Subset Selection for Improved Parameter Estimation in On-Line Identification of a Synchronous Generator

MICHAEL BURTH† GEORGE C. VERGHESE ‡ MIGUEL VÉLEZ-REYES‡‡

†Technical University Berlin ‡Massachusetts Institute of Technology
‡‡University of Puerto Rico

Abstract — **This paper examines subset selection for nonlinear least squares parameter estimation, and applies the methodology to a test system previously studied in the power system literature, involving the on-line identification of a synchronous generator model with many parameters. Subset selection partitions the parameters into well-conditioned and ill-conditioned subsets. We show for the test system that fixing the ill-conditioned parameters to prior estimates (even if these prior estimates are substantially in error), and estimating only the remaining parameters, significantly improves the performance of the estimation algorithm and greatly enhances the quality of the estimated parameters. It is shown that attempts to estimate all of the model parameters, as done in the original work with this test system, can yield extremely unreliable results.**

Key Words — **Least Squares, Estimation, Identification, Conditioning, Subset Selection, Synchronous Generator.**

I. INTRODUCTION

Parameter estimation is a key step in fitting a model to measurements, and is therefore central to the task of system identification. The parameter estimates are chosen so as to minimize a measure of the errors between: (i) the model's predictions of what values the available measurements will take, and (ii) the actual values taken by the measurements. Methods of solving this minimization problem are dependent on the structure of the model and on the error criterion. For models whose predictions are linear in the parameters, and with an error measure that is the sum of squared prediction errors, efficient and stable *linear least squares* estimation techniques are available (see, e.g. [1, 2, 3]). For models that are *nonlinear* in the parameters, least squares estimation involves iterative methods, of which the Gauss-Newton *iterated linearized* least squares method is among the most used [1, 4, 5, 6].

Several aspects of the model and measurements affect the performance of a parameter estimation algorithm and determine the quality of the estimates that are produced. What happens very often in the power system setting is that a component model is built up from detailed analysis of the underlying physical phenomena, and therefore involves a relatively large number of physically interesting and interpretable parameters; this is certainly the case with the synchronous generator models found in the literature, [7, 8]. On the other hand, the measurements available from online experiments in an interconnected power system — which is the setting that is ultimately of interest — are typically *not* rich enough to adequately reflect the individual effects of all the parameters in the various components of the system. This mismatch between the (high) detail of the models and the (low) richness in the measurements leads to very sensitive or ill-conditioned parameter estimation problems.

The purpose of this paper is to point out the manifestations of such ill-conditioning in the context of parameter estimation for a synchronous generator model, and to suggest a strategy for overcoming the ill-conditioning. Our results apply more broadly than to synchronous generator identification, but this particular identification problem is important enough to have been studied fairly extensively in the power systems literature — see for example [9, 10, 11, 12, 13, 14, 15, 16] and references therein — and therefore provides a fruitful context for our study. In particular, the test system that we use is taken from [16]. Some mention of conditioning in the generator identification problem appears in [11, 13, 14], but the theme does not seem to have been developed in any detail prior to now.

Our strategy for overcoming ill-conditioning is based on the *subset selection* approach proposed in [17, 18] for nonlinear least squares parameter estimation, extending subset selection for linear least squares estimation as described in [2]. Subset selection partitions the model parameters into *well-conditioned* parameters, which are likely to be estimated reliably from the given measurements, and *ill-conditioned* parameters, whose estimates are likely to be unreliable, and whose presence makes the parameter estimation problem very sensitive. Given this partitioning, we proposed in [17, 18] to fix the ill-conditioned parameters at prior estimates, in effect abandoning any attempt to estimate them from the available measurements, and to then solve a reduced-order and well-conditioned parameter estimation problem to determine the remaining parameters. This strategy is successful if the bias introduced by fixing the ill-conditioned parameters to prior estimates is more than made up for by the improvement in estimation of the remaining parameters. The application of this strategy to the estimation of induction machine speed and parameters in [17, 18] led to major performance improvements over full-order estimation.

We demonstrate in this paper, for the synchronous generator

PE-465-PWRS-0-04-1998 A paper recommended and approved by the IEEE Power System Dynamic Performance Committee of the IEEE Power Engineering Society for publication in the IEEE Transactions on Power Systems. Manuscript submitted August 4, 1997; made available for printing April 24, 1998.

model used in the identification experiments of [16], and for qualitatively similar "measurements" (synthesized from the model in [16], since the original data was not available to us), that the proposed strategy leads to a reduced-order estimation procedure and associated parameter estimates that are much better behaved than if all the parameters are estimated together. Section II of the paper briefly reviews the nonlinear least squares problem, highlighting the role of the Jacobian (or gradient or first derivative matrix) of the error vector with respect to the parameter vector in finding the least squares estimate by the Gauss-Newton method; the Hessian (or second derivative matrix) of the error criterion with respect to the parameter vector is also defined. The section then gives the main idea behind the subset selection procedure, applied to the Jacobian or Hessian, and finally specifies the algorithm in more detail. Section III describes the test system, explains how the "measurements" for our identification experiments were synthesized, and then presents the results of our various estimation experiments on the system. Some conclusions are stated in Section IV.

