Electric Power System State Estimation

A. MONTICELLI, FELLOW, IEEE

Invited Paper

This paper discusses the state of the art in electric power system state estimation. Within energy management systems, state estimation is a key function for building a network real-time model. A real-time model is a quasi-static mathematical representation of the current conditions in an interconnected power network. This model is extracted at intervals from snapshots of real-time measurements (both analog and status). The new modeling needs associated with the introduction of new control devices and the changes induced by emerging energy markets are making state estimation and its related functions more important than ever.

Keywords*—*Bad data analysis, generalized state estimation, network topology processing, observability analysis, parameter estimation, power systems, state estimation, topology estimation.

I. INTRODUCTION

Vertically integrated utilities provide bundled services to customers aiming at high reliability with the lowest cost. In the traditional environment utilities perform both power network and marketing functions. Although energy management systems (EMS) technology has been used to a certain extent, utilities were not pressed to utilize tools that demanded accurate real-time network models such as optimal power flows and available transfer capability determination [1]. This is bound to change in the emerging competitive environment.

In the new environment, the pattern of power flows in the network is less predictable than it is in the vertically integrated systems, in view of the new possibilities associated with open access and the operation of the transmission network under energy market rules. Although reliability remains a central issue, the need for the real-time network models becomes more important than before due to new energy market related functions which will have to be added to new and existing EMS. These models are based on the results yielded by state estimation and are used in network applications such as optimal power flow, available transfer capability, voltage, and transient stability. The new role of state estimation and other advanced analytical functions in competitive energy markets was discussed by Shirmohammadi *et al.* [2]. Hence, the implementation of real-time network analysis functions is crucial for the proper independent system operation. Based on these network models, operators will be able to justify technical and economical decisions, such as congestion management and the procurement for adequate ancillary services, and to uncover potential operational problems related to voltage and transient stability [1].

Reviews of the state of the art in state estimation algorithms were presented in [4]–[6]. Comparative studies of numerically robust estimators for power networks can be found in [7]. A review of the state of the art in bad data analysis was provided in [8]. A comprehensive bibliography on state estimation up to 1989 can be found in [9]. Generalized state estimation, which includes the estimation of states, parameters, and topology, is discussed in [10] and [11]. A review of external system modeling was presented in [12], and more recently, the state of the art on this subject was reviewed by the IEEE Task Force on External Network Modeling [13].

II. NETWORK REAL-TIME MODELING

Network real-time models are built from a combination of snapshots of real-time measurements and static network data. Real time measurements consist of analog measurements and statuses of switching devices, whereas static network data correspond to the parameters and basic substation configurations. The real-time model is a mathematical representation of the current conditions in a power network extracted at intervals from state estimation results. Ideally, state estimation should run at the scanning rate (say, at every two seconds). Due to computational limitations, however, most practical estimators run every few minutes or when a major change occurs.

A. Conventional State Estimation

In conventional state estimation, network real-time modeling is decomposed into: 1) the processing of logical data (the statuses of switching devices) and 2) the processing of analog data (e.g., power flow, power injection, and voltage magnitude measurements). During topology processing, the statuses of breakers/switches are processed using a bus-sec-

Manuscript received April 9, 1999; revised August 31, 1999.
The author is with UNICAMP, Campinas 13081-970, São Paulo, Brazil.
Publisher Item Identifier S 0018-9219(00)00840-9.

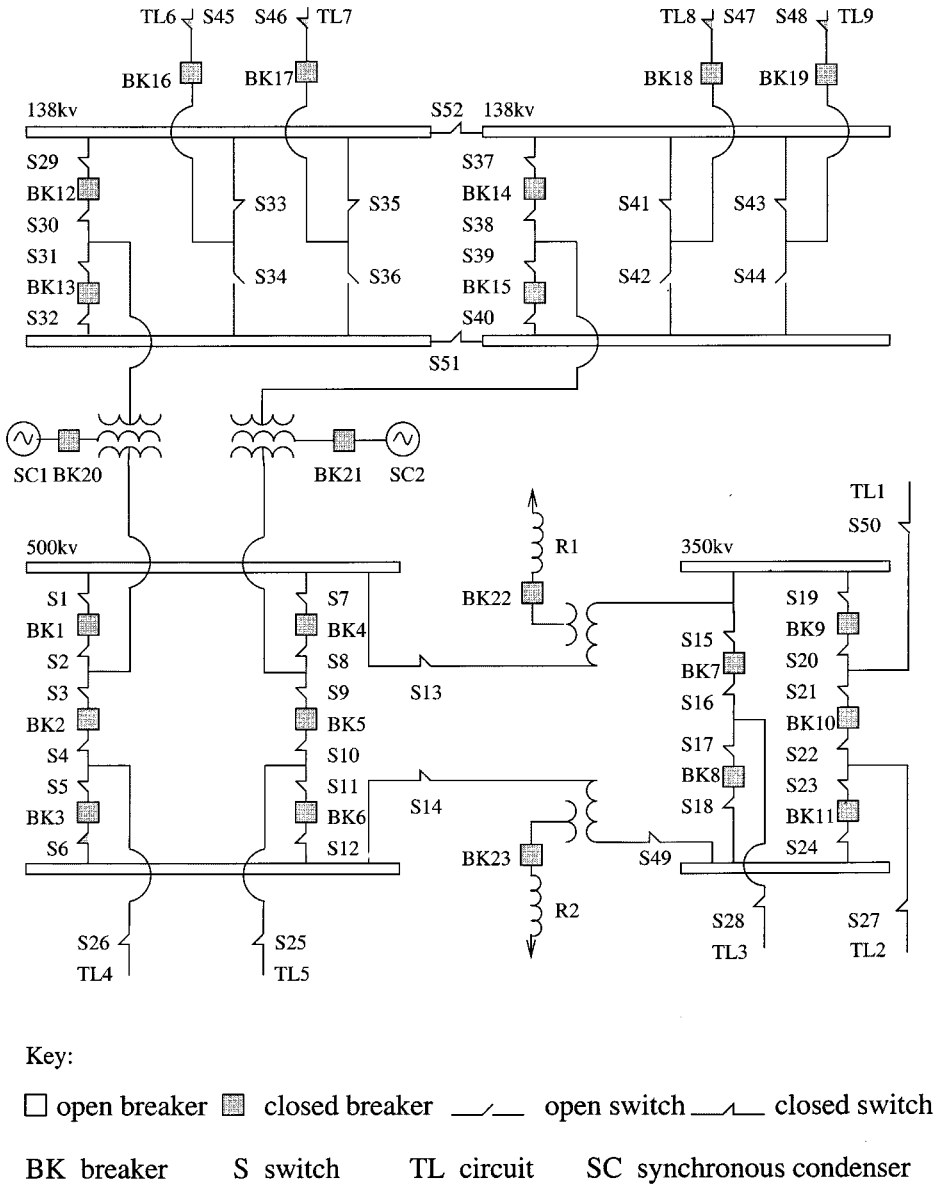


Fig. 1. A substation modeled at the physical level.

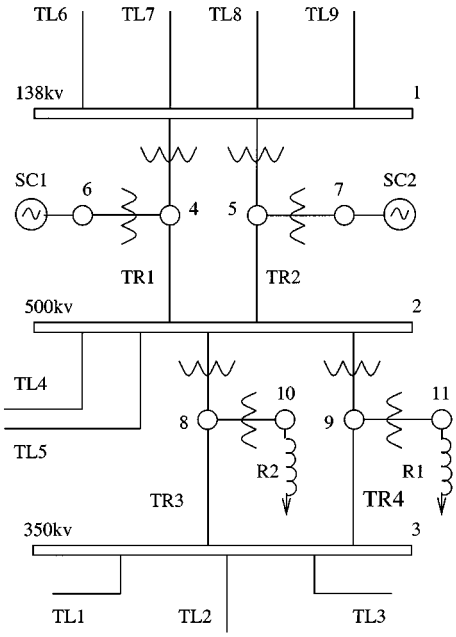
tion/switching-device network model of the type illustrated in Fig. 1. During observability analysis and state estimation, the network topology and parameters are considered as given, and analog data are processed using the bus/branch network model (Fig. 2). In the conventional approach, logical data are checked by the topology processor and the analog data are checked by the state estimator.

B. System of Interest

Fig. 3 shows a part of an interconnected network delineating both observable and unobservable areas of the system of interest as viewed from the EMS. This conceptual visualization can be used both for vertically integrated utilities, as well as for pools and independent system operators. Ideally, the control area for which a specific control center is responsible is observable although this is not always the case, since parts of it can be permanently or temporarily unobservable (e.g., lower voltage subnetworks). On the other hand, parts of the network outside the control area, which are normally unobservable, can be made observable by direct metering or data exchange.

Observable islands are handled with full state estimation including bad data analysis. State estimation can be extended to the rest of the system of interest through the addition of pseudomeasurements based on load prediction and generation scheduling. In executing state estimation for this augmented system, however, care must be taken to avoid corrupting the states estimated from telemetry data.

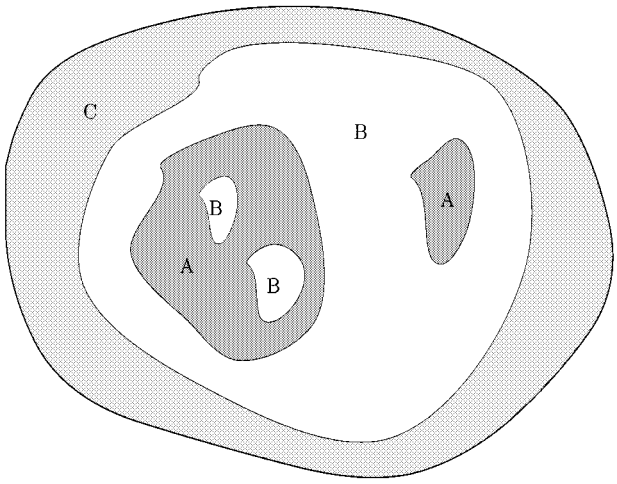
Hence, the state estimator is used to build the model for the observable part of the network and optionally to attach a model of the unobservable part to it. With adequate redundancy level, state estimation can eliminate the effect of bad data and allow the temporary loss of measurements without significantly affecting the quality of the estimated values. State estimation is mainly used to filter redundant data, to



Key:

TL transmission line	SC synchronous condenser
TR transformer	R reactor

Fig. 2. Bus/branch model of the substation in Fig. 1.



A Observable part of the system of interest

B Unobservable part of the system of interest

C Rest of the interconnected system

Fig. 3. Observability characterization of an interconnected network.

