J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*

POWER SYSTEM STATE ESTIMATION

Today's power systems need to be monitored in real time so that system conditions can be monitored at the energy control centers as closely as possible and proper operation and control actions can be taken in a timely manner. Utilities install Supervisory Control and Data Acquisition systems (*SCADA*) to bring the real-time monitoring data to their control centers. As illustrated in Fig. 1, typical real-time data include status of devices, especially the circuit breakers (*CB*), power generation from the power plants, load at important substations, and power flows at important lines. Since it is not possible to monitor every quantity in the system, and since the telemetered data may contain errors from time to time, one of the challenging tasks at the control center becomes "estimating" the current operating conditions of the system based on these telemetered data as reliably as possible.

Fred Scheweppe, who introduced state estimation to power system in 1968 (1), defined the state estimator as "a data processing algorithm for converting redundant meter readings and other available information into an estimate of the state of an electric power system." As this definition indicates, the state estimator (*SE*) helps to track the real-time state of the system as reliably as possible. The real-time system data provided by the SE is used in many of the on-line applications at the control center, such as security analysis, economic dispatch, and study of possible supervisory control strategies, like switching operations. Today, the SE is an essential tool in real-time monitoring of power systems.

The SE acts as a filter for cleaning up the errors in telemetered data. The real-time data obtained through the SCADA are contaminated due to various reasons, such as inaccurate transducer calibration and noise in communication channels. The filtering done by the SE is achieved mainly by making use of the relationship between the measurements and the state of the system. Several references on state estimation have been published (2,3,4). The basic introduction to SE in the next three sections draws heavily from these sources. Then new developments on state estimation are reviewed.

Problem Formulation

The SE estimates the state of the system based on a set of real-time measurements. The state of the system defines the operating conditions of the system and is usually chosen as the bus voltage magnitudes and angles. Thus, for a system of N buses, the state of the system x contains N bus voltage magnitudes and $N - 1$ bus voltage angles (one of the bus voltage angle is chosen as the reference). That is,

$$x = [\theta_2, \ldots, \theta_N \quad V_1, \ldots, V_N]^T \tag{1}$$

Note that once the state of the system is determined (in this case all the bus voltages), then any quantity in the system, such as line flows and power injection at any bus, can be calculated.

Since it is not feasible to measure and telemeter the system state directly (especially the relative phase angles of the bus voltages), in practice, other variables—mostly the ones that are important for the monitoring

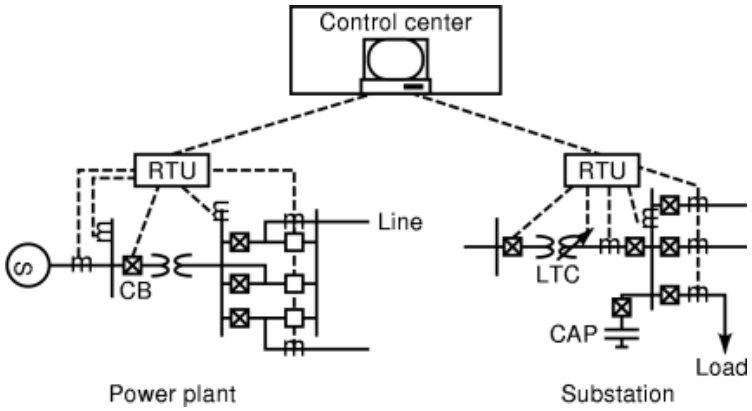

Fig. 1. Real-time system monitoring through SCADA.

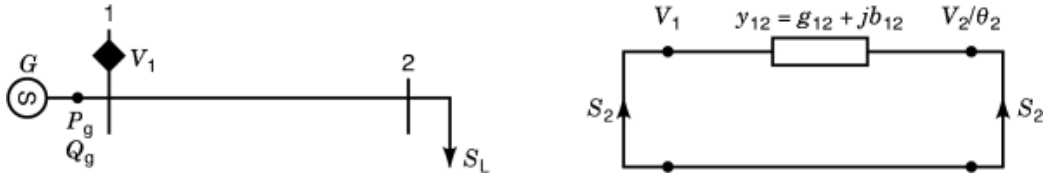

Fig. 2. A small system with measurements for SE.

of the system, such as power generation from generation plants, load at important substations, and power flows at important lines—are telemetered to the control center. Let z denote the set of such measurements. As indicated before, the SE estimates the system state x based on these measurements by making use of the relationships between the measurements and the state variables. Consider, for example, the measurements on the small system in Fig. 2.

If we represent the line by its total series impedance per phase y_{12}, then the system model for state estimation becomes the one shown in Fig. 2(b), and thus the measurements can be expressed in terms of the state variables $x = [\theta_2 V_1 V_2]^T$ as

$$z_i = h_i(x) \qquad i = 1, \ldots, m \tag{2}$$

where the function $h_i(x)$ is called the *measurement function* corresponding to the measurement z_i. For example, the measurement functions for the measurements of the system in Fig. 2 are

$$\begin{aligned} z_1 = P_g^m &\rightarrow h_1(x) = V_1^2 g_{12} - V_1 V_2 (g_{12} \cos(\theta_{12}) + b_{12} \sin(\theta_{12})) \\ z_2 = Q_g^m &\rightarrow h_2(x) = -V_1^2 b_{12} - V_1 V_2 (g_{12} \sin(\theta_{12}) - b_{12} \cos(\theta_{12})) \\ z_3 = V_1^m &\rightarrow h_3(x) = V_1 \end{aligned} \tag{3}$$

Note that the equality in Eq. (2) would hold if the measurement and the model were perfect. However, the actual measurements and the model may contain errors. Assuming that we can have an accurate system

model, Eq. (2) can be modified to include the measurement errors as

$$z_i = h_i(x) + e_i \qquad i = 1, \ldots, m \tag{4}$$

where e_i represents the error in measurement i. Hence, for the general case, we have a set of m equations relating m measurements to n states:

$$z = h(x) + e \tag{5}$$

Note that the assumptions underlying this *SE* model are: (1) the system is 3-phase balanced, (2) system topology is known, (3) system model (parameters of lines and transformers) are known, and (4) all the measurements are made at the same time (no time-skew).

State estimation involves estimating the system state $\hat{x}$ by using the preceding set of nonlinear equations. These equations are usually overdetermined (more measurements than number of states; i.e., $m > n$) to improve the accuracy of the estimation and to accommodate cases in which some of the measurements become unavailable due to meter loss. Also, the error terms are assumed to be varying randomly with some known statistical properties. There are various approaches to solve this type of problem. The conventional approach adopted for the power system state estimation—the weighted least squares (*WLS*) method—is presented first in the next section; other approaches are summarized later.