II. Least Squares Estimation and Subset Selection

A. The Nonlinear Least Squares Problem

Least squares fitting of a model to experimental data is a common procedure in engineering. In this method, the parameters of a model are determined such that they minimize the sum of squares of the components of the N-component error (or "residual") vector

$$\mathbf{r}(\theta) = \widehat{\mathbf{y}}(\theta) - \mathbf{y} \,, \tag{1}$$

where θ denotes the n-vector of model parameters, $\widehat{\mathbf{y}}(\theta)$ is the N-vector of model predictions for the measurements, and $\mathbf{y}$ is the N-vector of actual measurements. Stated mathematically, the vector of parameter estimates is specified as

$$\widehat{\theta} = \arg\min_{\theta} V(\theta) \,, \tag{2}$$

where the minimization criterion for least squares is defined by

$$V(\theta) = \frac{1}{2}\|\mathbf{r}(\theta)\|^2 = \frac{1}{2}\sum_{\ell=1}^{N} r_\ell^2(\theta) \,, \tag{3}$$

with $r_\ell(\theta)$ denoting the ℓth component of the error vector, and the factor of $\frac{1}{2}$ being simply for convenience in some later expressions. Minimization of the above criterion by the Gauss-Newton method involves iterated *linearization* of the nonlinear problem around the current best guess of the parameter estimates, and therefore requires the $N \times n$ *Jacobian* or gradient or matrix of first partial derivatives of the error vector with respect to the parameter vector:

$$\mathbf{J}(\theta) = \frac{\partial \mathbf{r}(\theta)}{\partial \theta} \,. \tag{4}$$

The Gauss-Newton method begins with an initial guess of the parameter estimate, say $\widehat{\theta}_0$. The next guess is then computed as

$$\widehat{\theta}_1 = \widehat{\theta}_0 + \alpha_1 \mathbf{p}_1 \,, \tag{5}$$

where α_1 is a scalar (of the order of unity) that fixes the step size in the Gauss-Newton *direction* $\mathbf{p}_1$. This direction is computed by solving a *linear* least squares problem associated with the linearization of the original problem around the initial guess. Specifically, $\mathbf{p}_1$ satisfies the so-called "normal" equations:

$$(\mathbf{J}'\mathbf{J})\,\mathbf{p}_1 = \mathbf{J}'\,\mathbf{r} \,, \tag{6}$$

where $'$ denotes matrix transposition, and both $\mathbf{J}$ and $\mathbf{r}$ in (6) are evaluated at the current estimate $\widehat{\theta}_0$. In principle, $\mathbf{p}_1$ could be found by inverting the matrix $(\mathbf{J}'\mathbf{J})$ that pre-multiplies it in the above equation, but from a numerical point of view there are better methods of actually computing $\mathbf{p}_1$, see [2]. Note that $(\mathbf{J}'\mathbf{J})$ is invertible, and correspondingly $\mathbf{p}_1$ is uniquely determinable, if and only if the n columns of $\mathbf{J}$ are independent. From the definition in (4), we see that this condition is equivalent to requiring that increments in the various parameters should perturb the error vector in n independent directions (in N-space).

The step size α_1 in (5) may be picked so as to obtain (close to) the greatest possible decrease of the criterion $V(\cdot)$ by movement in the specified direction. Once $\widehat{\theta}_1$ has been found, the entire procedure is repeated, but now linearizing around $\widehat{\theta}_1$. This iteration is continued until the desired degree of convergence has been achieved. A more detailed algorithm description may be found in [4]; a clear summary in the context of generator identification is given in [12].

Other Newton-type approaches to solving the minimization problem (2) involve the *Hessian*, which is the $n \times n$ matrix of second partial derivatives of the error criterion $V(\theta)$ with respect to the parameter vector θ, and is easily seen to be given by

$$\mathbf{H}(\theta) = \mathbf{J}'(\theta)\mathbf{J}(\theta) + \sum_{\ell=1}^{N} r_\ell(\theta)\,\frac{\partial r_\ell(\theta)}{\partial\theta\,\partial\theta'} \,. \tag{7}$$

For small residuals, the Hessian can evidently be approximated by

$$\mathbf{H}(\theta) \approx \mathbf{J}'(\theta)\mathbf{J}(\theta) \,, \tag{8}$$

which is the matrix on the left side of (6). In the remainder of this paper, we shall use the term "Hessian" and the symbol $\mathbf{H}$ to refer to this approximation of the strict Hessian, namely $\mathbf{J}'\mathbf{J}$.

Efficient methods exist for computing the Jacobian when the data is modeled as comprising time samples of the output of a state-space model, which is the case with our test system. For a detailed description of these methods, we refer the reader to [4]. Explicit treatments of the application of least squares parameter estimation in this manner to the identification of synchronous machines, including the calculation of gradient functions, can be found in [9, 12, 14].

B. Parameter Conditioning

The Hessian matrix $\mathbf{H} = \mathbf{J}'\mathbf{J}$ on the left of the normal equations (6) is symmetric and positive semidefinite, so all its eigenvalues are real and non-negative. Suppose $\mathbf{H}$ is actually singular, with just one eigenvalue at 0 and some associated eigenvector; this happens if and only if the n columns of $\mathbf{J}$ actually contain only $n-1$ independent vectors. An immediate implication of the singularity is that the step direction computed from the normal equations (6) can be varied in the direction of this eigenvector of $\mathbf{H}$ without affecting the error criterion (at least to first order). Such indeterminacy would be highly undesirable in a physical parameter estimation problem, because it would indicate that the parameters cannot be unambiguously estimated from the given measurements. Note that such parameter indeterminacy can exist even if the model predictions fit the measurements exactly, i.e. even if

$V(\widehat{\theta}) = 0$; the Hessian involves the second partials of $V(\cdot)$, not the value of $V(\cdot)$ itself. Thus the fact that one gets a small error between the model predictions and the measurements is not sufficient validation of the quality of the parameter estimates.

It is typically the case that $\mathbf{H}$ is not exactly singular. However, often $\mathbf{H}$ is *nearly* singular, in the sense that its smallest eigenvalue is very small relative to its largest eigenvalue. Such a situation corresponds to the n columns of $\mathbf{J}$ being nearly dependent, i.e., being almost confined to a subspace of dimension less than n in Euclidean N-space. This situation is also undesirable, because it reflects *near* indeterminacy in the parameter estimates, caused by having more parameters than can be reliably estimated from the available measurements. Nearness to singularity is usually measured by the *condition number*, [2], which (in the case of a symmetric, positive definite matrix) is the ratio of the largest to smallest eigenvalues; we shall denote the condition number of the Hessian by $\kappa(\mathbf{H})$.