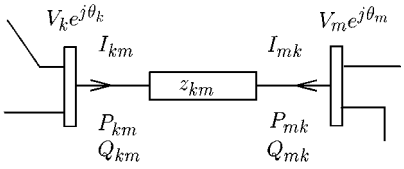


Fig. 4. Branch with unknown impedance z_{km}.

eliminate incorrect measurements, and to produce reliable state estimates, although, to a certain extent, it allows the determination of the power flows in parts of the network that are not directly metered. Not only do traditional applications such as contingency analysis, optimal power flow, and dispatcher training simulator rely on the quality of the real-time network model obtained via state estimation, but the new functions needed by the emerging energy markets will do so even more [2].

C. Generalized State Estimation

In generalized state estimation there is no clear-cut distinction between the processing of logical and analog data since network topology processing may include local, substation level, state estimation, whereas when state estimation is performed for the whole network, parts of it can be modeled at the physical level (bus-section/switching-device model). The term generalized is used to emphasize the fact that not only states, but also parts of the network topology, or even parameters, are estimated.

The explicit modeling of switches facilitates bad data analysis when topology errors are involved (incorrect status of switching devices). In this case, state estimation is performed on a model in which parts of the network can be represented at the physical level. This allows the inclusion of measurements made on zero impedance branches and switching devices. The conventional states of bus voltages and angles are augmented with new state variables. Observability analysis is extended to voltages at bus sections and flows in switching devices, and if their values can be computed from the available measurements they are considered to be observable.

For a zero impedance branch, or a closed switch (illustrated in Fig. 4 with $z_{km} = 0$), the following constraints, or pseudomeasurements, are included in state estimation [14]:

$$V_k - V_m = 0 \quad \text{and} \quad \theta_{km} = \theta_k - \theta_m = 0.$$

In this case, P_{km} and Q_{km} are used as additional state variables. These variables are independent of the complex nodal voltages $V_k e^{j\theta_k}$ and $V_m e^{j\theta_m}$, since Ohm's law (in complex form) cannot be used to compute the branch current as a function of these voltages.

For open switches ($z_{km} = \infty$ in Fig. 4), the same additional state variables are include in state estimation. In the case of open switches the pseudomeasurements are as follows:

$$P_{km} = 0 \quad \text{and} \quad Q_{km} = 0.$$

No pseudomeasurements are added in the case of switches with unknown status. (There are situations in which the wrong status of a switching device can affect state estimation convergence. In these cases it may be preferable to treat such status as unknown and proceed with state estimation, which hopefully will include the estimation of the correct status.)

The ideas above can be extended to branches with unknown impedances [11]. (The same comment regarding the impact of unknown status on state estimation convergence applies to branches with impedances with large errors.) Consider the diagram shown in Fig. 4, where the branch

impedance z_{km} is unknown, whereas for simplicity all branches incident to k and m are assumed to have known impedances. As with zero impedance branches and with closed/open breakers, Ohm's law cannot be used to relate the state variables $V_k e^{j\theta_k}$ and $V_m e^{j\theta_m}$, associated with the terminal nodes k and m, with the branch complex power flows $P_{km} + jQ_{km}$ and $P_{mk} + jQ_{mk}$. These power flows can be used as additional states, although they are not independent, since they are linked by the constraint $I_{km} + I_{mk} = 0$, which can be expressed by the two following pseudomeasurements:

$$P_{km}V_m + (P_{mk}\cos\theta_{km} - Q_{mk}\sin\theta_{km})V_k = 0$$
$$Q_{km}V_m + (P_{mk}\sin\theta_{km} + Q_{mk}\cos\theta_{km})V_k = 0.$$

A power injection measurement at node k can be expressed as the summation of the flow state variables $P_{km}+jQ_{km}$ and the flows in all other branches incident to k. Since only the flows in regular branches are functions of the nodal state variables, the unknown impedance will not form part of the measurement model. A similar analysis holds for power injection measurement at node m and power flow measurements made in the unknown impedance branch. Once the network state is estimated, the value of the unknown parameter can be computed from the estimates.

More complex network elements such as a transmission line π equivalent model requires the consideration of additional constraints (pseudomeasurements) in addition to the inclusion of flow state variables. Consider, for example, the equivalent π model in Fig. 5, where the series branch impedance is to be estimated. In this case, power flows P'_{km}, Q'_{km}, P'_{mk}, and Q'_{mk} are considered to be additional states. The terminal power flows P_{km}, Q_{km}, P_{mk}, and Q_{mk} are then expressed in terms of the new state variables rather than as a function of the terminal bus voltages; as a consequence, the series branch impedance will not appear in the measurement model, and this can be written as

$$P_{km} = P_{kk} + P'_{km} \quad \text{and} \quad Q_{km} = Q_{kk} + Q'_{km}$$
$$P_{mk} = P_{mm} + P'_{mk} \quad \text{and} \quad Q_{mk} = Q_{mm} + Q'_{mk}.$$

Notice that in the above equations, the power flows P_{kk}, Q_{kk}, P_{mm}, and Q_{mm} are written in terms of the shunt parameters, as usual. The bus injection measurements at buses k and m are expressed in terms of the terminal power flows as described above.

These added states are not entirely independent, so it is necessary to include the following relationship in the model:

$$P'_{km}V_m + (P'_{mk}\cos\theta_{km} - Q'_{mk}\sin\theta_{km})V_k = 0$$
$$Q'_{km}V_m + \left(P'_{mk}\sin\theta_{km} + Q'_{mk}\cos\theta_{km}\right)V_k = 0.$$

Now consider the situation represented in Fig. 6, which shows a π equivalent model in which the shunt elements are made dormant. The state variables and measurement model are the same as those in Fig. 5. The constraints linking the state variables in the case of a balanced π model is $y^{\mathrm{sh}}_{kk} = y^{\mathrm{sh}}_{mm}$. Expressing the shunt admittances in terms of the corresponding active and reactive power flows yields the two following pseudomeasurements:

$$P_{kk}V_m^2 - P_{mm}V_k^2 = 0$$
$$Q_{kk}V_m^2 - Q_{mm}V_k^2 = 0.$$

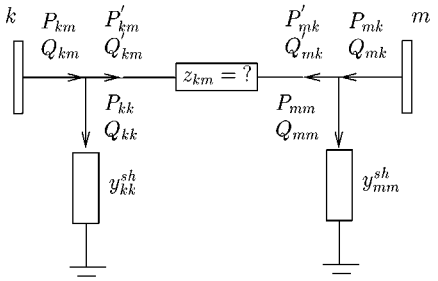


Fig. 5. Dormant parameter technique applied to the estimation of the series impedance of a π equivalent model.

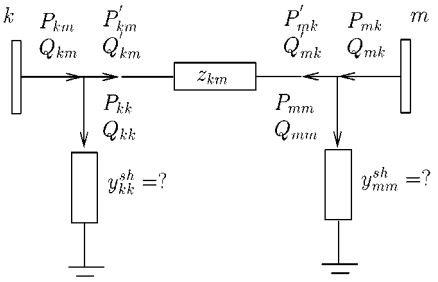


Fig. 6. Dormant parameter technique applied to the estimation of the shunt admittance of a balanced π equivalent model.

III. Network Topology Processor (NTP)

Conventional network topology processing identifies energized, de-energized, and grounded electrical islands and is performed before state estimation and other related functions (observability analysis and bad data processing) [15]–[17]. In this conventional approach, state estimation assumes that the topology is correct and proceeds to estimate the states and identify analog bad data whenever redundancy allows it. A complete description of the network model and the location of metering devices in terms of bus-sections and switching-devices is assumed to be available from a database. The NTP transforms the bus-section/switching-device model into the bus/branch model and assigns metering devices to the components of the bus/branch network model identified.

Hence, in more conventional implementations, the real-time modeling of a power network usually follows a six-step procedure involving: 1) data gathering; 2) network topology processing; 3) observability analysis; 4) state estimation; 5) processing of bad data; and 6) identification of network model. Step 1) assumes a bus-section/switching-device model. Steps 2) and 3) assume that switching device status is correct. Step 4) additionally assumes that the parameters are correct. Step 5) processes bad data assuming that they are caused by analog measurements.

A. Conventional NTP

The first task of the NTP is to convert raw analog measurements to the appropriate units and to verify operating limits, the rate of change of operating variables, the zero flows in open switching devices, and zero voltage differences across closed switching devices [19].

Next, bus sections are processed to determine the connectivity in bus section groups, which are sets of bus sections that become a single bus when all switches and breakers are considered closed. (e.g., in the substation shown in Fig. 1, if all the 500-kV switching devices are closed, then the bus sections connected by these switching devices will merge into a single bus).

In addition to switching devices, substations are associated with terminals of branch devices (e.g., transmission lines, transformers, phase shifters, and series devices), shunt devices (e.g., capacitors, reactors, synchronous condensers, static VAr compensators, loads and generators), and metering devices (e.g., power and current flow meters, power and current injection meters, and voltage magnitude meters) as well. The connection of these devices in a bus/branch model requires the determination of the network buses.

Bus section processing consists of merging bus sections of a bus section group into one or more network buses (buses of the bus/branch network model). Once such network buses are formed, data structures (pointers and links) are built to associate them with branch and shunt devices.

B. Topology Processor in Tracking Mode

In the tracking mode, the NTP updates the parts of the power network affected by status changes [20]. Only bus section groups where changes have occurred are processed and the associated data structures are updated accordingly. Switching device status changes can modify both the way bus sections are grouped into network buses and the association of branch and shunt devices with network buses. The location of metering devices may be affected as well. Thus, the new data structures relating network buses to various devices (branch, shunt, and metering) are compared with the corresponding structures saved from the previous run. Simple plausibility checks are then performed for changed bus section groups using Kirchhoff's laws.

There are cases in which changes in a bus section group also affect network connectivity (e.g., cases when a bus splits or a branch device switches from one bus to other). In these cases, the data structures describing network connectivity and network islands are updated by NTP in tracking mode. The benefits of the tracking mode are not limited to the topology processor itself: in the case of minor changes, or when no changes occur, matrix structures (including the optimal pivoting order) used in other applications can be reused up to a certain point, although eventually the cumulative effect of changes will require a complete matrix refactorization.

C. Generalized NTP

Identification of topology errors that pass undetected through configuration analysis using the conventional bus/branch network model is not always effective [18]. A more appropriate state estimation approach has been developed, the generalized state estimation, to cope with these critical cases. In this approach, an integrated estimation of state, status, and parameters is made.

Besides the regular functions performed by a conventional NTP, a generalized topology processor identifies extended islands in which switching devices that appear in the database as being open can be explicitly represented. Unknown or suspect statuses can also be represented explicitly in the model. For example, the status of switching devices of a bus section group in which one or more changes have occurred since the previous execution can be considered as suspect. Such explicit representation facilitates bad data processing, since possible status error will appear in state estimation as such and thus will not be disguised as errors in the analog measurements.

A local weighted least squares (LS) state estimator can be run using a bus-section/switching-device model for areas containing suspected data, such as a bus section group or a substation. Depending on the results of the state estimation, bad analog/status data can be removed or, if redundancy does not permit a safe decision, the suspect area is kept in the bus-section/switching-device model level for further evaluation during state estimation of the entire network.