WLS Estimator

The basic approach of the WLS method is to find an estimate of the system state $\hat{x}$ that minimizes the *measurement residuals*

$$r_i = z_i - h_i(x) \qquad i = 1, \ldots, m \tag{6}$$

by solving the following optimization problem:

$$\min_x J(x) = \sum_{i=1}^{m} w_i (z_i - h_i(x))^2 = [z - h(x)]^T W [z - h(x)] \tag{7}$$

where w_i represents the weight associated with measurement z_i. Weights are chosen in proportion to the accuracy of the measurements; the higher the accuracy of a measurement, the greater its weight. Note that this approach corresponds to the solution of the overdetermined system of equations

$$z = h(x) \tag{8}$$

which tries to find a state $\hat{x}$ that fits the data as closely as possible.

The solution of the optimization problem of Eq. (7) gives the estimated state $\hat{x}$ that must satisfy the following necessary optimality condition:

$$\frac{\partial J(x)}{\partial x} = 0 \Rightarrow H^T(\hat{x}) W [z - h(\hat{x})] = 0 \tag{9}$$

where

$$[H(x)]_{i,j} = \frac{\partial h_i(x)}{\partial x_j} \tag{10}$$

is the Jacobian matrix of the measurement function $h(x)$. The solution of the nonlinear equation [Eq. (9)] is usually obtained by an iterative method derived by linearizing $h(x)$ about x^k. The iteration starts with an initial guess for the state x^0, and at each iteration k, a set of linear equations of the following type, called the *normal equations* of the WLS problem, is solved to calculate the correction Δx^k:

$$[G(x^k)]\Delta x^k = H^T(x^k)W[z - h(x^k)] \tag{11}$$

where $G(x)$ is called the gain matrix and is usually chosen as

$$G(x) = H^{\mathrm{T}}(x)WH(x) \tag{12}$$

Then the state is updated:

$$x^{k+1} = x^k + \Delta x^k \tag{13}$$

If this updated state makes the *mismatches* [optimality condition of Eq. (9)] sufficiently small, then it is accepted as the solution $\hat{x}$; otherwise the iteration is repeated.

To illustrate this procedure, consider the small system in Fig. 2, and let the line admittance be $y_{12} = g_{12} + jb_{12} = 0.01 + j0.05$ p.u. (per unit) and the measurements be $z = [P^m{}_g Q^m{}_g V^m{}_1] = [1.0\ 0.9\ 1.0]pu$. Starting from a flat starting point for the state $x^0 = [\theta_2 V_1 V_2]^T = [0.0\ 1.0\ 1.0]$ and letting the weights ω_i for the measurements be 1 for the power measurements and 10 for the voltage measurement (assuming that the voltage measurement is more accurate than the power measurements), the iterative procedure yields

$$\hat{x} = \begin{bmatrix} \theta_2 \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} -2.485^o \\ 1.0 \\ 0.946 \end{bmatrix} \Rightarrow \begin{array}{l} \hat{V}_1 = 1/\underline{0} \\ \hat{V}_2 = 0.946/\underline{-2.485^o} \end{array} \text{ p.u.}$$

If we calculate the measurement residuals $r = z - h(\hat{x})$ for this example, they will be all zero. This will indicate that the data (the measurements) fit the model perfectly, which is a result of the fact that there is no measurement redundancy (i.e., the number of measurements are equal to the number of states), and therefore the SE will not filter the noise in measurements, in this case. The ability of SE to filter measurement noise is a very important feature and will be explored further later in a section entitled "Bad Data Processing."

Fast Decoupled WLS Estimator. Since the SE is expected to run as fast as possible (today the typical execution time is about 1 to 3 min), the early efforts toward decreasing the computational burden of the WLS SE led to the development of fast decoupled WLS estimators (5,6). These methods, as the fast decoupled power flow methods, make use of the well-known weak coupling between real and reactive power: that the real power flows mainly affect the bus phase angles, while the reactive power flows mainly affect the bus voltage magnitudes. To make use of this feature in WLS estimation, partition the measurement set z into two groups: one containing mainly the real power measurements z_p and the other reactive power and voltage magnitude measurements z_q, $z = [z_p z_q]$. Let the corresponding measurement vector be $\boldsymbol{h} = [h_p h_q]$, and partition the state vector into two parts as $\boldsymbol{x} = [\theta V]$. The weak coupling implies that the sensitivities of $h_p(x)$ to bus voltages are

much smaller than to angles and, similarly, sensitivities of $h_q(x)$ to bus angles are much smaller than to bus voltages. That is,

$$\left|\frac{\partial h_{pi}(x)}{\partial \theta_j}\right| \gg \left|\frac{\partial h_{pi}(x)}{\partial V_j}\right| \quad \left|\frac{\partial h_{qi}(x)}{\partial V_j}\right| \gg \left|\frac{\partial h_{qi}(x)}{\partial \theta_j}\right| \tag{14}$$

Therefore, the measurement Jacobian matrix can be partitioned and approximated as follows:

$$\begin{aligned} H &= \begin{bmatrix} \dfrac{\partial h_p(x)}{\partial \theta} & \dfrac{\partial h_p(x)}{\partial V} \\ \dfrac{\partial h_q(x)}{\partial \theta} & \dfrac{\partial h_q(x)}{\partial V} \end{bmatrix} \cong \begin{bmatrix} \dfrac{\partial h_p(x)}{\partial \theta} & 0 \\ 0 & \dfrac{\partial h_q(x)}{\partial V} \end{bmatrix} \\ &= \begin{bmatrix} H_{p\theta} & 0 \\ 0 & H_{qv} \end{bmatrix} \end{aligned} \tag{15}$$

This approximation is used to simplify the normal equations of Eq. (11) in various ways. In the so-called algorithm decoupled version, the right-hand side of the normal equations, the mismatch terms, are calculated exactly at each iteration, but the gain matrix is decoupled by using the decoupled H as

$$G = \begin{bmatrix} G_{p\theta} & 0 \\ 0 & G_{qv} \end{bmatrix} = \begin{bmatrix} H_{p\theta}^T W_p H_{p\theta} & 0 \\ 0 & H_{qv}^T W_q H_{qv} \end{bmatrix} \tag{16}$$

Furthermore, these matrices are constructed and computed once at the beginning of the iterations and then are kept constant in the rest of the iterations. Thus, the normal equations to be solved at iteration k become

$$\begin{aligned} G_{p\theta} \Delta\theta^k &= r_p(x^k) \\ G_{qv} \Delta V^k &= r_q(x^k) \end{aligned} \tag{17}$$

where

$$\begin{bmatrix} r_p \\ r_q \end{bmatrix} = r = H^T(x^k) W[z - h(x^k)] \tag{18}$$

A further approximation and decoupling on the normal equation is done in the model decoupled version. In this approach, the decoupled H is used in calculating the mismatch terms as well as the gain matrix. This leads to the total decoupling of the normal equations into the following two sets of equations:

$$\begin{aligned} G_{p\theta} \Delta\theta^k &= r_p(x^k) \cong H_{p\theta}^T W_p [z_p - h_p(x^k)] \\ G_{qv} \Delta V^k &= r_q(x^k) \cong H_{qv}^T W_q [z_q - h_q(x^k)] \end{aligned} \tag{19}$$

To keep the computations to a minimum, in this case both the gain matrices and the Jacobian matrices are calculated once at the beginning of the iterations and then they are kept constant. This last approximation leads to an approximate solution, since the mismatch terms are not calculated exactly in this case.

