection of the trajectory sensitivities.

TABLE I
PARAMETER ESTIMATION RESULTS

| Case | Values | | | Iter. |
|---|---|---|---|---|
| | $M_2$ (pu) | $D_2$ (pu) | $t_{3\text{-}16}$ (s) | |
| $M_2, D_2$ | 1.606 | -0.143 | - | 4 |
| $M_2, D_2, t_{3\text{-}16}$ | 1.556 | 0.028 | 4.52 | 7 |
| (Initial values) | 1.146 | 0.0 | 4.0 | |

An algorithm proposed in [5] provides a reliable approach to identifying a subset of parameters that ensures $S^tS$ is well-conditioned. The algorithm consists of the following steps:

- Given a set of potentially identifiable parameters $\overline{\theta}$, form $S(\overline{\theta})^tS(\overline{\theta})$ and compute its eigen-decomposition $S(\overline{\theta})^tS(\overline{\theta}) = U\Lambda U^t$.
- Determine $\rho$ such that the largest $\rho$ eigenvalues of $S^tS$ are much larger than the remaining $n - \rho$.
- Partition $U = [U_\rho \quad U_{n-\rho}]$ with $U_\rho$ containing the first $\rho$ columns of $U$.
- Determine a permutation matrix $P$ by constructing a $QR$ decomposition with column-pivoting [23] for $U_\rho^t$,

$$U_\rho^t P = QR$$

  where $Q$ is an orthogonal matrix, and the first $\rho$ columns of $R$ form an upper triangular matrix.
- Reorder the parameter vector according to $\tilde{\theta} = P'\overline{\theta}$.
- The well-conditioned parameters $\theta$ are the first $\rho$ elements of $\tilde{\theta}$.

The results of this algorithm are compared with the observations of Section III-C in the following section.

### D. Example (Continued)

The observations of Section III-C indicated that the well-conditioned parameters consist of: generator G2 inertia and damping constants ($M_2$, $D_2$), switching time for line 3–16 ($t_{3\text{–}16}$), and possibly the system-wide real power load index ($np$). For comparison, the algorithm from the previous section was applied to a set of 24 parameters, consisting of the inertia and damping constants for all generators, system-wide real and reactive power load indices, and the switching times for line 3–16 and generator G2. It was found that the eigenvalues of $S^tS$ were

$$13.51,\ 1.21,\ 0.11,\ 0.07,\ 0.04,\ 0.02,\ \ldots,\ 9.7 \times 10^{-21}.$$

Two eigenvalues are large, implying that two parameters (at least) are well-conditioned. The algorithm selected $M_2$, $D_2$ as the best candidates for estimation. This is consistent with the trajectory sensitivities of Fig. 4. The algorithm was also used to extend the selection to the four most well-conditioned parameters. The selected parameters in this case were $M_2$, $D_2$, $t_{3\text{–}16}$ and $np$. This was again exactly consistent with the observations of Section III-C.

The results of estimating $M_2$ and $D_2$ are given in Table I and shown in Fig. 5. The fit is very good, especially given the level of modeling approximation. Certainly the fit is much better than with the original parameter values, shown in Fig. 3. The estimation algorithm converged to a parameter update tolerance