Extended islands are crucial to generalized state estimation. Incorrect status may cause a conventional NTP to identify two islands where there is in fact only a single connected network, or it can lead to the identification of a de-energized island, whereas in reality the island is part of a larger, energized system. To cope with these situations the concept of extended island is introduced here. Fig. 7 illustrates the case in which the telemetered status of breaker 4–5 is incorrect: it reads open, whereas in the field, the breaker is actually closed. Rather than identifying two separate islands, the open breaker 4–5 is retained in an extended model. Hence, the incorrect status of breaker 4–5 can be detected and identified by the generalized state estimator.

A linear programming based state estimator for substation data validation was proposed in [21] and developed in [22]. The weighted least squares (WLS) estimator was extended to network represented at the physical level in [23].

IV. WLS State Estimator

Most state estimation programs in practical use are formulated as overdetermined systems of nonlinear equations and solved as WLS problems [3], [24].

A. Problem Formulation

Consider the nonlinear measurement model

$$z_j = h_j(\mathbf{x}) + e_j$$

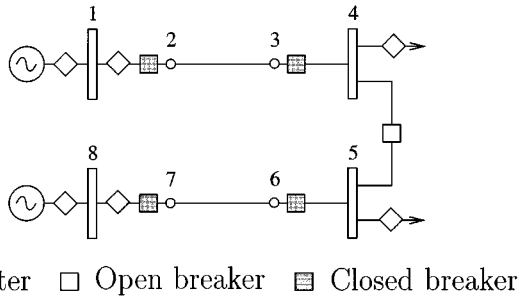


Fig. 7. Extended observable island.

where z_j is the jth measurement, $\mathbf{x}$ is the true state vector, $h_j(\cdot)$ is a nonlinear scalar function relating the jth measurement to states, and e_j is the measurement error, which is assumed to have zero mean and variance σ_j^2. There are m measurements and n state variables, $n < m$.

The WLS state estimation can be formulated mathematically as an optimization problem with a quadratic objective function and with equality and inequality constraints

$$\begin{aligned} \text{minimize} \quad & J(\mathbf{x}) = \frac{1}{2}\sum_{j=1}^{m} r_j^2/\sigma_j^2 \\ \text{subject to} \quad & g_i(\mathbf{x}) = \mathbf{0}; \quad i = 1, n_g \\ & c_i(\mathbf{x}) \le \mathbf{0}; \quad i = 1, n_c \end{aligned} \tag{1}$$

where r_j is the residual $z_j - h_j(\mathbf{x})$, $f(\cdot)$ is an objective function, and $g_i(\cdot)$ and $c_i(\cdot)$ are functions representing power flow quantities. Equality and inequality constraints are normally used to represent target values and operating limits in the unobservable parts of the network.

B. State Variables

Complex nodal voltages are the most commonly used variables. Turn ratios of transformers with taps that change under operating conditions are also treated as state variables. Power flows in branches that follow Ohm's law are dependent variables and can be determined from the state variables (nodal voltages and turn ratios). However, in branches for which application of Ohm's law is not fruitful, such as branches with unknown impedances, branches with zero impedances, or closed switches, flows cannot be determined from these state variables. In these cases one alternative consists in introducing power flows as additional states. In the case of open switches, although the flow is known, there is no relation between the voltage spread across the switch and the zero power flow; the flow state variable technique is then extended to open switches. (This helps detecting and identifying bad status information in switches that are wrongly considered as being open.)

Hence, the vector of state variables $\mathbf{x}$ usually includes the following states:

1) nodal voltage:
   a) voltage magnitude V_k;
   b) voltage angle θ_k;
2) transformer turns ratio:
   a) turns ratio magnitude t_{km};
   b) phase shift angle φ_{km};
3) complex power flow:
   a) active power flow P_{km} and P_{mk};
   b) reactive power flow Q_{km} and Q_{mk}.

C. Analog Measurements

The following measurements are normally included in practical state estimators:

1) voltage magnitude V_k;
2) voltage angle difference θ_{km};
3) active power:
   a) branch flow P_{km};
   b) branch-group flow $\sum P_{km}$ in a designated group of branches;
   c) bus injection P_k;
4) reactive power:
   a) branch flow Q_{km};
   b) branch-group flow $\sum Q_{km}$ in a designated group of branches;
   c) bus injection Q_k;
5) current magnitude flow $|I_{km}|$ in branch km, and injection $|I_k|$;
6) magnitude of turns ratio t_{km};
7) phase shift angle of transformer φ_{km};
8) active power flow P_{km}:
   a) in switches;
   b) in zero impedance branches;
   c) in branches of unknown impedance;
9) reactive power flow Q_{km}:
   a) in switches;
   b) in zero impedance branches;
   c) in branches of unknown impedance.

Remark 1: Not all measurements are obtained simultaneously and hence the measurements that are obtained in a measurement scan can be time-skewed up to a few seconds. Although in most of the cases this has no effect on the quality of state estimates, data inconsistencies may occur when the protection system is activated or when the system is ramping up or down at a rapid pace. In these cases, reliable estimates will be obtained again only when the system settles down. In this sense, the state estimate is not actually real time but is only a quasi-static representation of the conditions in the network. Truly dynamic, real-time estimation would require more sophisticated telemetering systems, probably based on time-tagged measurements obtained via Global Positioning System (GPS) [25].

D. Equality Constraints

Target values normally used in power flow studies can be included in state estimation in order to restore observability to those parts of the network which are permanently or temporarily unobservable. Branches where the application of Ohm's law is not appropriate can also be represented via equality constraints: e.g., the voltage difference across a short circuit branch is equal to zero.

1) Target or specified voltage magnitude V_k^{sp}.
2) Target voltage angle θ_k^{sp}.

3) Target flow P_{km}^{sp}.
4) Target reactive power flow Q_{km}^{sp}.
5) Target current magnitude flow $|I_{km}^{\mathrm{sp}}|$ and injection $|I_k^{\mathrm{sp}}|$.
6) Voltage magnitude difference $V_k - V_m$ in closed switches.
7) Voltage angle difference $\theta_k - \theta_m$ in closed switches.
8) Active power flow P_{km} in open switches.
9) Reactive power flow Q_{km} in open switches.
10) Current difference $I_{km} + I_{mk}$.
11) Admittance difference $\Delta Y_{km}^{\mathrm{sh}}$ in π equivalent models.

Equality constraints can be treated as such [26] or as pseudomeasurements with relatively high weights [12]. In the first case, in the WLS approach, the gain matrix becomes indefinite, which demands a special sparse factorization [39], [40] since the delayed-pivot approach does not necessarily guarantees nonzero pivots throughout the factorization process. In the second case, a numerically robust state estimator is normally required [27]–[30].

E. Inequality Constraints

Limits, such as minimum and maximum reactive power generation, can also be used to improve the representation of unobservable parts of the network. The most commonly used inequality constraints are the following:

1) VAr limit Q_k^{lim};
2) tap limit t_{km}^{lim};
3) phase-shift limit $\varphi_{km}^{\mathrm{lim}}$.

Inequality constraints are dealt with in the same way as in power flow and optimal power flow calculations [10], [12]. The constraints are initially relaxed, and as one approaches the solution those constraints that are violated are enforced on the corresponding limits, either as actual equality constraints or as pseudomeasurements with relatively high weights. Interior point algorithms have also been suggested in the literature [31].

V. Solution Approaches

The iterative normal equations method below is the standard approach to the solution of WLS state estimation in power systems. Ill-conditioning can occur, however, in connection with the use of widely different weighting factors, the presence of a large number of injection measurements, and the representation of low impedance branches which are incident to regular branches. All these problems are somehow related to the squared form of the gain matrix ($\mathbf{H}'\mathbf{H}$). The other methods discussed in this section were partly motivated by the need to unsquare the gain matrix in order to improve numerical robustness.

A. Normal Equations Method

The unconstrained state estimation problem can be formulated as a minimization of

$$J(\mathbf{x}) = \frac{1}{2}\sum_{j=1}^{m} r_j^2/\sigma_j^2 = \frac{1}{2}\mathbf{r}'\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r} = \frac{1}{2}\tilde{\mathbf{r}}'\tilde{\mathbf{r}} \qquad (2)$$

where $\tilde{\mathbf{r}} = \mathbf{R}_{\mathbf{z}}^{-1/2}\mathbf{r}$ is the weighted residual, $\mathbf{R}_{\mathbf{z}}$ being the diagonal weighting matrix of measurement variances. (The apostrophe denotes vector and matrix transposition throughout.)

The first-order optimal condition is as follows:

$$\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} = -\sum_{j=1}^{m} \frac{r_j}{\sigma_j^2}\frac{\partial h_j}{\partial \mathbf{x}} = -\mathbf{H}'\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r} = \mathbf{0} \qquad (3)$$

where $(\partial h_j/\partial \mathbf{x})'$ is the jth row of the Jacobian matrix $\mathbf{H}$. The root of (3) can be found using the Newton Raphson method. The Taylor expansion approximates the gradient function

$$\left.\frac{\partial J}{\partial \mathbf{x}}\right|_{(\mathbf{x}+\Delta\mathbf{x})} \simeq \left.\frac{\partial J}{\partial \mathbf{x}}\right|_{(\mathbf{x})} + \left.\frac{\partial^2 J}{\partial \mathbf{x}^2}\right|_{(\mathbf{x})}\Delta\mathbf{x} \qquad (4)$$

where the Hessian matrix is as follows:

$$\frac{\partial^2 J}{\partial \mathbf{x}^2} = \sum_{j=1}^{m}\sigma_j^{-2}\frac{\partial h_j}{\partial \mathbf{x}}\left(\frac{\partial h_j}{\partial \mathbf{x}}\right)' - \sum_{j=1}^{m}\frac{r_j}{\sigma_j^2}\frac{\partial^2 h_j}{\partial \mathbf{x}^2}. \qquad (5)$$

In the Newton Raphson method, the state update is obtained from (4) using the Hessian matrix in (5). This method presents quadratic convergence and has been adopted in a parameter estimation approach discussed in [32] (The importance of second order derivatives in the presence of erroneous data was emphasized in [33].) In the Gauss Newton method the second-order term of (5) is ignored. The effect in convergence is normally not significant, except in cases of large residuals caused by topology and parameter errors, combined with a strong nonlinearity of the measurement function $h_j(\mathbf{x})$. Most of practical implementations of state estimation in electric power systems is based on the Gauss Newton method. The state estimate $\hat{\mathbf{x}}$ is obtained by the following iterative procedure:

$$\mathbf{G}(\mathbf{x}^{\nu})\Delta\mathbf{x}^{\nu} = \mathbf{H}'(\mathbf{x}^{\nu})\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{r}(\mathbf{x}) \qquad (6)$$

$$\mathbf{x}^{\nu+1} = \mathbf{x}^{\nu} + \Delta\mathbf{x}^{\nu} \qquad (7)$$

where $\mathbf{G} = \partial^2 J/\partial \mathbf{x}^2$ is a gain matrix. In the Gauss Newton method $\mathbf{G} = \mathbf{H}'\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{H}$, whereas in the Newton Raphson method the second-order term of (5) should be included.