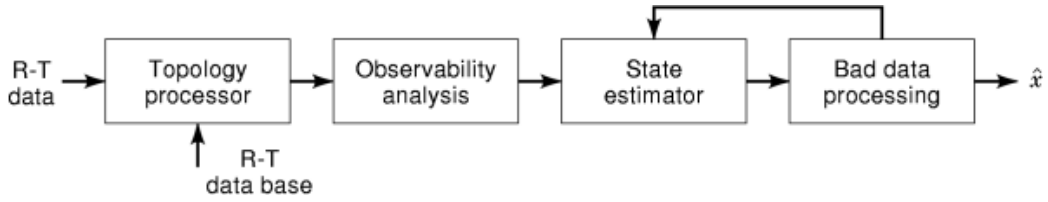

Fig. 3. Main components of a state estimation program.

In practice it has been observed, and later investigations have showed, that the performance of the decoupled estimators is sensitive to the approximations in the construction of the gain matrix (4,6). It is also shown that certain system and measurement conditions degrade the performance of these estimators significantly, as they cause the gain matrix to be ill conditioned (6). The well-known examples are the presence of a large number of injection measurements in a system, connection of a long transmission line (large impedance) with a short line (small impedance), and large disparity between the measurement weights. New numerical methods have been adopted to address these problems, and some of them will be outlined in the following sections.

Implementation Issues

There are various issues that need to be addressed for the SE to be successfully implemented in practice. First, the approach assumes that we know the exact network model [so that we can write the measurement functions $h(x)$]. To achieve this, in practice, a preprocessor program, called a topology processor, is used to mainly to determine the network connectivity and to gather the device model parameters from the database. The second preprocessor program, called the observability analysis, is used to see if the available measurements are enough for state estimation. Implementation of the WLS algorithm to solve practical cases is also an important issue; the method should be as computationally efficient as possible, and robust enough to handle various cases and provide accurate solutions. Finally, after the state estimation, a postprocessing procedure is needed to check how good the data fit the system model and then to detect the bad measurements (measurements with major errors), if necessary. This last phase is called the bad data processing. An SE implementation involving these stages is illustrated in Fig. 3.

A brief review of the various issues is provided in the following sections.

Network Topology Processor. The topology processor determines the current topology of the system (i.e., the way the devices are connected in the system) based on the circuit breaker status data obtained from SCADA. Since, as illustrated in Fig. 1, a power system mainly consists of substations interconnected by power lines, topology of the network is determined on a substation basis. As illustrated in Fig. 1, in a substation, most of the devices, such as generators, loads, shunt capacitors, transformers, and transmission lines, are connected to the bus sections. By opening/closing the CBs at a substation, different bus sections can be connected to form a common bus at a voltage level. For example, in Fig. 1, the two buses in the generation substation are separated by opening CBs on the de-energized side of the bus. The topology of a network is therefore mainly determined by the status of the CBs.

In the system database, connectivity of all the devices is defined in terms of bus sections. The network processor's task then becomes to use this description together with the status of CBs received from the SCADA to determine the network topology. Note that for the SE, and other analysis purposes, the network model is defined in terms of buses and branches, as illustrated in Fig. 2. Thus, the output of the topology processor is the data that describes a bus-branch-oriented network: each of the buses and the devices connected to these buses

(such as generators, loads, and shunt devices) and the connectivity between the buses by the transmission lines and the transformers. In addition, the topology processor identifies the network islands and discards the unenergized devices from the network description or those branches that have no power measurements associated with them.

A widely used approach to construct the network topology is based on a tree search method (3,7). The search consist of three passes. In the first pass, the bus sections of the same voltage level at the substations are processed to determine if they are connected together by closed CBs. In the second pass, all the energized network islands are identified. The search starts from a generator bus, and all the buses connected to this generator through the transmission lines and transformers are identified and numbered. When no more buses can be found to be added to the first island, a search is made for an unprocessed generator. If one found, it is used to start the search for the next island. Otherwise, all the energized islands are found. In the final pass, two tables are constructed: a bus list indicating all the devices connected to the buses; and a branch list indicating the buses the lines are connected to.

Observability Analysis. As pointed out previously, the observability analysis involves determining if the available set of measurements is enough and are evenly distributed across the network so that the state of the system can be estimated. If the measurement set is not enough, the problem in WLS estimation arises in the gain matrix of Eq. (12); it becomes singular. The system in this case is said to be unobservable. Because measurement availability as well as network topology may vary with time, it is necessary to perform an observability test each time there is a change in the available measurements or network topology. When the system becomes unobservable, it is necessary to identify the parts of the network that are unobservable (which are called unobservable islands). These islands either have to be removed from the system representation or pseudo measurements must be added at the unobservable buses to render the system observable. A recent survey of observability techniques and issues is given in Ref. 8.

The preceding introduction indicates that there are two possible ways to determine observability: by examining how much of the system the measured variables cover, or by checking the singularity of the gain matrix G (alternatively, by checking the rank of the measurement Jacobian H). These two approaches are called topological and numerical observability, respectively.

The topological observability approach was pioneered by Clements, Davis, and Krumpholz (9,10). They showed that the network can be represented by a graph, and this graph should contain at least one "observable spanning tree" for the system to be observable. A tree of a graph is observable if every one of its branches can be assigned to a measurement incident to it and each measurement is assigned only once. A spanning tree is a tree that covers all the nodes of a graph. This approach is combinatorial in nature in that it requires search of all possible trees. However, algorithms developed to minimize the search and thus reduce the computational burden of the approach (10,60).