[Our discussion could equivalently be framed in terms of the so-called *singular values* and associated singular vectors of $\mathbf{J}$, [2], rather than the *eigenvalues* and eigenvectors of $\mathbf{H}$, but we use the latter because eigenvalues and eigenvectors are more familiar objects. From a numerical point of view, it is actually preferable to work with the singular value decomposition (SVD) of $\mathbf{J}$ rather than to compute $\mathbf{H}$ according to (8) and then determine its eigendecomposition, but this is not a serious issue for the exposition here.]

There are several related manifestations of a high condition number in the Hessian. For example, from the definition of $\mathbf{H}$ as the matrix of second partial derivatives of the error criterion $V(\theta)$, it follows that the eigenvalues of the Hessian describe the *curvature* of the error criterion in the directions of the associated eigenvectors. A high condition number for the Hessian indicates that the error criterion varies much more slowly with θ in some directions than in others. The implication for parameter estimation is that the parameter vector is more poorly determined in directions where the curvature it small, relative to directions with high curvature.

Yet another consequence of a high condition number, which can be deduced from (6), is that a small fractional change in the error $\mathbf{r}$ can make a large fractional change in the step direction, with the ratio of these fractional changes possibly being as large as the condition number. Finally, a high condition number can result in a large number of iterations to convergence in the Gauss-Newton algorithm.

C. Subset Selection and Reduced-Order Estimation

The strategy suggested in [17, 18] for improving the behavior of nonlinear least squares parameter estimation is to determine which parameter axes lie closest to the ill-conditioned directions of the Hessian, and to fix the associated parameter values at prior estimates throughout the iterative estimation process. If the Hessian has ρ large eigenvalues and $n-\rho$ small ones, then what we do in the update step (5) is fix $n-\rho$ appropriately chosen components of the step direction vector to be 0, so that the associated parameters do not change. The resulting normal equations will then only involve the Jacobian of the error with respect to the ρ remaining or "active" parameters; we denote this Jacobian by $\mathbf{J}_\rho$, and note that its columns are just a subset of the columns of $\mathbf{J}$, namely those corresponding to the active parameters. The associated Hessian is $\mathbf{H}_\rho = \mathbf{J}'_\rho \mathbf{J}_\rho$.

Subset selection is aimed at recognizing which ρ parameters to keep active so as to obtain a corresponding Hessian $\mathbf{H}_\rho$ with as small a condition number as possible. A combinatorial search would be prohibitively expensive computationally. The following algorithm, which is essentially the one specified in [17] (which in turn is derived from the subset selection algorithm for linear least squares described in [2]), is much cheaper, and yields very good results:

Algorithm (Subset Selection and Reduced-Order Estimation)

- *Given an initial parameter vector estimate $\widehat{\theta}_0$, compute the eigendecomposition of $\mathbf{H}(\widehat{\theta}_0)$, yielding $\mathbf{H} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'$.*
- *Determine ρ such that the first ρ eigenvalues of $\mathbf{H}$ are much larger than the remaining $n-\rho$ ones.*
- *Make the partition $\mathbf{V} = [\ \mathbf{V}_\rho \quad \mathbf{V}_{n-\rho}\ ]$ with $\mathbf{V}_\rho$ containing the first ρ columns of $\mathbf{V}$.*
- *Determine a permutation matrix $\mathbf{P}$ by constructing a QR decomposition with column-pivoting, [2], for $\mathbf{V}'_\rho$, i.e. determine*
$$\mathbf{V}'_\rho \mathbf{P} = \mathbf{Q}\mathbf{R} \tag{9}$$
where $\mathbf{Q}$ is an orthogonal matrix, and the first ρ columns of $\mathbf{R}$ form an upper triangular matrix.
- *Use $\mathbf{P}$ to reorder the parameter vector θ according to $\widetilde{\theta} = \mathbf{P}'\theta$*
- *Make the partition $\widetilde{\theta} = [\ \widetilde{\theta}'_\rho \quad \widetilde{\theta}'_{n-\rho}\ ]'$ with $\widetilde{\theta}_\rho$ containing the first ρ elements of $\widetilde{\theta}$. Fix $\widetilde{\theta}_{n-\rho}$ to a prior estimate $\widehat{\widetilde{\theta}}_{n-\rho}$.*
- *Compute $\widehat{\widetilde{\theta}}$ by solving $\widehat{\widetilde{\theta}} = \arg\min_{\widetilde{\theta}} V(\widetilde{\theta})$ subject to $\widetilde{\theta}_{n-\rho} = \widehat{\widetilde{\theta}}_{n-\rho}$.*

[It is assumed that the initial estimate $\widehat{\theta}_0$ is accurate enough for the Algorithm to provide a subset selection that does not differ significantly from one based on the Hessian evaluated at the optimal estimate. If desired, the Algorithm can be restarted at any stage with the current best estimate of the parameters, to check if the subset selection has changed.]