B. Orthogonal Methods

The need to represent equality constraints as pseudomeasurements with relatively high weights has caused a search for an alternative to the normal equations approach. The orthogonal transformation method is one such alternative which has found wide acceptance in practice. An orthogonal method based on row-wise Givens rotations was first suggested in [27]. An early implementation of Givens rotations in a production grade state estimator was reported in [28]. A hybrid approach based on seminormal equations was then proposed in [34] and a theoretical discussion about seminormal equations in general can be found in [35]. An efficient ordering scheme to preserve sparsity and minimize the number of intermediate fill-ins, along with a modified Givens rotations method (the 2-multiplication scheme), was reported in [30], and the inclusion of equality constraints in orthogonal state estimators was described [36].

Consider an $m \times n$ weighted Jacobian matrix $\tilde{\mathbf{H}}(\mathbf{x}^\nu) = \mathbf{R}_\mathbf{z}^{-1/2}\mathbf{H}(\mathbf{x}^\nu)$ with rank n. The orthogonal transformation method avoids squaring the gain matrix $\tilde{\mathbf{H}}'\tilde{\mathbf{H}}$ by using the following decomposition of the Jacobian matrix:

$$\tilde{\mathbf{H}} = \mathbf{Q}'\mathbf{U} = (\mathbf{Q}_1' \quad \mathbf{Q}_2') \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{0} \end{pmatrix} \tag{8}$$

where $\mathbf{Q}$ is an orthogonal $m \times m$ matrix and $\mathbf{U}_1$ is an upper triangular $n \times n$ matrix.

Using the orthogonal decomposition in (8), the normal equations (6) can be rewritten as follows:

$$\mathbf{U}'\mathbf{Q}\mathbf{Q}'\mathbf{U}\Delta\mathbf{x}^\nu = \mathbf{U}'\mathbf{Q}\tilde{\mathbf{r}}(\mathbf{x}^\nu).$$

Since $\mathbf{U}_1$ is nonsingular (the system is assumed to be observable), and $\mathbf{Q}$ is orthogonal, i.e., $\mathbf{Q}'\mathbf{Q} = \mathbf{Q}\mathbf{Q}' = \mathbf{I}_m$, the result is

$$\mathbf{U}_1\Delta\mathbf{x}^\nu = \mathbf{Q}_1\,\tilde{\mathbf{r}}(\mathbf{x}^\nu).$$

This equation can be solved in two stages, as follows:

$$\mathbf{y}_1 = \mathbf{Q}_1\tilde{\mathbf{r}}(\mathbf{x}^\nu)$$
$$\mathbf{U}_1\Delta\mathbf{x}^\nu = \mathbf{y}_1.$$

Remark 1: Pivoting is normally necessary to preserve sparsity. In this case a permutation of $\tilde{\mathbf{H}}$ can be factorized instead, i.e., $\mathbf{P}_r\tilde{\mathbf{H}}\mathbf{P}_c = \mathbf{Q}'\mathbf{U}$, where $\mathbf{P}_r$ performs row permutations on $\mathbf{H}$, and $\mathbf{P}_c$ performs column permutations.

Remark 2: The Euclidean norm of the estimation residual is invariant under an orthogonal transformation $\mathbf{Q}\tilde{\mathbf{H}} = \mathbf{U}$.

C. Seminormal Equations

If the orthogonal decomposition $\tilde{\mathbf{H}} = \mathbf{Q}'\mathbf{U}$ is applied only to the left side of the normal equations above, the following results are obtained (a fast decoupled version of this hybrid method was presented in [34]):

$$\mathbf{U}'\mathbf{Q}\mathbf{Q}'\mathbf{U}\Delta\mathbf{x}^\nu = \tilde{\mathbf{H}}'\tilde{\mathbf{r}}(\mathbf{x}^\nu).$$

Since $\mathbf{Q}$ is orthogonal, and $\mathbf{U}$ is upper trapezoidal, the result is

$$\mathbf{U}_1'\mathbf{U}_1\Delta\mathbf{x}^\nu = \tilde{\mathbf{H}}'\tilde{\mathbf{r}}(\mathbf{x}^\nu).$$

This equation can then be solved as in the normal equations approach, except that the triangular factor $\mathbf{U}_1$ is obtained via orthogonal transformations.

The corrected seminormal equations method is summarized in the following:

$$\mathbf{U}_1'\mathbf{U}_1\Delta\mathbf{x}^{\text{tmp}} = \tilde{\mathbf{H}}'\tilde{\mathbf{r}}(\mathbf{x}^\nu)$$
$$\tilde{\mathbf{r}}_c = \tilde{\mathbf{r}}(\mathbf{x}^\nu) - \tilde{\mathbf{H}}\Delta\mathbf{x}^{\text{tmp}}$$
$$\mathbf{U}_1'\mathbf{U}_1\Delta\mathbf{x}_c = \tilde{\mathbf{H}}'\tilde{\mathbf{r}}_c$$
$$\Delta\mathbf{x}^\nu = \Delta\mathbf{x}^{\text{tmp}} + \Delta\mathbf{x}_c.$$

For linear estimators, under mild conditions, it can be shown that this approach is as accurate as a full orthogonal method [35]. For nonlinear models, the above correction can be carried out together with the next iteration, with no additional computational effort [34].

D. Equality Constrained WLS Estimator

Rather than modeling equality constraints as pseudomeasurements with relatively high weights, they can be introduced in the optimization problem as such [26]. This was the first attempt to unsquare the gain matrix. To cope with the indefiniteness of the augmented gain matrix (zero pivots) the authors proposed a delayed pivoting scheme. This approach has two potential advantages: the pseudomeasurements corresponding to the equality constraints are not squared and no weights are assigned to the equality constraints. Mathematically this can be expressed as follows:

$$\begin{aligned} \text{minimize} \quad & J(\mathbf{x}) = \frac{1}{2}\mathbf{r}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{r} \\ \text{subject to} \quad & \mathbf{c}(\mathbf{x}) = \mathbf{0} \end{aligned}$$

where $\mathbf{c}(\mathbf{x}) = \mathbf{0}$ represents a set of nonlinear constraints. This optimization problem can be expressed by the following Lagrangian function:

$$\mathcal{L}(\mathbf{x}, \mathbf{\Lambda}) = \frac{1}{2}\mathbf{r}'(\mathbf{x})\mathbf{R}_\mathbf{z}^{-1}\mathbf{r}(\mathbf{x}) - \mathbf{\Lambda}'\mathbf{c}(\mathbf{x}).$$

The corresponding Karush-Kuhn-Tucker (KKT) first-order necessary conditions are as follows:

$$\partial\mathcal{L}/\partial\mathbf{x} = -\mathbf{H}'(\mathbf{x})\mathbf{R}_\mathbf{z}^{-1}\mathbf{r}(\mathbf{x}) - \mathbf{C}'(\mathbf{x})\mathbf{\Lambda} = \mathbf{0} \tag{9}$$
$$\partial\mathcal{L}/\partial\mathbf{\Lambda} = -\mathbf{c}(\mathbf{x}) = \mathbf{0} \tag{10}$$

where $\mathbf{H}(\mathbf{x}) = \partial\mathbf{h}/\partial\mathbf{x}$ and $\mathbf{C}(\mathbf{x}) = \partial\mathbf{c}/\partial\mathbf{x}$.

This system of nonlinear equations can be solved iteratively by the Gauss Newton method using Taylor expansions as follows:

$$\mathbf{r}(\mathbf{x}) \simeq \mathbf{r}^\nu - \mathbf{H}(\mathbf{x}^\nu)\Delta\mathbf{x}^\nu$$
$$\mathbf{c}(\mathbf{x}) = \mathbf{c}(\mathbf{x}^\nu) + \mathbf{C}(\mathbf{x}^\nu)\Delta\mathbf{x}^\nu.$$

In view of these linear approximations, (9) and (10) can be rewritten as follows:

$$\begin{pmatrix} \tilde{\mathbf{H}}'\tilde{\mathbf{H}} & -\mathbf{C}' \\ -\mathbf{C} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Delta\mathbf{x}^\nu \\ \Lambda^{\nu+1} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{H}}'\tilde{\mathbf{r}}(\mathbf{x}^\nu) \\ \mathbf{c}(\mathbf{x}^\nu) \end{pmatrix}. \tag{11}$$

Remark 1: Notice that the coefficient matrix in (11) is indefinite, and if the pivot ordering is based on sparsity considerations only, the factorization process can break down due to the occurrence of zero pivots [26]. Delayed pivoting has been used to avoid zero or small pivots, although there is no guarantee that this procedure will avoid all possible zero pivots. Alternatively, blocked sparse matrices [37], [38] and mixed 1×1 and 2×2 pivoting [39], [40] have been used.

E. Sparse Tableau Formulation—Hachtel Method

The unsquared representation of equality constraints can be extended to regular measurements as follows:

$$\begin{aligned} \text{minimize} \quad & J(\mathbf{r}) = \frac{1}{2}\mathbf{r}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{r} \\ \text{subject to} \quad & \mathbf{r} = \mathbf{z} - \mathbf{h}(\mathbf{x}) \\ & \mathbf{c}(\mathbf{x}) = \mathbf{0}. \end{aligned}$$

The corresponding Lagrangian function is

$$\mathcal{L}(\mathbf{x}, \mathbf{r}, \mathbf{\Lambda}) = \frac{1}{2}\mathbf{r}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{r} - \Lambda'\mathbf{c}(\mathbf{x}) - \Gamma'(\mathbf{r} - \mathbf{z} + \mathbf{h}(\mathbf{x})).$$

The KKT first order necessary conditions for an optimal solution are expressed by the following augmented tableau (Hachtel tableau):

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & \tilde{\mathbf{H}} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \\ \tilde{\mathbf{H}}' & \mathbf{C}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Gamma^{\nu+1} \\ \Lambda^{\nu+1} \\ \Delta\mathbf{x}^{\nu} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{r}}(\mathbf{x}^{\nu}) \\ -\mathbf{c}(\mathbf{x}^{\nu}) \\ \mathbf{0} \end{pmatrix}. \quad (12)$$

The Hachtel method was first applied to power system state estimation in [39] and was further studied in [41], which extended the normalized residuals approach to Hachtel state estimators. The use of blocked matrices was simultaneously suggested in [37] and [38], and an extension of the blocked matrix approach to numerical observability analysis and bad data processing was presented in [42].