The numerical observability approach was pioneered by Monticelli and Wu (11,12). Since the gain matrix is factored (using LDU decomposition, for example) for use in the SE algorithm, its rank can be checked during this factorization. If the gain matrix is singular (as indicated by one of the diagonal entries, called pivots, becoming zero during the factorization), then the system is unobservable. Monticelli and Wu showed that by introducing pseudoinjections at the nodes corresponding to the zero pivot locations, the unobservable buses and islands of the network can also be identified by this approach. Later, an alternative approach based on a symbolic reduction of the measurement Jacobian matrix was proposed for observability analysis (13).

Both the topological and numerical observability methods have been adapted in practice. The topological methods use mostly symbolic manipulations but are algorithmically more complex. The numerical methods are algorithmically simpler and employ numerical routines used in WLS estimation. By checking the singularity of the gain matrix, they provide a necessary and sufficient condition for observability (see the discussion of Slutsker and Monticelli in Ref. 60). They are, however, prone to numerical errors and hence may need more numerically robust factorization algorithms (8).

Note that the observability of a system depends on the amount of real-time measurements and how they are geographically distributed in the system. The distribution of the meters in the system also has a direct effect on the performance of the SE. Therefore, placing meters properly in a system for state estimation is another important issue for SE; enough meters must be placed to cover the entire system so that the system remains observable even when a few of the meters are lost occasionally. Also, the meters should be uniformly distributed in the system to retain local (bus level) redundancy as well as global redundancy. Various approaches have been proposed for this important problem (8,14).

Other Implementation Issues. There are two main challenges in successful implementation of state estimators in practice: preparation of the data and model for the SE, and performance of the SE. Some utility experience with SE is reported in literature (15,16). These experiences indicate that data and model preparation is the most critical challenge, especially in the first installation phase of the SE. Once this challenge is met, then the performance of the SE—the accuracy and reliability of SE—takes precedence.

Creating and maintaining a real-time database that contains as accurate and up-to-date a network model as possible is crutial because WLS SE assumes that the network model is accurate. Efforts have been made to determine network model parameter errors by tracking the SE solutions (17), and also topology errors (61,62).

Maintaining adequate amounts of real-time measurements is another challenge in practice. As indicated before, the minimum amount of measurements needed for state estimation is n, the number of states of the system. However, in practice the typical measurements should be on the order of 1.5 n to maintain observability of the system and reliability of the SE. To increase the number of available measurements for state estimation, there are two types of measurements that may be included in z, in addition to the telemetered measurements. Virtual measurements are the perfect measurements known in the system, such as the zero injections at buses with no load and generation. High-voltage transmission systems usually have a large amount of zero injection buses (typically 60%). Pseudomeasurements are not actual but rather guessed quantities, such as the output of generators that are not telemetered or the tab loads on a transmission line that are not telemetered. The tap values of the tap-changing transformers are also treated as pseudomeasurements if they are not telemetered or if there is not enough telemetered data for SE to estimate the tap.

The selection of weights for the SE is also an important implementation issue. If we assume that all the measurements are coming from meters operating normally, then we can assume that the measurement errors e_i in Eq. (4) are normally distributed random variables with zero mean and that they are statistically independent. The accuracy of the measurement z_i can then be related to the variance ${\sigma^2}_i$ of the random variable e_i; a large variance indicates that the measurement is not very accurate (18). If we choose the weights as the variances

$$W = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_m^2 \end{bmatrix} = R \qquad (20)$$

where R is called the *covariance matrix* of the error vector $\boldsymbol{e}$ (i.e., $R = E[\boldsymbol{e}\boldsymbol{e}^T]$). Then the WLS estimator becomes the so-called the maximum likelihood estimator, which tries to find an $\hat{x}$ that maximizes the probability density function of the measurements z (2). Thus, the weights of the telemetered measurements are usually chosen based on the accuracy of the corresponding meters. Methods have been proposed to characterize and check the accuracy of analog measurements (58,59). The weights for the virtual and pseudomeasurements are then chosen relative to these actual measurements; virtual measurements are assigned with large weights, whereas the pseudomeasurements are assigned with small weights since they are the least accurate measurements. The assignment of very large and very small weights for these measurements should, however, be avoided, since big differences between the weights may cause convergence problems, as big disparities between the weights

causes the gain matrix G to become ill conditioned (4). Alternative approaches that avoid ill-conditioning problems are described later in this article.

BAD Data Processing

Measurements that are inaccurate due to meter, telemetry, or other types of errors will deteriorate the state estimate if they are not detected, identified, and eliminated (or corrected). When using the WLS estimation method, bad data detection and identification are done after the estimation of the system state by processing the measurement residuals.

One commonly used method of detecting and identifying bad data is the *largest normalized residual* (r^N) *test* (1,6,19,39). In this test, measurement residuals are calculated and normalized by using their estimated standard deviations. This test can best be described by first reviewing some statistical properties of the measurement residuals. Consider the linearized measurement model:

$$\Delta z = z - h(x) = H\Delta x + e$$

The WLS state estimate will then be

$$\Delta \hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} \Delta z$$

and the estimated value of Δz is

$$\Delta \hat{z} = H\Delta \hat{x} = K\Delta z$$

where $K = H(H^T R^{-1} H)^{-1} H^T R^{-1}$ is called the *hat matrix.*

The measurement residuals can be expressed as follows:

$$\begin{aligned} r &= \Delta z - \Delta \hat{z} \\ &= (I - K)(H\Delta x + e) \qquad (\text{note that } KH = H) \\ &= (I - K)e \\ &= Se \end{aligned}$$

where S is called the *residual sensitivity matrix.* Each entry S_{ij} represents the sensitivity of the residual for measurement i to an error in measurement j. The residual covariance matrix Ω can then be written as

$$\Omega = E[rr^T] = S \cdot E[ee^T] \cdot S^T = SRS^T = SR$$

Hence, the normalized value of the residual for measurement i will be given by

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}} = \frac{|r_i|}{\sqrt{R_{ii} S_{ii}}}$$

Assuming that the ith measurement residual has a normal distribution with zero mean and Ω_{ii} variance, its normalized value, $r^N{}_i$, will have a standard normal distribution.

The measurement set will be suspected to include bad data if any of the normalized residuals is above the detection threshold ϵ. This threshold is chosen based on the desired level of confidence in the hypothesis that none of the measurements are in any significant error. For example, a choice of 2.58 as ϵ will correspond to a confidence level of 99.0% in the standard normal distribution table.