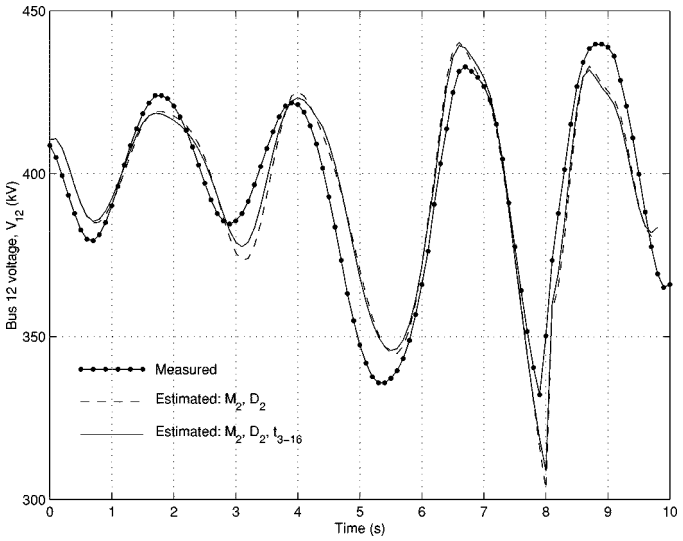



Fig. 5. Comparison of voltage recording with estimated model behavior.

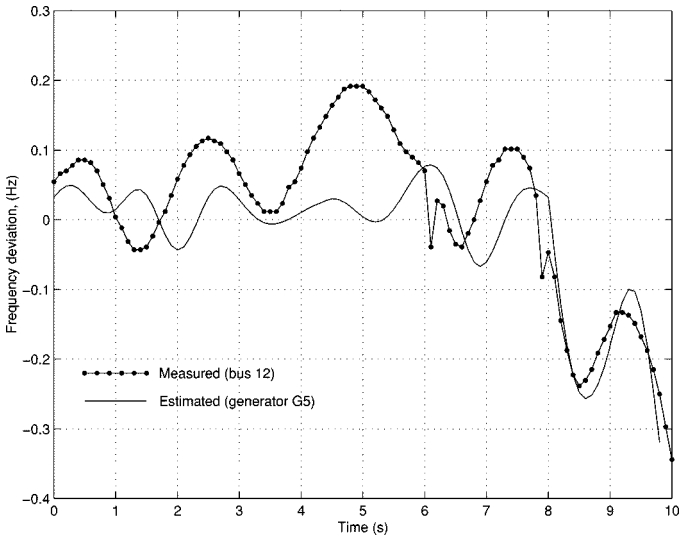



Fig. 6. Comparison of frequency recording with estimated model behavior.

of $|\Delta\theta_i| < 10^{-2}$ in 4 iterations, and a tolerance of $10^{-3}$ in 6 iterations.

The estimation of the four parameters $M_2$, $D_2$, $t_{3\text{–}16}$ and $np$ was considered next. In this case convergence had still not been obtained after 25 iterations. The first three parameters quickly converged, but $np$ continued to wander. It was found that as the first three parameters approached their optimal values, the sensitivity of $np$ diminished. It became ill-conditioned, resulting in the convergence difficulties.

Based on this experience, it was decided to estimate just the three parameters $M_2$, $D_2$ and $t_{3\text{–}16}$. Results are given in Table I and shown in Fig. 5. Again a very good fit was obtained. The estimated values were close to those of the previous case, where $np$ did not converge. This confirms that the initial guess of $np = 2$ did not bias the estimate.

Comparing the parameter estimates for the two successful cases, it appears that fixing $t_{3\text{–}16}$ biased the estimate of $D_2$, though had much less influence on the estimate of $M_2$. However Fig. 5 shows that the trajectories for the two cases are very similar.

The estimation process did not make use of the bus frequency recording given in Fig. 2(b). However that measurement was used as a cross check of model validity. As explained earlier, the simple system model did not provide bus frequency, so the frequency of the closest generator (G5) was used for comparison. Fig. 6 compares the measured (bus) frequency with the predicted (generator) frequency. The estimated values of $M_2$, $D_2$ and $t_{3\text{–}16}$ were used in this case. The smaller oscillations do not match closely, which is not surprising given the difference between bus and generator frequencies. However an important indication of model validity is given by the agreement in the large-scale deviation beyond 8 s.

It should be reiterated that the purpose of the example has been to illustrate the model evaluation process. Usually event reconstruction would have access to many more measurements, and a more elaborate system model. However, there will always be some discrepancy between measured and modeled response, due to measurement errors and model inadequacy. Separating those effects can be challenging, and generally involves exploring the influence of model structure (as against parameter values) on the quality of the measurement/model fit.