F. Blocked Sparse Tableau

The sparsity of the factors of the Hachtel tableau can be improved if the tableau is put into a blocked form with the same structure as the $\mathbf{Y}$-matrix and factorized as such. For networks modeled at the bus/branch model, the blocked form is obtained by grouping measurements as follows:

$$\mathbf{z} = \begin{pmatrix} \mathbf{z}_b \\ \mathbf{z}_n \end{pmatrix}$$

where $\mathbf{z}_b$ includes both branch flow measurements and nodal voltage measurements, and $\mathbf{z}_n$ contains the nodal injection measurements. The Jacobian matrix is partitioned accordingly, i.e.

$$\tilde{\mathbf{H}} = \mathbf{R}_{\mathbf{z}}^{-1/2}\mathbf{H} = \begin{pmatrix} \mathbf{R}_{\mathbf{b}}^{-1/2}\mathbf{H}_b \\ \mathbf{R}_{\mathbf{n}}^{-1/2}\mathbf{H}_n \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{H}}_b \\ \tilde{\mathbf{H}}_n \end{pmatrix}.$$

The Hachtel tableau in (12) can then be rewritten as follows:

$$\begin{pmatrix} \mathbf{I}_b & \mathbf{0} & \mathbf{0} & \tilde{\mathbf{H}}_b \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} & \tilde{\mathbf{H}}_n \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C} \\ \tilde{\mathbf{H}}'_b & \tilde{\mathbf{H}}'_n & \mathbf{C}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Gamma_b^{\nu+1} \\ \Gamma_n^{\nu+1} \\ \Lambda^{\nu+1} \\ \Delta\mathbf{x}^{\nu} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{r}}_b(\mathbf{x}^{\nu}) \\ \tilde{\mathbf{r}}_n(\mathbf{x}^{\nu}) \\ -\mathbf{c}(\mathbf{x}^{\nu}) \\ \mathbf{0} \end{pmatrix}.$$

Applying Gauss elimination to zeroize the submatrix $\tilde{\mathbf{H}}'_b$ results in the following tableau:

$$\begin{pmatrix} \mathbf{I}_n & \mathbf{0} & \tilde{\mathbf{H}}_n \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \\ \tilde{\mathbf{H}}'_n & \mathbf{C}' & -\tilde{\mathbf{H}}'_b\tilde{\mathbf{H}}_b \end{pmatrix} \begin{pmatrix} \Gamma_n^{\nu+1} \\ \Lambda^{\nu+1} \\ \Delta\mathbf{x}^{\nu} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{r}}_n(\mathbf{x}^{\nu}) \\ -\mathbf{c}(\mathbf{x}^{\nu}) \\ -\tilde{\mathbf{H}}'_b\tilde{\mathbf{r}}_b(\mathbf{x}^{\nu}) \end{pmatrix}.$$

If there is only one injection measurement or constraint per node, this tableau can be further arranged to have the same block structure of the $\mathbf{Y}$-matrix.

Remark 1: The selection of pivots based purely on sparsity may lead to block pivots with zero determinant. Hence, both the identification of singularities and the modification of the blocking (or of the ordering) scheme is still necessary. (See the closure of [37]).

VI. Observability Analysis

If there are enough measurements and they are well distributed throughout the network in such a way that state estimation is possible, the network is said to be observable. If a network is not observable, it is still useful to know which portion has a state which can be estimated, i.e., it is important to determine the observable islands. In the observable parts of a network, measurement redundancy is defined as the ratio of the number of measurements to the number of states; in most practical cases the redundancy is in the range 1.7–2.2. A critical measurement is a nonredundant measurement, i.e., a measurement that when removed turns the network unobservable.

There are three principal types of algorithms for observability analysis: topological; numerical; and hybrid. The concept of topological observability in power network state estimation was introduced in [55] as a partial requirement for state estimation solvability and further developed in [56]. Practical solvability also depends on numerical aspects [57] of state estimation, and a numerical approach to observability analysis which could take both topological and numerical aspects into consideration was suggested [58], [59].

A. Extended Observable Islands

The inclusion of breakers, switches, zero impedance branches, and branches with unknown impedances in the generalized state estimation model has motivated the following extended definitions of islands and observable islands.

1) *Definition A:* An island is a contiguous part of a network with bus sections as nodes and lines, transformers, open switches, closed switches, and switches with unknown status as branches.
2) *Definition B:* An observable island is an island for which all branch flows can be calculated from the available measurements independent of the values adopted for reference pseudomeasurements [60].

According to Definition A, the network shown in Fig. 7 forms an island, even considering that the breaker 4–5 is open. And according to Definition B, this network forms a single observable island, since the power flows in all branches are observable. (Notice that this would not be so if the injections at buses 4 and 5 were unmetered, in which case we would have two observable islands.) Fig. 8 further illustrates the extended concepts of observability. In this case the two complex nodal voltages will be known since the voltage magnitude at bus-1 is metered. This, however, does not mean the network is observable since the flow distribution between the two zero impedance branches is indeterminate; there could be an arbitrary circulating flow in the two branches and still they will match the injection measurements. Hence, according to Definition B, the network is unobservable, i.e., it is impossible to determine the power flows from the available measurements.

B. Topological Observability Analysis

The concept of topological observability was originally proposed for networks represented by bus/branch models and is linked to the idea of maximal spanning tree, which is a tree that contains every node in the graph representing a network [55], [56]. To each branch of a maximal spanning tree is assigned a measurement. A measurement assignment satisfies: 1) different branches are always assigned to different measurements; 2) if the flow in a branch is measured the branch

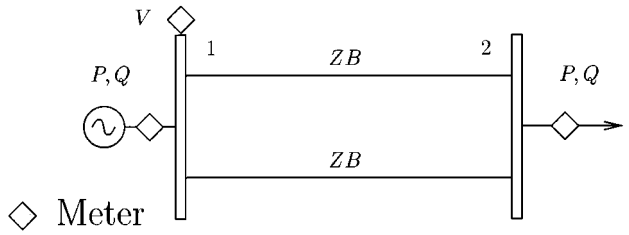


Fig. 8. Unobservable network. (ZB is a zero impedance branch.)

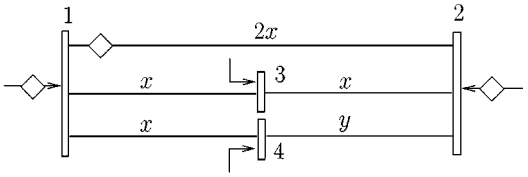


Fig. 9. Topologically observable network. Solvability requires that $y \neq x$.

is assigned to that measurement; 3) a branch with unmeasured flow is assigned to an injection measurement at a node incident to the branch [61]. Although it is generally agreed that the topological approach can be extended to handle the additional states and pseudomeasurements of the generalized state estimation, further research is still needed (The same can be said regarding the consideration of current magnitude measurements [62], [63].)

Topological observability does not necessarily guarantee solvability of the state estimation problem. Consider, for example, the dc model for the system in Fig. 9, where the reactances of lines are marked alongside. In this case, there are four state variables $(\theta_k, k = 1, 2, 3, 4)$ and four measurements ($\theta_1^{\text{ref}} = 0$, P_{12}^{meas}, P_1^{meas}, and P_2^{meas}). If $y \neq x$, the rank of the Jacobian matrix is 4, whereas for $y = x$ the rank is 3, in which case the gain matrix becomes singular and the system is unsolvable (although it is observable in the topological sense). Numerical problems can also occur for values of y close to x, which may render the problem to be numerically unsolvable. Although numerical coincidences are relatively rare, they do occur in practice. (See [64] for an example involving transformers.)

C. Numerical Observability Analysis

The numerical observability algorithm below was designed to handle networks that are totally or partially represented at the physical level. The basic modifications regarding the algorithm for networks modeled at the bus/branch level is the addition of new state variables and new pseudo measurements as discussed in Section II.C. The extended algorithm is also based on the presence of zero pivots that may occur during triangular factorization of the gain matrix. The difference is that the gain matrix will include additional information as well as new states, and since power flows are also state variables, zero pivots may occur in connection with these variables as well. When this happens, state variables corresponding to these zero pivots are added as pseudomeasurements with arbitrary values, just as in the bus/branch model. Irrelevant measurements are again identified, i.e., injection measurements with incident branches with estimated flows being a function of the arbitrary values assigned to the pseudo measurements added (nonzero flows) are considered irrelevant [64].

Algorithm

1) Initialization.
   a) Initialize the measurement set of interest as consisting of all available measurements and pseudo measurements representing switches and short circuits.
   b) Initialize the power network of interest as consisting of all branches incident to at least one measurement or pseudo measurement.
2) Form gain matrix $\mathbf{G}$ and perform triangular factorization $\mathbf{G} = \mathbf{U}'\mathbf{U}$.
3) Introduce angle/flow pseudo measurements whenever a zero pivot is encountered.
4) Solve the dc state estimator equation for the angle/flow state variables considering all the measured values equal to zero, except for the added angle/flow pseudomeasurements that are assigned arbitrary values.
5) System update.
   a) Remove from the power network of interest all the branches with nonzero flows.
   b) Update the measurement set of interest by removing power injection measurements adjacent to the removed branches along with the corresponding pseudomeasurements.
   c) If modifications have been made, update the triangular factor $\mathbf{U}$ and go to 3).
6) Form islands with nodes connected by branches with zero flows.

Remark 1: The algorithm above can be extended to other methods such as the orthogonal [65] method and the Hachtel method [42]. The results provided by numerical observability analysis, however, can be drastically affected if singularity occurs for reasons other than unobservability (topological or numerical), as may be the case with the blocking approach and with the equality constrained approach based on pivot delay. As a rule, reliable factorization methods are always necessary when indefinite matrices are used.

D. Hybrid Observability Analysis

The hybrid algorithm exploits the best of both topological and numerical approaches: a basic topological algorithm with simple injection conversion to obtain one or more islands which are as large as possible, and a numerical algorithm for application to the reduced system [10], [66]. The topological algorithm is initially used to process flow measurements and injection measurements for which all except one of the incident branches are observable, whereas injection measurements for which branch assignment is not straightforward are left to be treated by the numerical algorithm using a reduced model. Only the boundary nodes of the islands obtained via the topological algorithm are retained for the numerical analysis. A tree of angle difference

pseudomeasurements are then associated with these nodes in order to take into account the effect of the measurements that have already been processed by the topological algorithm. Such measurements, in addition to the unprocessed injection measurements (measurements for which branch assignment was not straightforward), are then processed by the numerical algorithm.

VII. Bad Data Analysis

It may occur that the errors affecting state estimates are not compatible with their standard deviations. Most frequently, this happens due to the presence of bad data among the measured quantities. Alternatively, the proposed model may be unfit to explain the measured quantities. In either case, it is important to determine a plausible explanation, or a set of plausible explanations, for the observed data inconsistency. Methods used for detecting and identifying bad data are discussed in this section [67].