Largest Normalized Residual Test. It can be shown that if there is a single bad data in the measurement set, then the largest normalized residual will correspond to bad data. This also applies to multiple bad data cases as long as none of the bad measurements interact. The following are the steps of the largest normalized residual test for identification of single and noninteracting multiple bad data:

(1) Solve the WLS estimation and obtain the elements of the measurement residual vector:

$$r_i = z_i - h_i(\hat{x}) \qquad i = 1, \ldots, m$$

(2) Compute the normalized residuals:

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}} \qquad i = 1, \ldots, m$$

(3) Find k such that $r^N{}_k$ is the largest among all $r^N{}_i$, $i = 1, \ldots, m$.

(4) If $r^N{}_k > \epsilon$, then the kth measurement will be suspected as bad data. Else stop, no bad data will be suspected. Here, ϵ is a chosen identification threshold (e.g., 2.6).

(5) Eliminate the kth measurement from the measurement set and go to step 1.

Classification of Measurements and Their Properties. Measurements can be broadly classified as *critical* and *redundant* (or *noncritical*) with the following properties:

- A critical measurement is the one whose elimination from the measurement set will result in an unobservable system.
- The row/column of S corresponding to a critical measurement will be zero.
- The residuals of critical measurements will always be zero, and therefore errors in critical measurements cannot be detected.
- Only noncritical measurements will have nonzero measurement residuals, allowing detection and identification of their errors.

Types of Bad Data and Their Identification. When there is single bad data, the largest normalized residual will correspond to the bad measurement, provided that the measurement is not critical. Multiple bad data may appear in three ways (20):

- Noninteracting: If $S_{ik} \approx 0$, then measurements i and k are said to be noninteracting. In this case, even if bad data appear simultaneously in both measurements, the largest normalized residual test can identify them sequentially, one pass at a time.
- Interacting, nonconforming: If S_{ik} is significantly large, then measurements i and k are said to be interacting. However, if the errors in measurement i and k are not consistent with each other, then the largest normalized residual test may still indicate the bad data correctly.
- Interacting, conforming: If two interacting measurements have errors that are in agreement (that is, they satisfy power flow equations), then the largest normalized residual test may fail to identify either one.

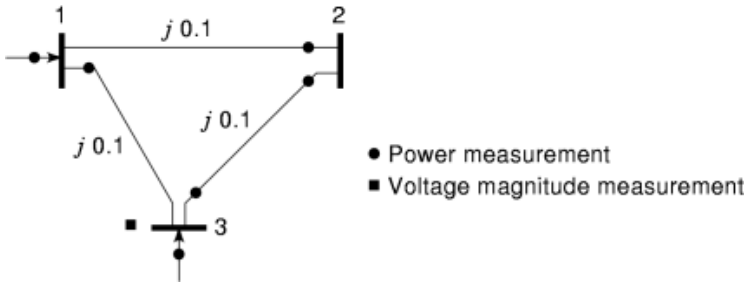

Fig. 4. A small system with measurements to illustrate the bad data processing capabilities of *WLS SE*.

Table 1. Bad Data Simulation Results on the Small System

	Bad Data					
			Multiple Interacting			
	Single		Nonconforming		Conforming	
Meas. Type	z_i	r_i^N	z_i	r_i^N	z_i	r_i^N
Flow 1-3	0	7.1	0	17.6	0	31.7
Flow 2-1	0	18.7	0	11.2	0	26.2
Flow 3-2	1	**88.0**	−1	**120.9**	1	55.0
Flow 2-3	0	25.6	0	7.3	0	**58.6**
Inj. 1	0	15.5	0	41.5	0	10.4
Inj. 3	0	41.7	1	111.3	1	27.9

A small test system given in Fig. 4 will be used to illustrate the performance of the largest normalized residual test for the aforementioned types of bad data. In Fig. 4, all branches have identical reactances, $x_i = 0.1$, and all measurements have the same error variance, $\sigma_i = 0.01$. Note that all measurements are assumed to be zero except for the bad data. A dc model is used, and therefore the true solution should yield zero phase angles at all buses and zero flows along all branches. Single, multiple interacting nonconforming and multiple interacting conforming bad data are introduced into the measurements, and the results are given in the following table. The largest normalized residual for each case appears in boldface in Table 1. As the results in the table indicate, the main difficulty in the bad data processing scheme arises due to the smearing effect of gross errors on the residuals.

Other Bad Data Identification Methods. Failure of the largest normalized residual technique to identify interacting conforming bad data motivated the development of alternative methods that attempt to overcome this deficiency. Hypothesis testing identification (*HTI*) is one such technique suggested by Mili, Van Cutsem, and Ribbens-Pavella (21,22). In this technique, a suspected set of measurements are initially set aside based on their normalized residuals. Then a hypothesis testing is done based on the estimated errors of the chosen suspect set of measurements. This technique works well as long as all bad measurements are included in the suspect set. Another approach is introduced by Monticelli and Wu (23). This approach uses the framework of decision theory to formulate the bad data identification problem as a combinatorial optimization problem, where there are 2^m possibilities for m measurements, each one having a possibility of being bad or good. An efficient branch and bound method is developed to search the decision tree leading to the optimal solution. A geometric approach to detecting and identifying multiple bad data is presented by Clements and Davis (24). A procedure for selecting the suspect set of measurements and carrying out the identification tests

based on the projection of the residual vector on the subspace determined by these suspect measurements is introduced.

Numerically Stable Estimators

The standard weighted least squares formulation of the state estimation problem yields the so-called normal equations as given in Eq. (11). The gain matrix $G(x)$ shown in Eq. (12) is known to be numerically ill conditioned. Ill conditioning may be due to the following reasons:

- Existence of very long and very short lines terminating at the same bus
- Assigning very high weights to some measurements in order to enforce equality constraints
- Existence of a large number of injection measurements

The ill-conditioning problem has been addressed by two categories of methods. One is based on orthogonalization of the matrix H. The other category includes methods that formulate the state estimation problem with equality constraints. These are briefly described next.

Orthogonalization Methods. The WLS estimator minimizes the following objective function at each iteration:

$$\begin{aligned} J &= (\Delta z - H\Delta x)^T W(\Delta z - H\Delta x) \\ &= (\Delta z' - H'\Delta x)^T(\Delta z' - H'\Delta x) \end{aligned}$$

where $H' = W^{1/2}\,H$ and $\Delta z' = W^{1/2}\,\Delta z$.

Choosing an orthonormal matrix Q (i.e., $Q^TQ = I$) such that $QH' = U$ (U being an upper triangular matrix), the objective function can be rewritten as (25,26)

$$J = (\Delta y - U\Delta x)^T(\Delta y - U\Delta x) \qquad (21)$$

where $\Delta y = Q\,\Delta'_z$. Let

$$U = \begin{bmatrix} R \\ 0 \end{bmatrix}$$

where R is a square upper triangular matrix and

$$\Delta y = \begin{bmatrix} \Delta y_r \\ \Delta y_0 \end{bmatrix}$$

Then the minimum of J will be found by solving

$$R\Delta x = \Delta y_r \qquad (22)$$

and the minimum value of J will be given by $(\Delta y_0)^T(\Delta y_0)$. Q is usually much denser compared to H and found using the Given transformations (27). An efficient implementation of this method is developed in Ref. 28.