### *E. Parameter Estimation From More Than One Measurement*

Normally a DAS (or a number of DASs scattered around a power system) will provide many measurements of a disturbance. Ideally all the available information should be used to give the best estimate of the parameters.

Assume there are $\ell$ measurement sequences, $m^1$, $m^2$, …, $m^\ell$ and the corresponding model flows are $\breve{y}^1$, $\breve{y}^2$, …, $\breve{y}^\ell$. A sensitivity matrix $S^i$ corresponding to each $\breve{y}^i$ can be defined as in (16). We now define

$$\breve{y}(\theta) = \begin{bmatrix} \breve{y}^1 \\ \breve{y}^2 \\ \vdots \\ \breve{y}^\ell \end{bmatrix}, \qquad m = \begin{bmatrix} m^1 \\ m^2 \\ \vdots \\ m^\ell \end{bmatrix}$$

and make corresponding changes in the definition of $e(\theta)$ at (13). The sensitivity matrices are arranged as

$$S(\theta) = \begin{bmatrix} S^1(\theta) \\ S^2(\theta) \\ \vdots \\ S^\ell(\theta) \end{bmatrix}.$$

Again (17) and (18) can be used to obtain the (locally) optimal value of $\theta$. In this case, the optimal $\theta$ provides an overall best fit of the model to *all* the measurements.

Note that if some measurements are known more accurately than others, then weighting factors can be used to weight the relative importance of the measurements.

## V. Conclusion

The process of evaluating nonlinear dynamic models from measurements of disturbances involves, 1) identifying

parameters that can be reliably estimated, and 2) obtaining a best estimate for those parameters. Trajectory sensitivities provide information that is valuable for determining well-conditioned parameters. The paper establishes a close link between trajectory sensitivities and parameter identifiability.

A nonlinear least-squares algorithm has been proposed for estimating power system parameters from measurements of system disturbances. The algorithm is based on the Gauss–Newton iterative method. Trajectory sensitivities provide gradient information at each iteration. The algorithm is reliable, even when measurements are corrupted by significant noise.

The model evaluation process has been discussed in terms of system-wide event reconstruction. However the process is directly applicable to the development/refinement of component models, where measurements are obtained from a sequence of tests. In fact, identifiability analysis can be used to design tests that enhance parameter estimation reliability and accuracy. The refined component models form the starting point for reconstruction of more widespread system disturbances.

The estimation algorithm can be easily adapted for model reduction purposes [25]. The "measurements" in this case are replaced by trajectories of the full model. The estimated (reduced model) parameters achieve a best fit between the full and reduced models over scenarios of interest. Controller tuning can also be achieved through a conceptually similar extension. In this application, the optimal parameters result in the best fit between the system response and a pre-specified reference behavior. These and other *inverse problems* are considered in [13]. In all cases, trajectory sensitivities provide gradient information that underlies iterative algorithms.

## Appendix A
## Trajectory Sensitivity Equations

Away from events, where system dynamics evolve smoothly, the sensitivities $x_{x_0}$ and $y_{x_0}$ are obtained by differentiating (5) and (6) with respect to $x_0$. This gives

$$\dot{x}_{x_0} = f_x(t)x_{x_0} + f_y(t)y_{x_0} \tag{19}$$

$$0 = g_x(t)x_{x_0} + g_y(t)y_{x_0} \tag{20}$$

where $f_x \equiv \partial f/\partial x$, and likewise for the other Jacobian matrices. Note that $f_x$, $f_y$, $g_x$, $g_y$ are evaluated along the trajectory, and hence are time varying matrices. It is shown in Appendix B that the solution of this (potentially high order) time-varying DAE system can be obtained as a by-product of solving the original DAE system (5) and (6).