III. Application to a Test System

A. The System Model

Our test system is extracted from the description in [16], and involves the following small-signal model of a 100 MVA turbogenerator connected to an infinite bus. The model consists of incremental equations describing the synchronous machine around an operating point. The d-axis incremental equations in standard notation, [7, 8], and as presented in [16], are

$$\begin{aligned}\frac{d\Delta E'_q}{dt} &= \frac{1}{T'_{do}}\Delta E'_q - \frac{x_d - x'_d}{T'_{do}x''_d}\Delta E''_q \\ &\quad + \frac{x_d - x'_d}{T'_{do}x''_d}\Delta V_q + \frac{k}{T'_{do}}\Delta V_{fd}\end{aligned} \tag{10}$$

$$\begin{aligned}\frac{d\Delta E''_q}{dt} &= \left(\frac{1}{T''_{do}} - \frac{1}{T'_{do}}\right)\Delta E'_q \\ &\quad - \left(\frac{1}{T''_{do}} + \frac{x_d - x'_d}{T'_{do}x''_d} + \frac{x'_d - x''_d}{T''_{do}x''_d}\right)\Delta E''_q\end{aligned} \tag{11}$$

$$+\left(\frac{x_d - x'_d}{T'_{do}x''_d} + \frac{x'_d - x''_d}{T''_{do}x''_d}\right)\Delta V_q + \frac{k}{T'_{do}}\Delta V_{fd}$$

$$\Delta i_d = \frac{1}{x''_d}(\Delta E''_q - \Delta V_q) \tag{12}$$

The inputs to the above equations are ΔV_q and ΔV_{fd}, while Δi_d is a measured output. The parameters to be estimated are x_d, x'_d, x''_d, T'_{do}, T''_{do} and $k = x_{ad}/R_{fd}$. Similarly, the incremental equations along the q-axis are

$$\frac{d\Delta E''_d}{dt} = -\frac{x_q}{T''_{qo}x''_q}\Delta E''_d - \frac{x_q - x''_q}{T''_{qo}x''_q}\Delta V_d \tag{13}$$

$$\Delta i_q = \frac{1}{x''_q}\Delta E''_d + \frac{1}{x''_q}\Delta V_d\ . \tag{14}$$

The input is ΔV_d, Δi_q is a measured output, and x_q, x''_q and T''_{qo} are the parameters to be estimated.

The generator is assumed to be connected to an infinite bus of voltage magnitude V_B through a line reactance x_e; these two quantities are parameters to be estimated. The connection imposes the phasor relationships represented in Fig. 1, where I denotes the current flowing from the generator into the infinite bus, V_t denotes the generator's terminal voltage, and ϕ, θ, and the power angle δ_t are as defined in the figure. Linearizing these relationships around an operating point leads to incremental equations relating ΔI, $\Delta\phi$, ΔV_t, $\Delta\theta$, ΔV_d and ΔV_q.

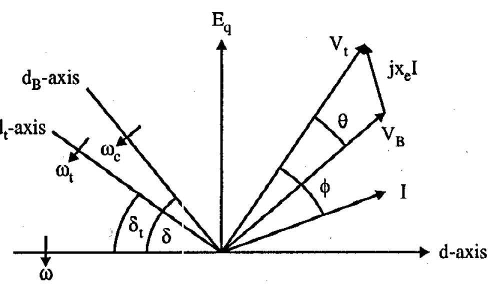

Figure 1: **Phasor relationships.**

The incremental motion equations of the system are

$$\frac{d\Delta\omega}{dt} = -\frac{D}{H}\Delta\omega - \frac{\omega_s}{2H}\Delta P \tag{15}$$

$$\frac{d\Delta\phi}{dt} = \Delta\omega \tag{16}$$

$$\Delta\delta_t = \Delta\phi - \Delta\theta\ , \tag{17}$$

where ω is the angular velocity of the rotor, ω_s is synchronous frequency, $P = V_d i_d + V_q i_q$ is the electrical power out of the generator, H is the inertia constant of the machine set and D is a mechanical damping coefficient. The inputs to the above equations are ΔP and $\Delta\theta$, while $\Delta\delta_t$ is a measured output. Both H and D are parameters to be estimated.

Putting together the incremental relationships above and eliminating the appropriate variables, we obtain a small-signal state-space model for the system in the standard form

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \tag{18}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)\ , \tag{19}$$

with state vector, output measurements and input respectively given by

$$\mathbf{x} = [\ \Delta E'_q \quad \Delta E''_q \quad \Delta E''_d \quad \Delta\omega \quad \Delta\phi\ ]' \tag{20}$$

$$\mathbf{y} = [\ \Delta i_d \quad \Delta i_q \quad \Delta\delta_t \quad \Delta V_d \quad \Delta V_q\ ]' \tag{21}$$

$$u = \Delta V_{fd}\ . \tag{22}$$

For convenience, we specify the operating point around which the small-signal model is constructed by specifying the active power P, reactive power $Q = V_q i_d - V_d i_q$, and the angle α_B from V_B to the d-axis. Our numerical experiments used the operating condition corresponding to $P = 1.0$ p.u., $Q = 0.1$ p.u., and $\alpha_B = 70°$.

B. Formulation of Estimation Experiments

A nonlinear least squares parameter estimation approach for on-line identification of synchronous machines is presented in [16], along with the results of application to several actual large generators. With the generator connected to a large power system and loaded normally, one applies a sudden change of excitation ΔV_{fd} through some appropriate means. The transients in the line voltages V_{ab}, V_{cb}, phase currents i_a, i_b and i_c, field voltage V_{fd} and power angle δ_t are recorded, and used as the basis for parameter estimation.

Table 1 lists the parameters that have to be estimated. All the parameters listed in the table, with the exception of V_B and x_e, were identified in [16] by applying the Gauss-Newton method, using the electrical equations (10-14) and the mechanical equations (15-17) to model the data. The electrical subsystem model involves 9 parameters; what we shall demonstrate is that estimating all 9 of these parameters without fixing a subset of them can yield extremely unreliable results for the given estimation problem.