A hybrid version of the preceding method was introduced in Ref. 29, where storage of Q is avoided. It is based on the observation that the upper triangular factor R in Eq. (22) is the same as the upper triangular factor of the original gain matrix G. Hence, after the orthogonal factorization of H, Δx can be solved using the following equation:

$$R^T R \Delta x = H^T W \Delta z \tag{23}$$

Normal Equations with Constraints. As stated earlier, use of very large weights for certain measurements is one of the reasons for ill conditioning of the normal equations. Zero injections are examples of such measurements. Instead of trying to enforce these measurements by introducing large weights, exact enforcement of equality constraints can be achieved by using the method of Lagrange multipliers (30). The WLS estimation problem to be solved at each iteration can be rewritten as

$$\text{Minimize } (\Delta z - H\Delta x)^T W(\Delta z - H\Delta x) \tag{24}$$

$$\text{Subject to } C\Delta x = -c(x) \tag{25}$$

where H and C are the Jacobians for unconstrained measurements and equality constraints respectively, $\Delta z = z - h(x)$, and $-c(x)$ is the equality constraint vector evaluated at the current iteration. The solution of the preceding constrained optimization problem will be given by solving the following linear equation (30):

$$\begin{bmatrix} H^T W H & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \lambda \end{bmatrix} = \begin{bmatrix} H^T W \Delta z \\ -c(x) \end{bmatrix} \tag{26}$$

where λ is the Lagrange multipliers vector and x is the value of the state at the last iteration.

The preceding equation can be rewritten in sparser form by using $\mu = W(\Delta z - H\Delta x)$ as an additional unknown:

$$\begin{bmatrix} 0 & 0 & C \\ 0 & W^{-1} & H \\ C^T & H^T & 0 \end{bmatrix} \begin{bmatrix} -\lambda \\ \mu \\ \Delta x \end{bmatrix} = \begin{bmatrix} -c(x) \\ \Delta z \\ 0 \end{bmatrix} \tag{27}$$

This method is known as Hachtel's augmented matrix method (31). Even though the number of unknowns is increased by the introduction of μ, the overall matrix remains sparse and the explicit formation of $H^T WH$ is avoided. This improves the condition number of the matrix, yielding a numerically more robust system of equations.

One of the drawbacks of this method is the fact that the coefficient matrix in Eq. (7) is symmetric but indefinite. Hence, an ordering based strictly on sparsity considerations may fail. Block pivoting with blocks of 1×1 or 2×2 may have to be used, ensuring numerical nonsingularity of pivot blocks during factorization of this indefinite matrix (32,33,34).

Alternative State Estimation Methods

All bad data processing methods discussed so far share a common shortcoming: the fact that the detection decisions are based on the measurement residuals. In the case of multiple bad data, masking the effect of these bad measurements on their residuals may prohibit identification of bad data using residual based methods. Starting with Schweppe (35), several methods have been proposed to address this shortcoming and improve the robustness of state estimators. Robustness of an estimator refers to its insensitivity to such errors in the measurements when solving for the state estimate (36,37).

State estimators that downweight measurements having large residuals are thus proposed in Refs. 38,39, 40, and 57. More recently, state estimators that minimize the sum of the absolute values of the measurement residuals have been introduced. These estimators are referred to as the least absolute value (*LAV*) estimators.

Least Absolute Value State Estimation. Least absolute value estimation is an alternative method that can be used to estimate the state of the system. Use of the LAV state estimation in power systems was initially proposed in Refs. 41 and 42, motivated by the automatic bad data rejection property of these estimators.

The formulation of the LAV estimation problem is based on minimizing the sum of the absolute values of the measurement residuals:

$$\text{Minimize } c^T|r| \tag{28}$$

$$\text{Subject to } z - Hx = r \tag{29}$$

where H is the $m \times n$ measurement Jacobian, $c \in R^m$ is the vector of measurement weights (which may all be set equal to 1 for equal weighting of measurements) and $r \in R^m$ is the vector of measurement residuals.

Solution of the optimization problem given by Eqs. (28) and (29) interpolates n out of m available measurements, where n and m are the number of state variables and measurements, respectively. Those measurements that are not exactly satisfied (having nonzero residuals) by the LAV estimator are referred to as rejected measurements. In the absence of leverage measurements, if there are any bad data in the measurement set, they will be rejected automatically by the LAV estimator. This characteristic is attractive in dealing with bad data. Leverage measurements are discussed within the context of robust estimators in the next subsection.

Alternative solution algorithms for solving the LAV estimation problem exist. These can be broadly classified as follows:

(1) Simplex-based algorithms
(2) Interior-point-based algorithms

Simplex-Based Algorithms. Several variations of the well-known simplex method of solving linear programming (*LP*) problems can be applied to the LAV estimation problem. In applying the simplex method, the special structure of the LAV estimation problem can be exploited for numerical efficiency (43,44,45). This can be accomplished both at the initialization and the actual optimization stages of the algorithm.

Interior-Point-Based Algorithms. A new algorithm for solving LP problems was introduced by N. K. Karmarkar in 1984 (46). Several variants of the original Karmarkar's algorithm have been introduced since then. The collection of these methods constitutes what is referred to as the interior point methods for linear programming. The distinguishing feature of these algorithms, as compared to the simplex algorithms, is the way they reach the solution. While simplex algorithms trace the extreme points of the feasible region along its exterior, interior point methods trace a path interior to the feasible region. There are three variants of the interior point algorithm (47):

(1) Affine-scaling primal algorithm
(2) Affine-scaling dual algorithm
(3) Primal-dual algorithm

Interior point methods typically perform better than their simplex-based counterparts when the problem size increases. Details of the three algorithms can be found in Ref. 47, and their applications to power system state estimation are discussed in Refs.48 and 49.