Initial conditions for $x_{x_0}$ are obtained from (9) as

$$x_{x_0}(t_0) = I$$

where $I$ is the identity matrix. Initial conditions for $y_{x_0}$ follow directly from (20),

$$0 = g_x(t_0) + g_y(t_0)y_{x_0}(t_0).$$

Equations (19) and (20) describe the evolution of the sensitivities $x_{x_0}$ and $y_{x_0}$ between events. However at an event, the sensitivities are generally discontinuous. It is necessary to calculate *jump conditions* describing the step change in $x_{x_0}$ and $y_{x_0}$. For clarity, consider a single switching/reset event, so the model (1)–(4) reduces to the form

$$\dot{x} = f(x, y) \tag{21}$$

$$0 = \begin{cases} g^-(x, y) & s(x, y) < 0 \\ g^+(x, y) & s(x, y) > 0 \end{cases} \tag{22}$$

$$x^+ = h(x^-, y^-) \quad s(x, y) = 0. \tag{23}$$

Let $(x(\tau), y(\tau))$ be the point where the trajectory encounters the hypersurface $s(x, y) = 0$, i.e., the point where an event is triggered. This point is called the *junction point* and $\tau$ is the *junction time*. Assume that the trajectory encounters the triggering hypersurface transversally.

Just prior to event triggering, at time $\tau^-$, we have

$$x^- = x(\tau^-) \ = \phi_x(x_0, \tau^-)$$
$$y^- = y^-(\tau^-) = \phi_y(x_0, \tau^-)$$

where

$$g^-(x^-, y^-) = 0.$$

Similarly, $x^+$, $y^+$ are defined for time $\tau^+$, just after the event has occurred. It is shown in [10] that the jump conditions for the sensitivities $x_{x_0}$ are given by

$$x_{x_0}(\tau^+) = h_x^* x_{x_0}(\tau^-) - \left(f^+ - h_x^* f^-\right)\tau_{x_0} \tag{24}$$

where

$$h_x^* = \left(h_x - h_y(g_y^-)^{-1}g_x^-\right)\Big|_{\tau^-}$$

$$\tau_{x_0} = -\frac{\left(s_x - s_y(g_y^-)^{-1}g_x^-\right)\Big|_{\tau^-} x_{x_0}(\tau^-)}{\left(s_x - s_y(g_y^-)^{-1}g_x^-\right)\Big|_{\tau^-} f^-}$$

$$f^- = f\left(x(\tau^-), y^-(\tau^-)\right)$$

$$f^+ = f\left(x(\tau^+), y^+(\tau^+)\right).$$

The sensitivities $y_{x_0}$ immediately after the event are given by

$$y_{x_0}(\tau^+) = -\left(g_y^+(\tau^+)\right)^{-1} g_x^+(\tau^+)x_{x_0}(\tau^+).$$

Following the event, i.e., for $t > \tau^+$, calculation of the sensitivities proceeds according to (19) and (20), until the next event is encountered. The jump conditions provide the initial conditions for the post-event calculations.

## Appendix B
## Efficient Trajectory Sensitivity Computation

Consider the DAE system (5) and (6) which describes behavior over the periods between events. The trapezoidal approach to numerical integration approximates the differential equations (5) by a set of algebraic difference equations. These algebraic equations are coupled with the original algebraic equations (6) giving

$$x^{k+1} = x^k + \frac{\eta}{2}\left(f(x^k, y^k) + f(x^{k+1}, y^{k+1})\right) \tag{25}$$

$$0 = g(x^{k+1}, y^{k+1}) \tag{26}$$

where the superscripts $k$, $k+1$ index the time instants $t_k$, $t_{k+1}$ respectively, and $\eta = t_{k+1} - t_k$ is the integration time step.

Equations (25) and (26) describe the evolution of the states $x$, $y$ from time instant $t_k$ to the next time instant $t_{k+1}$.