A. Largest Normalized Residual (LNR)

The normalized residual r_j^n is defined for noncritical measurements z_j as the ratio of the residual estimate $\hat{r}_j = z_j - \hat{z}_j$, to the standard deviation of the residual, ρ_{jj} [24]. Assuming that all other measurements are perfect and that z_j is a bad datum, it can be shown that no other measurement can have a normalized residual larger than that of z_j. This means that the wrong measurement is responsible for the largest normalized residual, i.e., although other measurements with the same residual magnitude may exist, none will have a residual larger than that of the ith measurement.

Assuming that the measurement errors e_j, $j = 1, \ldots, m$ are normal, independent with zero mean and variance σ_j^2, the vector of estimation residuals

$$\hat{\mathbf{r}} = \mathbf{S}\mathbf{e}$$

follows a ν-variate normal distribution; $\nu = m - n$ is the number of degrees of freedom, where m is the number of measurements and n is the number of state variables.

Since $\hat{r}_j = \mathbf{S}_{j\bullet}\mathbf{e}$, where $\mathbf{S}_{j\bullet}$ is the jth row of the sensitivity matrix $\mathbf{S}$, $\hat{r}_j$ is normally distributed with mean zero and variance ρ_{jj}^2, where $\rho_{jj}^2 = \mathbf{R}_{\hat{\mathbf{r}}}(j,j)$ (the diagonal elements of the covariance matrix of the residuals). Hence, the marginal distribution of each normalized residual $r_j^n = \rho_{jj}^{-1}\hat{r}_j$, is normal with mean zero and unit variance, i.e., $E\{r_j^n\} = 0$ and $E\{(r_j^n)^2\} = 1$ [24].

The largest normalized residual method was extended to the normal equations approach with equality constraints [41] and to the blocked sparse matrix approach [42]. The method was further extended in [68] to the Lagrange multipliers associated with equality constraints.

B. The $J(\hat{\mathbf{x}})$ Performance Index

If x_j, $j = 1, \ldots, m$ are random independent variables which follow a normal distribution with mean zero and unit variance, the chi-square distribution with m degrees of freedom (χ_m^2) is the distribution of the random variable y defined as $y = \sum_{j=1}^{m} x_j^2$. It can be shown that the χ_m^2 distribution has mean m and variance $2m$. When random variables x_j are constrained by n independent equations, then y follows a χ_{m-n}^2 distribution with the number of degrees of freedom reduced by the number of constraints. As the number of degrees of freedom ν increases, the χ_ν^2 distribution tends to a normal distribution (Central Limit Theorem).

Under the conditions above, the expected value of $J(\hat{\mathbf{x}})$ and its variance are as follows: $E\{J(\hat{\mathbf{x}})\} = m - n$ and $E\{[J(\hat{\mathbf{x}}) - (m-n)]^2\} = 2(m-n)$. Note that for an observable system with $m = n$, the state estimate $\hat{\mathbf{x}}$ fits the measurement model perfectly, i.e., all the residuals will be zero and $J(\hat{\mathbf{x}}) = 0$.

C. Hypotheses Testing

Once the estimates $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ are determined by the state estimator, one has to check whether these estimates are compatible in accuracy to their standard deviations. To this end, the $J(\hat{\mathbf{x}})$ test and the $\mathbf{r}^n$ test can be used.

1) The $J(\hat{\mathbf{x}})$ Test: Consider that an observation $J(\hat{\mathbf{x}})$ of the random variable $J(\mathbf{x})$ is yielded by a state estimation run. Based on this observation, the decision as to whether or not it actually belongs to the hypothesized χ^2 distribution must be made. The hypotheses made regarding the behavior of the random variables e_j will also be tested—indirectly—for normal distribution, $N(0, \sigma_j^2)$, or the effect of gross errors or biases.

Consider a null hypothesis H_0 and an alternative hypothesis H_1 as follows:

1) H_0, $E\{J(\hat{x})\} = m - n$;
2) H_1, $E\{J(\hat{x})\} > m - n$.

The alternative hypothesis suggests the way to perform the hypothesis test, i.e.:

1) if $J(\hat{\mathbf{x}}) > C$, reject H_0;
2) if $J(\hat{\mathbf{x}}) \leq C$, accept H_0

where C is a constant to be determined. If α is the significance level of the test, then constant C is given by

$$C = \chi_{m-n,1-\alpha}^2.$$

This means that if H_0 is true, the probability of $J(\hat{\mathbf{x}}) > C$ is α (or $\alpha \times 100\%$), i.e.

$$\int_0^C f(t)\,dt = 1 - \alpha \quad \text{with } f(t) = \frac{t^{\frac{\nu}{2}-1}e^{-\frac{t}{2}}}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}$$

where $f(t)$ is the probability density function of the χ_ν^2 distribution, where $\nu = m - n$ degrees of freedom and Γ is the Gamma function.

2) The $\mathbf{r}^n$ Test: Consider now the test for normalized residuals. As above, two hypotheses are considered: H_0: $E\{r_j^n\} = 0$ and H_1: $E\{|r_j^n|\} > 0$. H_0 is rejected if $|r_j^n| > C$, where C is a constant to be determined. The significance level of the test is given by α and corresponds to the probability of false alarm, i.e., $C = K_{\frac{\alpha}{2}}$. This means that α is the probability that $|r_j^n| > C$, i.e.

$$\int_{-C}^{C} f(t)\,dt = 1 - \alpha \quad \text{with } f(t) = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}$$

where $f(t)$ is the probability density function of the standard normal distribution with mean zero and unit variance.

Remark 1: Notice that the significance level α is also equivalent to the probability of a false alarm, or the probability of the occurrence of a Type 1 error: reject H_0 when it is actually true. (Type 2 error occurs if H_0 is accepted when it is actually false [67].)

3) Testing Equality Constraint Hypotheses: Equality constraints can be represented as either hard or soft constraints: 1) hard constraints consist of representing the equality constraints explicitly in the state estimation minimization problem [31] and 2) soft constraints consist of representing equality constraints by means of pseudomeasurements with the appropriate weightings [10]. When nonlinear models are used, both methods can suffer from convergence problems when the constraint represents false information, such as reporting a breaker status as open when it is actually closed in the field or using a bus as a transition bus (zero injection) when it is actually a load bus.

An alternative approach consists of testing hypotheses about equality constraints without explicitly adding these constraints to the state estimation problem or considering them as pseudomeasurements. Assume that the results of a state estimation run are available and that one wishes to test the following null hypothesis

$$\mathbf{Ax} = \mathbf{a}$$

where $\mathbf{x}$ is an n-vector of state variables, $\mathbf{a}$ is a constant p-vector, and $\mathbf{A}$ is a constant $p \times n$ matrix. When testing this hypothesis, the expected value of $\mathbf{A}\hat{\mathbf{x}}$ and the corresponding covariance matrix are as follows: $E\{\mathbf{A}\hat{\mathbf{x}}\} = \mathbf{a}$ and $\mathrm{Cov}\{\mathbf{A}\hat{\mathbf{x}}\} = \mathbf{A}(\mathbf{H}'\mathbf{W}\mathbf{H})^{-1}\mathbf{A}'$, where $\mathbf{W} = \mathbf{R}_\mathbf{z}^{-1}$. Since $\hat{\mathbf{x}}$ is n-variate normal, if $\mathrm{Rank}(A) = p \leq n$, then $\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}$ is p-variate normal with covariance given by the above expression, and the performance index is expressed as

$$J(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}) = (\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})'(\mathrm{Cov}\{\mathbf{A}\hat{\mathbf{x}}\})^{-1}\,(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a}).$$

This follows a χ_p^2 distribution, i.e., chi-square distribution with p degrees of freedom. The statistic defined by $J(\mathbf{A}\hat{\mathbf{x}} - \mathbf{a})$ can then be used to test the null hypothesis $\mathbf{Ax} = \mathbf{a}$, in the same way of the $J(\hat{\mathbf{x}})$-test describe above [69].

D. Dormant and Perfect Measurements

In performing bad data analysis, it may be necessary to investigate the effect on state estimates of considering a measurement either as a bad datum or as a perfect datum. In this section, an analytical approach to do this type of analysis is summarized.

First, consider the following partition of the measurement set: 1) $\mathbf{r}^o$ denotes a $o \times 1$ vector of residuals, with $o \geq n$, corresponding to a selected set of measurements that assures network observability (if, for example, $o = n$, the system is minimally observable, i.e., there are no redundant measurements) and 2) $\hat{\mathbf{r}}_k$ denotes a $k \times 1$ vector of residuals, with $k \leq m - n$, corresponding to a set of redundant measurements. Given this partition, the sensitivity relation $\hat{\mathbf{r}} = \mathbf{Se}$ can be written as

$$\begin{pmatrix} \hat{\mathbf{r}}_k \\ \hat{\mathbf{r}}_o \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{k\bullet} \\ \mathbf{S}_{o\bullet} \end{pmatrix} (\,\mathbf{e}\,)$$

where submatrices $\mathbf{S}_{k\bullet}$ and $\mathbf{S}_{o\bullet}$ are formed by the rows of the sensitivity matrix corresponding to measurements k and o, respectively.

Since the sensitivity matrix $\mathbf{S}$ is idempotent, i.e., $\mathbf{S} = \mathbf{S}^2$, then

$$\begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{k\bullet} \\ \mathbf{S}_{o\bullet} \end{pmatrix} (\,\mathbf{S}_{\bullet k} \quad \mathbf{S}_{\bullet o}\,)$$

where $\mathbf{S}_{\bullet k}$ and $\mathbf{S}_{\bullet o}$ are formed by the columns of the matrix $\mathbf{S}$, corresponding to measurements k and o, respectively.

Hence, $\mathbf{S}_{kk} = \mathbf{S}_{k\bullet}\mathbf{S}_{\bullet k}$. A similar relation holds for the weighted sensitivity matrix $\mathbf{S}^w = \mathbf{R}_\mathbf{z}^{-1/2}\mathbf{S}\mathbf{R}_\mathbf{z}^{1/2}$, i.e., $\mathbf{S}_{kk}^w = \mathbf{S}_{k\bullet}^w\mathbf{S}_{\bullet k}^w$ ($\mathbf{R}_\mathbf{z} = \mathbf{W}^{-1}$ is the measurement covariance matrix).