Leverage Measurements and Other Robust Estimators. Leverage measurements are those measurements that have a very strong influence on the estimated states (50,57). Leverage measurements in power systems can be identified based on the measurement Jacobian H, whose rows correspond to individual measurements. It has been observed that the following conditions usually lead to creation of leverage measurements (51):

- Having flow measurements along relatively short lines
- Having injection measurements at either terminal of a short line
- Having injection measurements at buses with several incident lines

The main motivation behind the application of LAV estimation methods to the power system state estimation problem has been their expected robustness against bad data in the measurement set. However, LAV estimators were shown to be susceptible to leverage points in the measurements (51,52,53,54). A new estimator that minimizes the least median of squares of the measurement residuals is proposed in Ref. 55 and applied to the power system state estimation problem in Refs. 51 and 56. More recently, a method of identifying leverage measurements based on a measure called projection statistics was introduced, and a robust iteratively reweighted least squares estimation algorithm was presented in Ref. 57.

New State Estimation Applications

Recently, the use of global positioning systems made it possible to obtain accurate synchronized measurements over a transmission system. This technology made possible the use of phasor measurements in state estimation (63), and even in harmonic state estimation (64,65). At the power distribution level, efforts to introduce automation at the distribution level led the extension of state estimation to distribution systems. Methods have been proposed (66,67) to address, and make use of, the special features of distribution systems such as system unbalance, radial topology, lack of enough real-time measurements, availability of mostly current measurements, and so on.

BIBLIOGRAPHY

1. F. C. Schewppe E. J. Handschin Static state estimation in electric power systems, *Proc. IEEE*, **62**: 972–982, 1974.
2. A. J. Wood B. F. Wollenberg *Power Generation, Operation, and Control*, New York: Wiley, 1984.
3. A. Bose K. A. Clements Real-time modeling of power networks, *Proc. IEEE*, 75: 1607–1622, 1987.
4. F. F. Wu Power system state estimation: A survey, *Int. J. Elect. Power Energy Syst.*, **12** (2): 80–87, 1990.
5. A. Garcia A. Monticelli P. Abreu Fast decoupled state estimation and bad data processing, *IEEE Trans. Power Appar. Syst.*, **PAS-98**: 1645–1652, 1979.
6. A. Monticelli A. Garcia Fast decoupled state estimators, *IEEE Trans. Power Syst.*, **5**: 556–563, 1990.
7. T. E. DyLiacco K. A. Ramarao A. W. Weiner Network status analysis for real-time systems, *Proc. 8th PICA Conf.*, 1973, pp. 356–362.

8. K. A. Clements Observability methods and optimal meter placement, *Int. J. Electr. Power Energy Syst.*, **12** (2): 88–93, 1990.
9. G. R. Krumpholz K. A. Clements P. W. Davis Power system observability: A practical algorithm using network topology, *IEEE Trans. Power Appar. Syst.*, pp. 1534–1542, 1980.
10. K. A. Clements G. R. Krumpholz P. W. Davis Power system state estimation with measurement deficiency: An algorithm that determines the maximal observable subnetwork, *IEEE Trans. Power Appar. Syst.*, **PAS-101**: 3044–3052, 1982.
11. A. Monticelli F. F. Wu Network observability: theory, *IEEE Trans. Power Appar. Syst.*, **PAS-104**: 1042–1048, 1985.
12. A. Monticelli F. F. Wu Network observability: Identification of observable islands and measurement placement, *IEEE Trans. Power Appar. Syst.*, **PAS-104**: 1035–1041, 1985.
13. I. W. Slutsker J. M. Scudder Network observability analysis through measurement jacobian matrix reduction, *IEEE Trans. Power Syst.*, **PWRS-2**: 331–338, 1987.
14. M. E. Baran *et al.* A meter placement method for state estimation, *IEEE Trans. Power Syst.*, **10**: 1704–1710, 1995.
15. T. E. DyLiacco Real-time implementation and role of state estimation with the EMS, *J. Electr. Power Energy Syst.*, pp. 75–90, 1990.
16. M. Assadian *et al.* Field operational experiences with on line state estimator, *IEEE PES 1993 Winter Meet.*, 1993, paper 190-9.
17. E. W.-H. Liu F. F. Wu S.-M. Lun Estimation of parameter errors for measurement residuals in state estimation, *IEEE Trans. Power Syst.*, **7**: 81–89, 1992.
18. J. J. Allemong L. Radu A. M. Sasson A fast and reliable state estimator algorithm for AEP's new control center, *IEEE Trans. Appar. Power Syst.*, **PAS-101**: 933–944, 1982.
19. F. Broussole State estimation in power systems: Detecting bad data through the sparse inverse matrix method, *IEEE Trans. Power Appar. Syst.*, **PAS-97**: 678–682, 1978.
20. A. Monticelli F. F. Wu *Decision Theory Approach to Identifying Bad Data in State Estimation for Power Systems*, EPRI Rep., Palo Alto, CA: EPRI, 1986, Res. Proj. 1999-6, EL-4325.
21. L. Mili Th. Van Cutsem, M. Ribbens-Pavella Hypothesis testing identification: A new method for bad data analysis in power system state estimation, *IEEE Trans. Power Appar. Syst.*, **PAS-103**: 3239–3252, 1984.
22. L. Mili Th. Van Cutsem Implementation of HTI method in power system state estimation, *IEEE Trans. Power Syst.*, **3**: 887–893, 1988.
23. A. Monticelli F. F. Wu M. Yen Multiple bad data identification for state estimation by combinatorial optimization, *IEEE Trans. Power Deliv.*, **PWRD-1**: 361–369, 1986.
24. K. A. Clements P. W. Davis Multiple bad data detectability and identifiability: A geometric approach, *IEEE Trans. Power Deliv.*, **PWRD-1**: 355–360, 1986.
25. A. Simoes-Costa V. H. Quintana A robust numerical technique for power system state estimation, *IEEE Trans. Power Appar. Syst.*, **100**: 691–698, 1981.
26. A. Simoes-Costa V. H. Quintana An orthogonal row processing algorithm for power system sequential state estimation, *IEEE Trans. Power Appar. Syst.*, **100**: 3791–3800, 1981.
27. A. George M. T. Heath Solution of sparse linear least squares problems using givens rotations, *Linear Algebra Its Appl.*, **34**: 69–83, 1980.
28. N. Vempati I. Slutsker W. F. Tinney Enhancements to givens rotations for power system state estimation, *IEEE Trans. Power Syst.*, **6**: 842–849, 1991.
29. A. Monticelli C. A. F. Murari F.F. Wu A hybrid state estimator: Solving normal equations by orthogonal transformations, *IEEE Trans. Power Appar. Syst.*, **105**, 3460–3468, 1985.
30. F. C. Aschmoneit N. M. Peterson E. C. Adrian State Estimation with Equality Constraints, *Proc. 10th PICA Conf.*, Toronto, 1977, pp. 427–430.
31. A. Gjelsvik S. Aam L. Holten Hachtel's augmented matrix method: A rapid method improving numerical stability in power system static state estimation, *IEEE Trans. Power Appar. Syst.*, **PAS-104**: 2987–2993, 1985.
32. P. A. Machado G. P. Azavedo A. J. Monticelli A mixed pivoting approach to the factorization of indefinite matrices in power system state estimation, *IEEE Trans. Power Syst.*, **6**: 676–682, 1991.
33. F. L. Alvarado W. F. Tinney State estimation using augmented blocked matrices, *IEEE Trans. Power Syst.*, **5**: 911–921, 1990.
34. R. R. Nucera M. L. Gilles A blocked sparse matrix formulation for the solution of equality constrained state estimation, *IEEE Trans. Power Syst.*, **6**: 214–224, 1991.