Notice that (25) and (26) form a set of implicit nonlinear algebraic equations. To solve for $x^{k+1}$, $y^{k+1}$ given $x^k$, $y^k$ requires the use of a nonlinear equation solver. Newton-based iterative techniques are commonly used. The solution process involves forming and factorizing the Jacobian

$$F_{\mathcal{H}} = \begin{bmatrix} \frac{\eta}{2} f_x - I & \frac{\eta}{2} f_y \\ g_x & g_y \end{bmatrix}. \tag{27}$$

Now consider the sensitivity equations (19) and (20). Using trapezoidal integration, they are approximated by

$$x_{x_0}^{k+1} = x_{x_0}^k + \frac{\eta}{2}\left(f_x^k x_{x_0}^k + f_y^k y_{x_0}^k + f_x^{k+1} x_{x_0}^{k+1} + f_y^{k+1} y_{x_0}^{k+1}\right)$$
$$0 = g_x^{k+1} x_{x_0}^{k+1} + g_y^{k+1} y_{x_0}^{k+1}.$$

Rearranging gives

$$\begin{bmatrix} \frac{\eta}{2} f_x^{k+1} - I & \frac{\eta}{2} f_y^{k+1} \\ g_x^{k+1} & g_y^{k+1} \end{bmatrix} \begin{bmatrix} x_{x_0}^{k+1} \\ y_{x_0}^{k+1} \end{bmatrix} = \begin{bmatrix} -\frac{\eta}{2}\left(f_x^k x_{x_0}^k + f_y^k y_{x_0}^k\right) - x_{x_0}^k \\ 0 \end{bmatrix}. \tag{28}$$

Therefore, $x_{x_0}^{k+1}$ and $y_{x_0}^{k+1}$ are obtained as the solution of a linear matrix equation. But notice that the matrix to be inverted in solving (28) is exactly the Jacobian (27) used in solving for $x^{k+1}$ and $y^{k+1}$. Because that matrix has already been built and factorized to calculate $x^{k+1}$ and $y^{k+1}$, the solution of (28) involves little extra computation.

To improve simulation speed, (25) and (26) are often solved using a quasi-Newton method. As a result, the factors of $F_{\mathcal{H}}$ may not be available for solving (28) directly. However a number of computationally efficient techniques have been proposed in [26]–[28].

## REFERENCES


[1] D. N. Kosterev, C. W. Taylor, and W. A. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Trans. Power Systems*, vol. 14, no. 3, pp. 967–979, Aug. 1999.

[2] J. F. Hauer, W. A. Mittelstadt, W. H. Litzenberger, C. Clemens, D. Hamai, and P. N. Overholt, "Wide area measurements for real-time control and operation of large electric power systems," DOE Final Report, Apr. 1999.

[3] R. H. Craven, T. George, G. B. Price, P. O. Wright, and I. A. Hiskens, "Validation of dynamic modeling methods against power system response to small and large disturbances," in *Proceedings of CIGRÉ General Session*, Paris, Aug. 1994.

[4] L. Ljung, *System Identification. Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[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 Trans. Power Systems*, vol. 14, no. 1, pp. 218–225, Feb. 1999.

[6] P. M. Frank, *Introduction to System Sensitivity Theory*. New York: Academic Press, 1978.

[7] J. J. Sanchez-Gasca, C. J. Bridenbaugh, C. E. J. Bowler, and J. S. Edmonds, "Trajectory sensitivity based identification of synchronous generator and excitation system parameters," *IEEE Trans. Power Systems*, vol. 3, no. 4, pp. 1814–1822, Nov. 1988.

[8] 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.

[9] C. Merckx, K. Karoui, J. Dubois, and M. Stubbe, "Optimization methods for parameter identification and controller tuning," in *Proceedings of the 12th CEPSI*, Pattaya, Thailand, Nov. 1998.