1) Dormant Measurement: Consider now that all the measurements, except those in the k-set, have no errors, i.e., $\mathbf{e} = (\mathbf{e}_k\mathbf{e}_o)'$, with $\mathbf{e}_o = \mathbf{0}$. In this case the sensitivity relation $\hat{\mathbf{r}} = \mathbf{Sz}$ can be written as

$$\begin{pmatrix} \hat{\mathbf{r}}_k \\ \hat{\mathbf{r}}_o \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} \begin{pmatrix} \mathbf{e}_k \\ \mathbf{0} \end{pmatrix}.$$

As with the situations involving single and multiple noninteracting bad data, discussed in the previous section, an error estimate is given by

$$\hat{\mathbf{e}}_k = \mathbf{S}_{kk}^{-1}\hat{\mathbf{r}}_k. \tag{13}$$

This error estimate can be used to correct the measurements in the k-set, yielding

$$\begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} \begin{pmatrix} \mathbf{e}_k - \mathbf{S}_{kk}^{-1}\hat{\mathbf{r}}_k \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

The correction in (13) zeroizes the estimation residuals corresponding to the k measurements. Since all the other measurements are assumed to be perfect, the residuals of the measurements of the o-set become zero as well. In general, if the o-set measurements are affected by errors, then only the residuals corresponding to the k-set measurements become zero after measurement correction.

In the unidimensional case (the k-set has a single element) (13) can be rewritten as follows:

$$\hat{e}_j = \frac{\sigma_j^2\hat{r}_j}{\rho_{jj}^2}$$

where $\rho_{jj}^2 = S_{jj}\sigma_j^2$.

If the error estimate is used to correct the corresponding measurements the impact of this correction on the performance index is as follows:

$$J(\hat{\mathbf{x}}^{\mathrm{new}}) = J(\hat{\mathbf{x}}) - \left(r_j^n\right)^2.$$

Assuming that only a single bad datum is present, the removal of the data presenting the largest normalized residual will eliminate the effect of the suspect data and provide the maximum reduction in the performance index. Hence, in these cases, detecting and identifying bad data by the largest normalized residual criterion turns out to be the same as

selecting the measurement whose removal would produce the largest reduction in the performance index.

The estimation of measurement errors and the correction of measured values were discussed in [70]–[72]. The interpretation of the residual estimates as the result of a state estimation based on parts of the sensitivity relation $\mathbf{S}\hat{\mathbf{r}} = \mathbf{e}$ was introduced in [73] and [74] and applied to a hypotheses testing procedure designed to detect and identify interacting multiple bad data in [75].

2) Perfect Measurement: Consider the following condition is imposed:

$$\begin{pmatrix} \Delta\mathbf{z}_k \\ \hat{\mathbf{r}}_o^{\text{new}} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{kk} & \mathbf{S}_{ko} \\ \mathbf{S}_{ok} & \mathbf{S}_{oo} \end{pmatrix} \begin{pmatrix} \mathbf{e}_k + \Delta\mathbf{z}_k \\ \mathbf{e}_o \end{pmatrix} \tag{14}$$

where $\Delta\mathbf{z}_k$ represents fictitious corrections on the measurements of the k-set such that the corresponding residual becomes $\hat{r}_k = \Delta\mathbf{z}_k$, i.e., the component of the residual $\hat{r}_k$ due to the errors $\mathbf{e}_k$ and $\mathbf{e}_o$ is zeroized. This is the same as assuming the measurements in the k-set are imposed as equality constraints or as pseudomeasurements with weights equal to infinity. Now, considering that

$$\hat{\mathbf{r}}_k = \mathbf{S}_{kk}\mathbf{e}_k + \mathbf{S}_{ko}\mathbf{e}_o$$

the first row of (14) yields

$$\Delta\mathbf{z}_k = \hat{\mathbf{r}}_k + \mathbf{S}_{kk}\Delta\mathbf{z}_k.$$

Hence, introducing corrections into the measurement vector yields

$$\Delta\mathbf{z}_k = (\mathbf{I}_k - \mathbf{S}_{kk})^{-1}\hat{\mathbf{r}}_k \tag{15}$$

where $\mathbf{I}_k$ is the unit matrix of order k and is equivalent to modeling the k measurements as equality constraints (or as pseudomeasurements with weights equal to infinity in a WLS estimator).

Remark 1: In (13), if the set σ guarantees observability, then $\det[\mathbf{S}_{kk}] \neq 0$, whereas in (15), if the set K is formed by nonredundant measurements, then $\det[\mathbf{I} - \mathbf{S}_{kk}] \neq 0$.

E. Identification Test

When bad data are detected, bad data identification normally has to be performed. This amounts to determining one or more plausible explanations for the data inconsistencies revealed by state estimation results. In order to do that, it may be useful to distinguish between the reliability of a measurement and its precision; measurement reliability refers to the probability of failure of a meter, whereas measurement precision refers to the magnitude of error when that meter is working. For example, a meter designed for great accuracy can still be quite unreliable. Some of the estimation methods discussed above take into account measurement precision by using weighted residuals in their objective functions; measurement reliability, however, is not explicitly considered in these formulations.

If the jth measurement is assumed to be a bad data it is denoted by $d_j = 0$, otherwise it is denoted by $d_j = 1$ [76]. To each possible combination of good and bad measurements is associated with a decision vector

$$\mathbf{d} = (d_1, d_2, d_3, \ldots, d_m).$$

For an observable system with m measurements there are 2^m possible decision vectors.

Any given decision vector is a plausible explanation for the observed data inconsistencies if, after the removal of the suspect data or of their effect on state estimation, no further bad data are detected. If, for example, the $J(\hat{\mathbf{x}})$ test is used, this means that $J(\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}) < C$, where $\hat{\mathbf{x}} + \Delta\hat{\mathbf{x}}$ is the new state estimate and C is the appropriate detection threshold. Hence, in any practical problem, a number of plausible explanations are likely to exist. The next step is to find the most plausible explanation or the set of most plausible explanations according to a prespecified criterion.

Let p_j be the probability that meter j is up and $q_j = 1 - p_j$ the probability that it is down. Let U be the set of all meters (and the corresponding part of the data gathering system) that is up at any given time t and D the set that is down. The probability associated with this particular decision vector $\mathbf{d}$ is

$$\text{Prob}(\mathbf{d}) = \prod_{j \in U} p_j \prod_{i \in D} q_i. \tag{16}$$

An optimal decision $\mathbf{d}$ is sought that maximizes $\text{Prob}(\mathbf{d})$. Now, if it is assumed that: 1) p_j is close to one and 2) the number of elements in U is usually much larger than that in D, and all meters have the same reliability, i.e., q_j is constant for all j, it can be shown [76] that a solution that maximizes $\text{Prob}(\mathbf{d})$ also maximizes the objective function

$$F(\mathbf{d}) = \sum_{j=1}^{m} (1 - d_j). \tag{17}$$

Hence, given the above assumptions, the optimal solution (or solutions if there is more than one) is the one that leads to the identification of the minimum number of bad measurements.

Let $N(\mathbf{d})$ denote the network and corresponding meter configuration for a given decision vector $\mathbf{d}$, $J(\hat{\mathbf{x}}(\mathbf{d}))$ denote the performance index, and $C(\mathbf{d})$ denote the detection threshold. The identification problem can then be formulated as follows:

$$\begin{array}{ll} \text{minimize} & F(\mathbf{d}) \\ \text{subject to} & N(\mathbf{d}) \text{ is observable} \\ & J(\hat{\mathbf{x}}(\mathbf{d})) < C(\mathbf{d}). \end{array} \tag{18}$$

Remark 1: When there is a single optimal solution, the determination of only that solution may be desirable. But when multiple optimal solutions are possible, the determination of all of them may be preferred. Alternatively, a set of all the solutions within a certain distance from the optimal ones may be of interest, such as when the optimal solution involves the removal of two measurements, and all the plausible solutions involving the removal of up to three measurements may be of interest.

Remark 2: In principle, the combinatorial problem above can be formulated to deal with analog, topology, and parameter errors simultaneously. This is a very complex combinatorial problem, however, for which further research is needed [11].

F. Pocketing and Zooming

As discussed above, the identification of interacting, conforming bad data is a hard combinatorial problem. The time required to reach a plausible solution will vary with the case under consideration. This can become an issue in a real-time environment. The use of pockets aims to focus the analysis to a limited part of the network and thus to reduce the computational burden of bad data identification (gross errors normally propagate poorly in state estimation) and can be formed around a predetermined set of suspect bad data, which can be determined by an approximate method such as the sequential largest weighted residual criterion. This allows the determination of the parts of the region of interest affected by bad data. Once suspect measurements are flagged, their neighborhood can be located using the topological algorithm based on the Jacobian matrix structure. A measurement corresponds to a row entry in the Jacobian matrix and columns of the nonzero elements in a row define the states that are adjacent to the measurement. A state corresponds to a column entry in the Jacobian matrix, and the rows of the nonzero elements in the column define the measurements that are adjacent to the state [10].

A common cause of multiple bad data that are both interacting and conforming is the presence of topological errors. Parts of the network which may contain topological errors can be zoomed in depending on the results of bad data analysis performed at the bus/branch level. These areas are then modeled at the physical level [10]. A topological algorithm of the same type described above to form pockets can be used to determine areas to be zoomed in. Zoomed-in areas can also be pre-established as areas in which the network is represented at the physical level throughout state estimation.

VIII. Alternative Formulations

In WLS estimators the influence of a measurement on the state estimate increases with the size of its residual, whereas nonquadratic estimators are designed to bound the influence of large residuals on state estimation (in the hope that these residuals actually correspond to bad data, but which is by no means guaranteed). In an unidimensional case, for example, the WLS estimator yields the mean value of the measurements, and assuming a finite number of measurements and that all measurements except one are correct, if the bad data tend to infinity, so will the state estimate (the average of the measured values). If instead the state estimate is defined as the median of the sample, this would not happen, since the median would not change when an arbitrary error is introduced in a single measurement; in fact, the median would remain finite when an arbitrary error is introduced in $(m-1)/2$ measurements. This is typical of ideally robust estimators which are insensitive to changes in $(m-1)/2$ measurements and can be achieved to a certain degree by changing the objective function as illustrated in Fig. 10.

A. Nonquadratic Estimators

The unconstrained state estimation problem can be formulated as follows:

$$\underset{\hat{x}}{\text{minimize}} \quad J(\mathbf{x}) = \sum_{j=1}^{m} \rho(\tilde{r}_j)$$

where $\rho(\cdot)$ is a scalar objective function and $\tilde{r}_j$ is the jth weighted residual. (In the WLS estimator the following quadratic objective function is adopted $J(\mathbf{x}) = \sum_{j=1}^{m} \tilde{r}_j^2$.)

To bound the influence of large residuals, ρ can be defined as a quadratic function only for small weighted residuals and as a function with constant or decreasing derivative for large weighted residuals [see Fig. 10(b)–(c)] [43]. Since the measurements with large weighted residuals are not known *a priori*, the process is normally initialized with a regular WLS estimator, and the measurements with large residuals have their weights gradually reduced, according to a nonquadratic objective function. The problem then becomes how to select the set of suspect measurements. If this modification on weights is based on the magnitude of weighted residuals only, good data may be rejected whereas bad data can be classified as good data. One alternative is to use the normalized residuals, which will work for cases of single and multiple nonconforming bad data.