Since the experimental data from [16] were not available to us, and also to permit a more controlled assessment and comparison of parameter estimation strategies, we used the system model from the previous subsection to produce synthetic "measurements" that serve as the basis for our estimation studies. To generate these "measurements", we took as the "true" or underlying parameters the final identified parameters from [16], listed in the bottom row of Table 1, chose a reasonable operating point, and excited our model with a field voltage perturbation ΔV_{fd} — shown in Fig. 2 — that yielded "measured" waveforms qualitatively similar to the experimental measurements in [16]. These "measured" waveforms, which serve as inputs to our estimation process, are Δi_d, Δi_q, ΔV_d, ΔV_q and $\Delta\delta_t$, shown in Figures 3-5. (The waveforms in Figure 4 are actually shown with some additive noise imposed, at levels used later in the paper to study the sensitivity of the parameter estimates to noise.)

The reader should keep in mind that our purpose is to compare *methodologies* rather than to make specific numerical comparisons of identified parameters with the results in [16]. Therefore, the fact that for "true" parameters we use values that were the result of the potentially ill-conditioned estimation process in [16] should not be a real concern, particularly since the values listed in the bottom row of Table 1 are physically quite reasonable.

For comparison, the manufacturer's nameplate values for most of these parameters, as quoted in [16], are listed above the "true" values in the table. We used these manufacturer's values as initial estimates for the parameter values; the starting values for the four parameters not provided by the manufacturer were chosen arbitrarily.

Table 1: **Parameter Values.** The second row (Manuf.) lists the parameter values declared in [16] as provided by the manufacturer. The third row ("True") lists the parameters estimated in [16], and used as "true" values in synthesizing the "measurements" for our experiments.

	x_d (p.u.)	x_d' (p.u.)	x_d'' (p.u.)	T_{do}' (sec)	T_{do}'' (sec)	k	x_q (p.u.)	x_q'' (p.u.)	T_{qo}'' (sec)	H (sec)	D	V_B (p.u.)	x_e (p.u.)
Manuf.	1.806	0.286	0.183	6.2	0.24	1271	-	-	-	9	3	1.22	-
"True"	1.414	0.333	0.208	5.85	0.194	1552	1.302	0.396	0.955	11.2	1.89	0.99	0.016

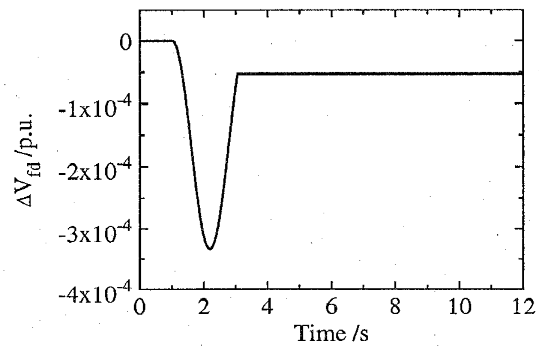

Figure 2: **Transient field voltage ΔV_{fd}.**

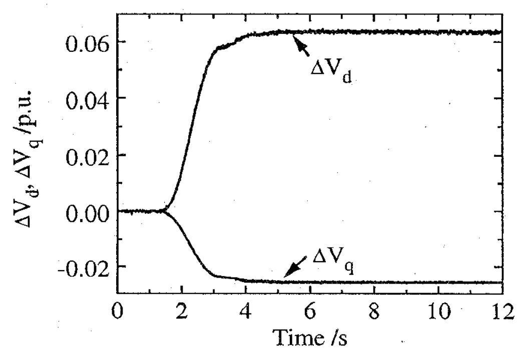

Figure 4: **Transient voltages ΔV_d, ΔV_q (with added noise).**

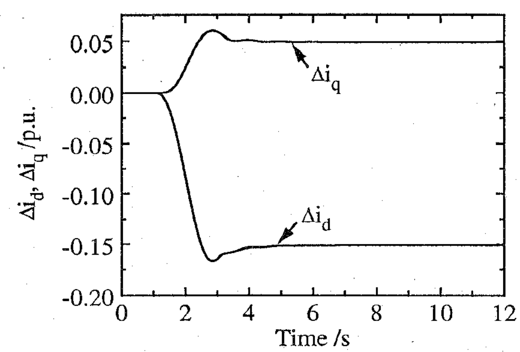

Figure 3: **Transient currents Δi_d, Δi_q.**

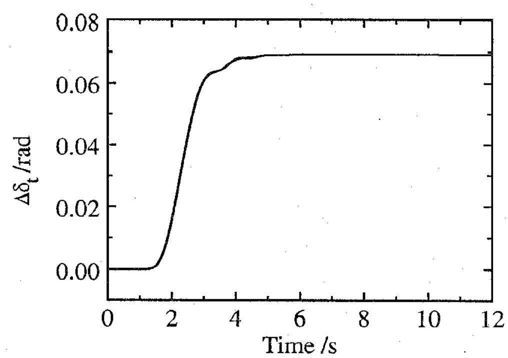

Figure 5: **Transient power angle $\Delta\delta_t$.**

C. Estimation of V_B, x_e, H and D

The electrical subsystem of the test model is the focus of our studies of subset selection, but we first deal briefly with the estimation of V_B, x_e, H and D, before concentrating on the electrical subsystem in the next subsection.

In [16] it is suggested that V_B and x_e be determined by linear regression from the following relation:

$$V_t = b_1 V_B^2 + b_2 x_e^2 + b_3 x_e \ , \tag{23}$$

where $b_1 = 1/V_t$, $b_2 = -I^2/V_t$ and $b_3 = 2I\cos(\pi/2 - \phi)$. Measurements provide us with the values of V_t, I and ϕ. However, since (23) is nonlinear in x_e, a nonlinear parameter estimation method should be used to get reliable results. Furthermore, the estimation problem is ill-conditioned when the measured transients are small compared to their operating-point values. We therefore propose estimating V_B and x_e by the following two steps, involving incremental equations.