35. F. C. Schweppe J. Wildes D. B. Rom Power system static state estimation, parts I–III, *IEEE Trans. Power Appar. Syst.*, **PAS-89**: 120–135, 1970.
36. F. R. Hampel *et al. Robust Statistics: The Approach Based on Influence Functions*, New York: Wiley, 1986.
37. P. J. Rousseeuw A. M. Leroy *Robust Regression and Outlier Detection*, New York: Wiley, 1987.
38. H. M. Merrill F. C. Schweppe Bad data suppression in power system state estimation, *IEEE Trans. Power Appar. Syst.*, **PAS-90**: 2718–2725, 1971.
39. E. Handschin *et al.* Bad data analysis in power system estimation, *IEEE Trans. Power Appar. Syst.*, **PAS-94**: 329–337, 1975.
40. U.S. Department of Energy, *Contribution to Power System State Estimation and Transient Stability Analysis*, Report, Washington, DC: USDOE, 1984, DOE/ET/29362-1.
41. M. R. Irving R. C. Owen M. J. H. Sterling Power System State Estimation Using Linear Programming, *Proc. Inst. Electr. Eng.*, 125: 879–885, 1978.
42. W. W. Kotiuga M. Vidyasagar Bad data rejection properties of weighted least absolute value techniques applied to static state estimation, *IEEE Trans. Power Appar. Syst.*, **PAS-101**: 844–851, 1982.
43. A. Abur M. K. Celik A fast algorithm for the weighted least absolute value state estimation, *IEEE Trans. Power Syst.*, **6**: 1–8, 1991.
44. K. A. Clements P. W. Davis K. D. Frey An efficient algorithm for computing the weighted least absolute value estimate in power system static state estimation, *Proc. IFAC Int. Symp. Power Syst. Power Plant Control*, Seoul, Korea, pp. 785–790, 1989.
45. I. Barrodale F. D. K. Roberts An improved algorithm for discrete $\mathscr{l}_1$ linear approximation, *SIAM J. Numer. Anal.*, **10** (5): 839–848, 1973.
46. N. K. Karmarkar A new polynomial time algorithm for linear programming, *Combinatorica*, **4**: 373–395, 1984.
47. A. Arbel *Exploring Interior-Point Linear Programming Algorithms and Software*, Cambridge, MA: MIT Press, 1993.
48. K. A. Clements P. W. Davis K. D. Frey An interior point algorithm for weighted least absolute value power system state estimation, *IEEE PES Winter Meet.*, New York, Pap. No. 235-2-PWRS, 1991.
49. H. Singh F. L. Alvarado Weighted least absolute value state estimation using interior point methods, *IEEE Trans. Power Syst.*, **9**: 1478–1484, 1994.
50. S. P. Ellis S. Morgenthaler Leverage and breakdown in L_1 regression, *J. Amer. Stat. Assoc.*, **87** (417): 143–148, 1992.
51. L. Mili V. Phaniraj P. J. Rousseeuw Least median of squares estimation in power systems, *IEEE Trans. Power Syst.*, **6**: 511–523, 1991.
52. A. Abur A bad data identification method for linear programming state estimation, *IEEE Trans. Power Syst.*, **5**: 894–900, 1990.
53. D. M. Falco S. M. de Assis Linear programming state estimation: Error analysis and gross error identification, *IEEE Trans. Power Syst.*, **3**: 809–815, 1988.
54. L. Mili V. Phaniraj P. J. Rousseeuw Robust estimation theory for bad data diagnostics in electric power systems, *Adv. Control Dyn. Syst.*, *36*: 1990.
55. P. J. Rousseeuw Least median of squares regression, *J. Amer. Stat. Assoc.*, **79**: 871–880, 1984.
56. M. G. Cheniae L. Mili P. J. Rousseeuw Identification of multiple interacting bad data via power system decomposition, *IEEE Trans. Power Syst.*, **11**: 1555–1563, 1996.
57. L. Mili *et al.* Robust state estimation based on projection statistics, *IEEE Trans. Power Syst.*, **11**: 1118–1127, 1996.
58. M. M. Adibi D. K. Thorne Remote measurement calibration, *IEEE Trans. Power Syst.*, **PWRS-1**: 194–203, 1986.
59. M. M. Adibi J. P. Stovall On estimation of uncertainties in analog measurements, *IEEE Trans. Power Syst.*, **5**: 1222–1230, 1990.
60. R. R. Nucera M. L. Gilles Observability analysis: A new topological algorithm, *IEEE Trans. Power Syst.*, **6**: 466–475, 1991.
61. K. A. Clements A. S. Costa Topology error identification using normalized Lagrange multipliers, *IEEE Trans. Power Syst.*, **13**: 347–353, 1998.
62. A. Abur H. Kim M. K. Celik Identifying the unknown circuit breaker statuses in power networks, *IEEE Trans. Power Syst.*, **10**: 2029–2037, 1995.
63. A. G. Phadke J. S. Thorp K. J. Karimi State estimation with phasor measurements, *IEEE Trans. Power Syst.*, **RWRS-1**: 233–241, 1986.

64. A. P. Sakis Meliopoulos F. Zhang S. Zelingher Power system harmonic state estimation, *IEEE Trans. Power Delivery*, **9**: 1701–1709, 1994.
65. B. Gou A. Abur A tracking state estimator for nonsinusoidal periodic steady-state operation, *IEEE PES Winter Meet.*, Tampa, FL, Feb. 1998, paper no: PE-368-RWRD, to appear in IEEE PES Transactions.
66. M. E. Baran A. W. Kelley A branch current based state estimation method for distribution systems, *IEEE Trans. Power Syst.*, **10**: 483–491, 1995.
67. C. N. Lu Meliopoulos J. H. Teng W-H E. Liu Distribution system state estimation, *IEEE Trans. Power Syst.*, **10**: 229–240, 1995.

MESUT E. BARAN
North Carolina State University
ALI ABUR
Texas A&M University