[10] I. A. Hiskens and M. A. Pai, "Trajectory sensitivity analysis of hybrid systems," *IEEE Trans. Circuits and Systems I*, vol. 47, no. 2, pp. 204–220, Feb. 2000.

[11] I. A. Hiskens and M. Akke, "Analysis of the Nordel power grid disturbance of January 1, 1997 using trajectory sensitivities," *IEEE Trans. Power Systems*, vol. 14, no. 3, pp. 987–994, Aug. 1999.

[12] I. A. Hiskens and M. A. Pai, "Hybrid systems view of power system modeling," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Geneva, Switzerland, May 2000.

[13] I. A. Hiskens, "Power system modeling for inverse problems," IEEE Trans. Circuits and Systems I, submitted for publication.

[14] I. A. Hiskens and A. J. A. Koeman, "Power system parameter estimation," *Journal of Electrical and Electronics Engineering, Australia*, vol. 19, no. 1, pp. 1–8, June 1999.

[15] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*. Upper Saddle River, NJ: Prentice-Hall, 1998.

[16] M. Akke, "Power system stabilizers in multimachine system," Licentiate Thesis CODEN:LUTFD2/ TFRT-3201, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, 1989.

[17] J. F. Hauer, "The use of Prony analysis to determine modal content and equivalent models for measured power system response," in *Eigenanalysis and Frequency Domain Methods for System Dynamic Performance*: IEEE Publication 90 TH0292-3-PWR, 1989, pp. 105–115.

[18] F. L. Lewis and V. L. Syrmos, *Optimal Control*, 2nd ed. New York: Wiley, 1995.

[19] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 2nd ed. New York: Springer, 1993.

[20] R. H. Craven and M. R. Michael, "Load representation in the dynamic simulation of the Queensland power system," *Journal of Electrical and Electronic Engineering Australia*, vol. 3, no. 1, pp. 1–7, 1983.

[21] I. A. Hiskens and J. Milanović, "Locating dynamic loads which significantly influence damping," *IEEE Trans. Power Systems*, vol. 12, no. 1, pp. 255–261, Feb. 1997.

[22] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995.

[23] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed: John Hopkins, 1989.

[24] I. A. Hiskens, "Identifiability of hybrid system models," in *Proceedings of the 9th IEEE Conference on Control Applications*, Anchorage, AK, Sept. 2000.

[25] R. J. Galarza, J. H. Chow, W. W. Price, A. W. Hargrave, and P. M. Hirsch, "Aggregation of exciter models for constructing power system dynamic equivalents," *IEEE Trans. Power Systems*, vol. 13, no. 3, pp. 782–788, Aug. 1998.

[26] W. F. Feehery, J. E. Tolsma, and P. I. Barton, "Efficient sensitivity analysis of large-scale differential-algebraic systems," *Applied Numerical Mathematics*, vol. 25, pp. 41–54, 1997.

[27] S. Li, L. Petzold, and W. Zhu, "Sensitivity analysis of differential-algebraic equations: A comparison of methods on a special problem," *Applied Numerical Mathematics*, vol. 32, pp. 161–174, 2000.

[28] D. Chaniotis, M. A. Pai, and I. A. Hiskens, "Sensitivity analysis of differential-algebraic systems using the GMRES method—Application to power systems," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Sydney, Australia, May 2001.



**Ian A. Hiskens** (S'77–M'80–SM'96) received the B.Eng.(Elec) and B.App.Sc.(Math) degrees from the Capricornia Institute of Advanced Education, Rockhampton, Australia in 1980 and 1983, respectively. He received the Ph.D. degree from the University of Newcastle, Australia in 1990. He was with the Queensland Electricity Supply Industry from 1980 to 1992, and a Senior Lecturer at the University of Newcastle from 1992 to 1999. He is currently a Visiting Associate Professor in the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign. His major research interests lie in the area of power system analysis, in particular system dynamics, security, and numerical techniques. Other research interests include nonlinear systems and control.