- By linearizing (23) around the operating point, x_e can be determined from

$$\Delta V_t = \frac{a_1 x_e + a_2 x_e^2}{a_3 + a_4 x_e} \tag{24}$$

  through application of a nonlinear estimation method, e.g. Gauss-Newton. The a_i's in (24) involve known operating point values and the transients of δ_t, i_d and i_q.

- Once x_e has been estimated, V_B can be directly calculated from the relationships in Figure 1.

The estimation of H and D was done by applying the Gauss-Newton method to the state-space incremental model (15-17). The condition of the associated Hessian is $\kappa(\mathbf{H}) = 79.3$, which indicates a rather well-conditioned estimation problem. As might be expected from this conditioning, both H and D can be estimated *reliably.*

D. Estimation of the Electrical Parameters

In this section we show that estimation of the parameters involved in the electrical subsystem (10-14) can be improved significantly by subset selection. The parameter vector to be estimated is

$$\theta = \left[\begin{array}{ccccccccc} x_d & x'_d & x''_d & T'_{do} & T''_{do} & k & x_q & x''_q & T''_{qo} \end{array} \right]' , \quad (25)$$

as listed in Table 1. The full-order Hessian evaluated at the "true" parameters possesses the following eigenvalues:

$$86.9,\ 4.94,\ 1.20,\ 0.670,\ 0.126,\ 0.0252, \quad (26)$$
$$0.0168,\ 3.46 \times 10^{-5},\ 6.47 \times 10^{-13},$$

and has a condition number of 1.34×10^{14}, which indicates *severe* ill-conditioning. [Evaluating the Hessian at other nearby sets of parameters yields qualitatively similar results; for our purposes, it suffices to evaluate it at the "true" values.] It is evident that the last two eigenvalues are comparatively very small; there is a huge gap between the 7th and the 8th eigenvalues. This suggests that there are two ill-conditioned parameters whose values we should fix at prior estimates in order to get a well-conditioned problem.

The application of our Algorithm yields the following partitioning of parameters into 7 well-conditioned parameters and 2 ill-conditioned ones:

$$\widetilde{\theta} = \left[\begin{array}{ccccccc|cc} x''_q & T''_{qo} & x_q & T''_{do} & x''_d & x'_d & x_d & k & T'_{do} \end{array} \right]' . \quad (27)$$

Since k and T'_{do} have been shifted to the last two positions of the reordered parameter vector $\widetilde{\theta}$, these are the parameters to fix.

Rather than looking at the results for just the case of two fixed parameters, it is illuminating to examine the whole range of possibilities. We therefore applied our Algorithm for all choices of ρ from 9 down to 1, to determine which $9-\rho$ parameters to fix. The remaining unfixed parameters were estimated using noise-free "measurements", and with a termination criterion of the form

$$V(\widehat{\theta}) < c\,, \quad (28)$$

with c being a small constant threshold.

Table 2 shows: (i) the number of iterations needed to meet the termination criterion; (ii) the condition number $\kappa(\mathbf{H}_\rho)$ of the reduced-order problem obtained by fixing the parameters specified by the Algorithm; and (iii) the ideal reduced-order condition number $\kappa_\rho(\mathbf{H}) = \lambda_1(\mathbf{H})/\lambda_\rho(\mathbf{H})$, with $\lambda_i(\mathbf{H})$ denoting the ith-largest eigenvalue of $\mathbf{H}$. This ideal reduced-order condition number is what would be obtained if we held constant the $n-\rho$ worst *linear combinations* of parameters, rather than the $n-\rho$ *individual* parameters. The close approximation of $\kappa_\rho(\mathbf{H})$ by $\kappa(\mathbf{H}_\rho)$ reflects how effective our Algorithm is at picking the $n-\rho$ worst individual parameters.

It can be seen that the fixing of only two parameters decreases the number of iterations by approximately a factor of 10, and reduces the condition number $\kappa(\mathbf{H}_\rho)$ by 11 orders of magnitude compared to the full-order case. Fixing more than two parameters gives modest further improvements. These results imply that fixing the two parameters k and T'_{do} is reasonable.

The preceding results were for noise-free "measurements". To assess the effects of noise on the estimation process, normally distributed zero-mean white noise signals were added to the noise-free "measurement" signals in the electrical model (10-14). The variances of the noise signals were chosen to be 0.5% of the steady-state magnitudes of the quantities they were added to. Figure 4 shows as an example the disturbed versions of ΔV_d and ΔV_q.

With the disturbed signals, 20 runs of parameter estimation were carried out for both the full-order case and the reduced-order case (with k and T'_{do} fixed to their "true" values). The results are presented in Table 3. The mean errors and the standard deviations of the estimated parameters, expressed as percentages of their "true" values, are used as indicators of the quality of the parameter estimations.

It can be seen that in the full-order case there are huge mean deviations (up to 791% in x_d") from the "true" values, whereas the maximum mean deviation for the reduced-order case is only 9.7% (in T''_{do}). Furthermore, the reproducibility of parameter estimates for the reduced-order case is very high, i.e. the normalized standard deviations are very small compared to those in the full-order case. This shows clearly the great improvement achieved by fixing a subset of only two parameters. It also shows that parameter estimation without fixing the ill-conditioned parameters can be absolutely unreliable in the case of noisy measurements.

It should be emphasized that all the parameter estimation runs used the same termination criterion (28). This means that all the sets of estimated parameters, in both the full-order and reduced-order cases, produce an objective function $V(\theta) < c$ for some fixed small c, i.e. all the final error vectors $\mathbf{r}(\theta)$ have small magnitude. That is why the the magnitude of $\mathbf{r}(\theta)$ cannot be used as the only measure of evaluation for parameter estimation, as is done in [16] and elsewhere in the literature. This measure is only reasonable in conjunction with other features like those used in Table 3.

In a practical application, parameters cannot be fixed to their "true" values because these are unknown. Instead an *a priori* estimate must be used. To illustrate the effects of fixing k and T'_{do} to values different from their "true" values, the following cases were investigated:

Case A k and T'_{do} vary $\pm 5\%$ from their "true" values
Case B k and T'_{do} vary $\pm 10\%$ from their "true" values
Case C k and T'_{do} vary $\pm 20\%$ from their "true" values

For each of the above cases, 20 runs of parameter estimation were carried out using the same noisy input and output signals as before. The mean relative errors of the estimated parameters with respect to their "true" values are presented in Table 4.

Comparing the results in Tables 3 and 4, one can see that fixing the ill-conditioned parameters to values different from their "true" values results in biased estimates of the remaining parameters. However, even in Case C, the estimated parameters are much better than the parameters obtained for the full-order case. The standard deviations for Cases A–C are not presented because they are very similar to those in the bottom row of Table 3, i.e. they are again very small compared to the full-order case.

Table 3 as well as Table 4 show that the parameters x_q, x''_q and T''_{qo} are always identified very well. One could have expected this by looking at the reordered parameter vector $\widetilde{\theta}$ in (27). The mentioned parameters hold the first three positions, which indicates that they are the best-conditioned of the parameters.

Even if the fixed parameters differ from their "true" values, the corresponding biased parameter estimates obtained from the reduced-order estimation can predict the system behavior more

Table 2: **Successive Fixing of Parameters.** Parameter estimation with noise-free "measurements" and a successive fixing of parameters to their design values. $\kappa(\mathbf{H}_\rho)$ is the condition of the Hessian achieved by our Algorithm. The last column shows the ideal reduced-order condition number $\kappa_\rho(\mathbf{H}) = \lambda_1(\mathbf{H})/\lambda_\rho(\mathbf{H})$.

Fixed Parameters	Number of Iterations	$\kappa(\mathbf{H}_\rho)$	$\kappa_\rho(\mathbf{H})$
0	223	1.34×10^{14}	1.34×10^{14}
1	229	2.51×10^{6}	2.51×10^{6}
2	22	5.17×10^{3}	5.17×10^{3}
3	24	3.46×10^{3}	3.45×10^{3}
4	23	796.35	690.70
5	23	216.35	129.67
6	17	72.91	72.41
7	16	19.49	17.61
8	10	1.00	1.00

Table 3: **Disturbed Data.** 20 runs of parameter estimation with noisy simulated "measurements" for the full-order and reduced-order cases, respectively. Row *err* shows the mean relative errors of the estimated parameters with respect to their "true" values (in %). Row *std* shows the normalized standard deviations of the estimated parameters (in %), normalized by the "true" values.

		x_d	x'_d	x''_d	T'_{do}	T''_{do}	k	x_q	x''_q	T''_{qo}
err	full	340.83	358.51	791.36	-97.17	-74.17	-61.16	-0.07	5.33	2.05
	reduced	-0.20	-0.89	8.74	-	9.96	-	0.07	5.09	2.43
std	full	18.47	12.18	24.18	19.85	12.00	8.83	0.44	0.85	1.18
	reduced	0.08	0.25	1.41	-	2.38	-	0.02	0.39	0.24

precisely than the full-order estimates. To show this, the state variables of the electrical model were simulated using the mean estimates underlying Tables 3 and 4. To slightly shift the generator's operating point, the excitation ΔV_{fd} was magnified by a factor of 1.5 for this investigation.

Figure 6 shows as an example the simulated state variable $\Delta E''_q$ for the full-order case as well as Case C. One can see that the state variables for Case C are much closer to the real state, which is the dashed line, than the state variable corresponding to the full-order case.

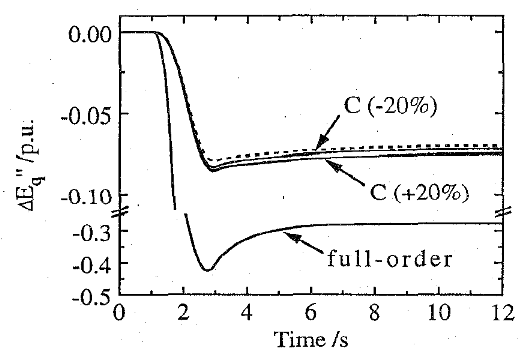

Figure 6: **Transient state variables $\Delta E''_q$ in modified experiment, and with errors in fixed parameters.**

IV. Conclusion

This paper has examined subset selection for nonlinear parameter estimation, and illustrated its application to identifying a synchronous generator model with many parameters. A reduced-order and well-conditioned estimation problem was obtained by fixing certain ill-conditioned parameters to prior estimates. Fixing just two carefully chosen parameters of the nine-parameter electrical model led to major improvements in estimation performance — in terms of numbers of iterations as well as standard deviations of the estimated parameters — compared to the full-order case, especially in the presence of added noise.

It is our belief that much work remains to be done in the area of matching model complexity to the quality of the available measurements in power systems, and in showing how to use the resulting models for various types of systems studies. As interconnected power systems move towards deregulation, probably with less sharing of information among the various players, the need for sound approaches to on-line identification will become increasingly felt, and the notion of parameter conditioning will almost certainly play an important role in the development of these approaches.

Acknowledgments

The first author is grateful to his thesis supervisor, Prof. D. Filbert, for encouraging and facilitating his extended visit to MIT, and to SIEMENS AG for support during this time. The second author is grateful for partial support from Electricité de France and the National Science Foundation (NSF) under Grant ECS-9634767 that keeps him involved in power

Table 4: **Effects of Errors in Fixed Parameters.** 20 runs of parameter estimation with noisy "measurements" for Cases A–C. The table shows mean relative errors of the estimated parameters with respect to their "true" values (in %).

Case	x_d	x'_d	x''_d	T''_{do}	x_q	x''_q	T''_{qo}	T'_{do}	k
A	1.92	-0.28	9.14	12.07	0.07	5.09	2.43	5	
	-2.34	-1.53	8.23	7.69	0.07	5.09	2.43	-5	
B	4.04	0.28	9.49	14.02	0.07	5.09	2.43	10	
	-4.50	-2.22	7.79	5.22	0.07	5.09	2.43	-10	
C	8.22	1.31	10.11	17.52	0.07	5.09	2.43	20	
	-8.88	-3.75	6.57	-0.40	0.07	5.09	2.43	-20	

systems research. The third author has been partially supported by the NSF under Grant ECS-9702860.

REFERENCES

[1] A. BJÖRCK, "Least Squares Methods", *Handbook of Numerical Analysis*, Vol. 1, pp. 466-645, 1990.

[2] G.H. GOLUB and C.F. VAN LOAN, *Matrix Computations*, John Hopkins, 2nd edition, 1989.

[3] C.L. LAWSON and R.J. HANSON, *Solving Least Squares Problems*, Prentice Hall, 1974.

[4] Y. BARD, *Nonlinear Parameter Estimation*, Academic Press, 1974.

[5] J.E. DENNIS and R.B. SCHANABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice Hall, 1983.

[6] G.A.F. SEBER and C.J. WILD, *Nonlinear Regression*, John Wiley & Sons, 1989.

[7] A.R. BERGEN, *Power Systems Analysis*, Prentice Hall, 1986.

[8] P. KUNDUR, *Power System Stability and Control*, McGraw-Hill, 1994.

[9] C.C. LEE and O.T. TAN, "A Weighted-Least-Squares Parameter Estimator For Synchronous Machines", *IEEE Trans. Power Apparatus and Systems*, Vol. PAS-96, No. 1, pp. 97-101, January/February 1977.

[10] M. NAMBA, T. NISHIWAKI, S. YOKOKAWA, K. OHTSUKA, Y. UEKI, "Identification of Parameters for Power System Stability Analysis Using Kalman Filter", *IEEE Trans. Power Apparatus and Systems*, Vol. PAS-100, No. 7, pp. 3304-3310, July 1981.

[11] E. EITELBERG and R.G. HARLEY, "Estimating Synchronous Machine Electrical Parameters from Frequency Response Tests", *IEEE Trans. Energy Conversion*, Vol. EC-2, No. 1, pp. 132-136, March 1987.

[12] 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. PWRS-3, No. 4, pp. 1814-1820, November 1988.

[13] L.X. LE and W.J. WILSON, "Synchronous Machine Parameter Identification: A Time Domain Approach", *IEEE Trans. Energy Conversion*, Vol. EC-3, No. 2, pp. 241-248, June 1988.

[14] C.-T. HUANG, Y.-T. CHEN, C.-L. CHANG, C.-Y. HUANG, H.-D. CHIANG and J.-C. WANG, "On-Line Measurement-Based Model Parameter Estimation For Synchronous Generators", *IEEE Trans. Energy Conversion*, Vol. EC-9, No. 2, pp. 330-343, June 1994.

[15] J. MA, B.W. HOGG, N. ZHIYUAN and Y. YIHAN, "On-Line Decoupled Identification of Transient and Sub-Transient Generator Parameters," *IEEE. Trans. Power Systems*, Vol. PWRS-9, No. 4, 1908-1914, November 1994.

[16] Z.M. ZHAO, F.S. ZHENG, J.D. GAO and L.Y. XU, "A Dynamic On-Line Parameter Identification and Full-Scale System Experimental Verification for Large Synchronous Machines", *IEEE Trans. Energy Conversion* Vol. EC-10, No. 3, pp. 392-398, September 1995.

[17] M. VÉLEZ-REYES, "Decomposed Algorithms for Parameter Estimation", PhD thesis, Massachusetts Institute of Technology, 1992.

[18] M. VÉLEZ-REYES and G.C. VERGHESE, "Subset Selection in Identification and Application to Speed and Parameter Estimation for Induction Machines", *Proceedings of the 4th IEEE Conference on Control Applications*, pp. 991-997, Albany, September 1995.

BIOGRAPHIES

Michael Burth received his Diploma Engineer in 1994 from the Technical University Berlin (Germany). Since 1995 he has been working on his doctoral studies at the Institute for Measurement Technology and Control Engineering at the same university. He contributed to this paper during a period as a visiting engineer at MIT's EECS Department in summer 1996. His research interests include technical diagnosis of low power electric motors, parameter identification, and modern signal processing methods.

George C. Verghese (S '74, M '78, SM '97) received his B.Tech. from IIT, Madras, India, his M.S. from SUNY, Stony Brook, and his Ph.D. from Stanford University. He joined the EECS Department at MIT in 1979, where he is now Professor of Electrical Engineering, attached to the Laboratory for Electromagnetic and Electronic Systems. He has served as an Associate Editor of *Automatica* and of the *IEEE Trans. Automatic Control*, and is currently an Associate Editor of the *IEEE Trans. Control Systems Technology*. He has also served on the AdCom of the IEEE Power Electronics Society. He is coauthor of *Principles of Power Electronics*, Addison-Wesley, 1991.

Miguel Vélez-Reyes received his B.S. from UPR Mayagüez in 1985, M.S. and E.E. degrees from MIT in 1988, and Ph.D. from MIT in 1992. He joined the ECE Department at the University of Puerto Rico Mayagüez Campus in 1992 where he is an Associate Professor and Associate Researcher at the Center for Computing Research and Development. His research interests are in the areas of identification, estimation, and nonlinear control.