B. Least Absolute Value Estimator (LAV)

Linear programming-based state estimation was originally proposed in [44] and further developed in [45] and [46]. The application of this method to topology estimation is discussed in [47] and [48]. The weighted least absolute value objective function [see Fig. 10(d)] is as follows:

$$J(\mathbf{x}) = \sum_{j=1}^{m} |\tilde{r}_j|$$

and the solution is given by n basic measurements which fits perfectly the state estimate. These measurements have zero residuals, whereas the remaining $m-n$ measurements, the nonbasic ones, can present nonzero residuals. In the unidimensional example discussed above, the least absolute value estimator yields a state estimate equal to one of the measured values, regardless of the magnitude of the bad data. For large and complex systems, however, things in general are more complicated due to the computational effort and to the existence of leverage points which make unwanted bad data to be selected as good data as explained in the following paragraph.

C. Leverage Points

Certain measurements called leverage points can have an abnormally high influence on state estimation. In [50], leverage points are defined as the points of a regression that

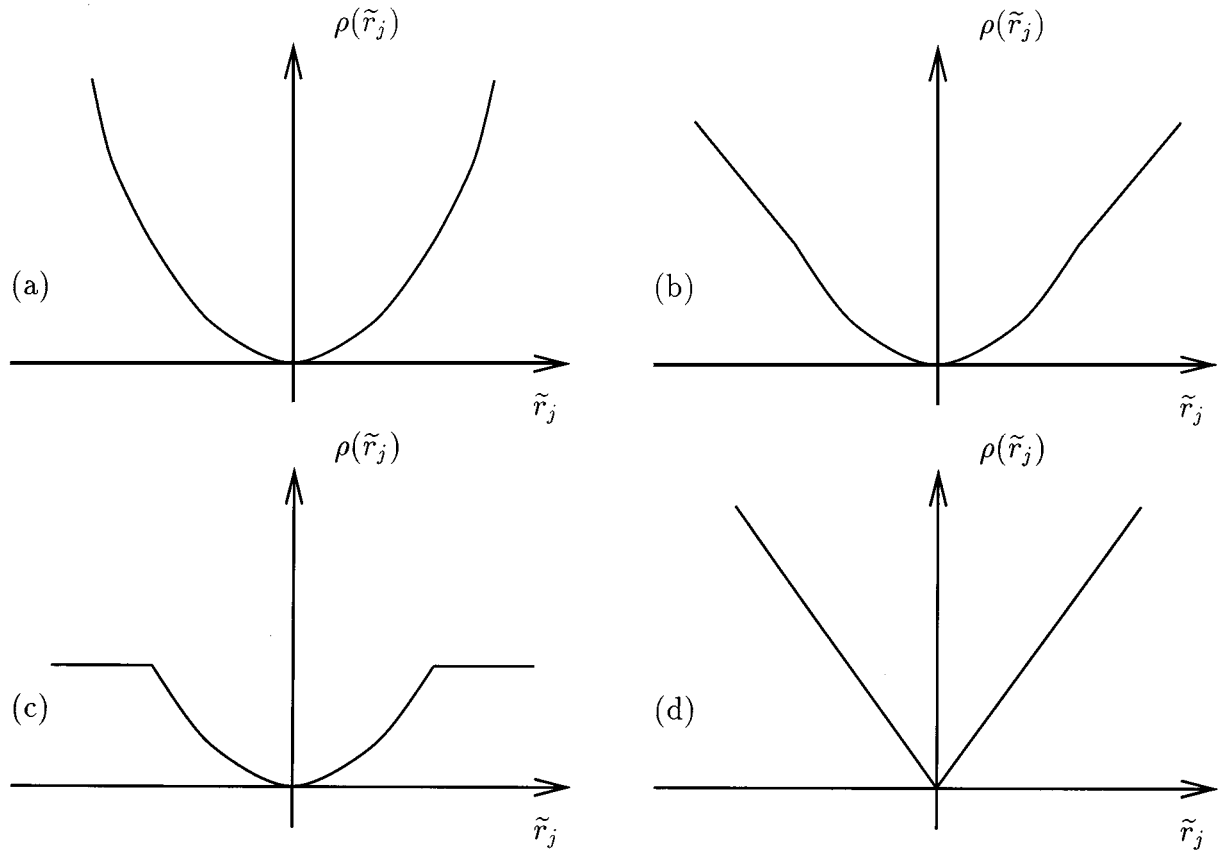


Fig. 10. Quadratic/nonquadratic objective functions. [$\tilde{r}_j = (z_j - h_j(\mathbf{x}))/\sigma_j$ is the weighted residual.]

are far away from the bulk of the data points in the factor space. A typical situation is shown in Fig. 11(b). (For a more detailed discussion, see [52]). Although the characterization of a leverage point depends only on the independent variable x, their classification as good or bad data depends also on the measured values z_i (as well as on the corresponding variances).

In power system state estimation, the factor space is the n-dimensional space spanned by the rows $\underline{\ell}'$ of the $m \times n$ Jacobian matrix (the jth row of the Jacobian matrix is $\underline{\ell}'_j = (\partial h_j/\partial \mathbf{x})'$). The leverage points of a linearized model are measurement points $(z_j, \underline{\ell}_j)$ whose vectors $\underline{\ell}_j$ define outliers in the factor space, [53].

In Fig. 11(a) the error is in the z direction (it is not a leverage point); in this case the LAV method automatically rejects the outlier, whereas the LS method gives estimates that are affected by the gross error. This property has served as a motivation for calling the LAV method robust. Notice that, when the LS method is used, the error can be correctly identified by the largest normalized residual criterion and eliminated from estimation, resulting, in this example, the same estimates yielded by the LAV estimator.

Things are harder when the outlier is in the x direction (the factor space), i.e., the outlier is a leverage point, as illustrated in Fig. 11(b). In this case, the LAV estimator will take the leverage point as a perfect measurement (a basic measurement with zero residual). Hence, we say that the LAV estimator looses its robustness when leverage points are present. This is why the identification and the removal of the effect of leverage point has become a active research area (a partial list of reference can be found in [49]). The LS estimator is also badly affected by the leverage point in this example, although the largest normalized residual criterion will correctly flag the bad data and, after its removal, the estimation will be correct.

Fig. 12 illustrates a somehow more dramatic situation originally studied in [50]. In case (a), the measurement errors are relatively small, and the five points in the initial part of the graph clearly defines a linear trend. Hence, the two leverage points on the right will appear as bad data with respect to this trend. The same is not necessarily true in case (b), in which case the trend is not as well defined as above, and where no bad data would be detect at all since estimates for a and b can be found that are compatible with all measured values (in view of the wider 3σ ranges). In case (a) both the LAV method and the LNR method will fail in identifying the bad leverage points. Both methods will classify four of the first five measured points as bad data. The combinatorial approach, however, will be able to find the desired solution, i.e., a solution that rejects only two bad data (the leverage points). The same is true regarding the LMS method discussed below.

Leverage measurements normally occur in connection with low impedance branches and nodal injection measurements, i.e.: 1) flows in low impedance branches; 2) injections at nodes adjacent to a low impedance branch;

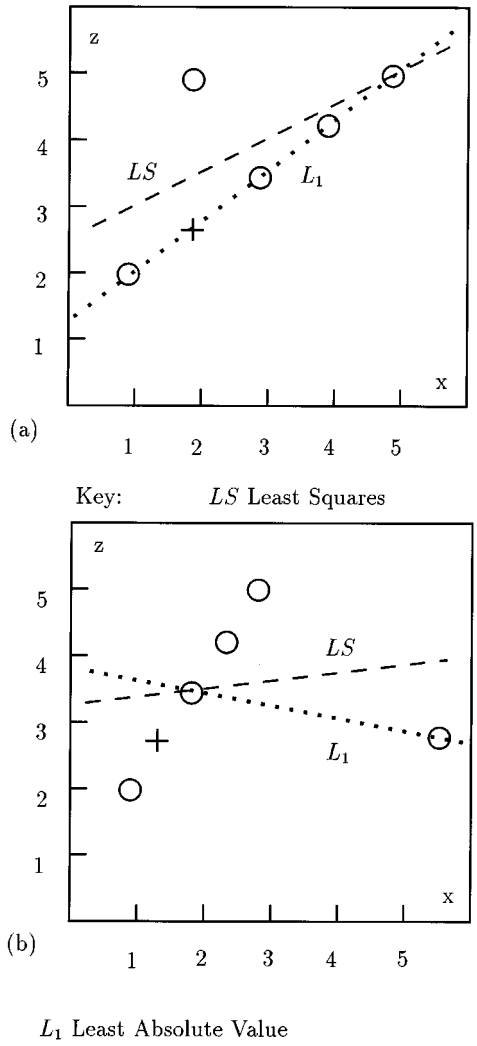


Fig. 11. Linear regression $z = a + bx$, estimates $\hat{a}$ and $\hat{b}$ are sought. (a) Outlier in the z direction. (b) Outlier in the x direction—leverage point [52].

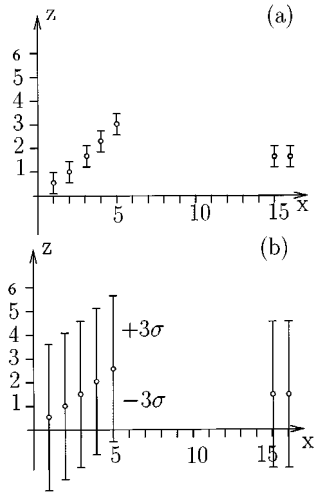


Fig. 12. Linear regression, $z = a + bx$, with two leverage points. Estimates for a and b are sought (σ is the standard deviation of the measured values z_i).

and 3) injections into nodes with a large number of incident branches. Reference [53] suggests a method for identifying leverage points based on projection statistics. This method has been used in a reweighted LS estimator by means of which bad leverage points are deweighted (Schweppe-type estimator) [54]. In order to avoid the occurence of certain types of leverage points, low impedance branches can also be modeled as zero impedance branches, with the corresponding through-flows considered as additional state variables.

D. Least Median of Squares Estimators (LMS)

In this case the objective is not based on the sum of the squared residuals or of their absolute values but on a single residual. This estimator is formulated as follows:

$$\underset{\hat{x}}{\text{minimize}} \quad \underset{j}{\text{median}}\, \tilde{r}_j^2$$

where what is sought is the state $\hat{x}$ that presents the minimum median squared weighted residual. Since nonzero residuals occur only when there are redundant measurements, i.e., $m > n$, the objective function above is used only for $n = 1$ or $n = 2$, whereas for multidimensional states $(n > 2)$ the median is given by $\mathrm{md} = [m/2] + [(n+1)/2]$, where $[]$ denotes the integer part of the argument.

The problem can then be reformulated as follows:

$$\underset{\hat{x}}{\text{Minimize}} \quad \tilde{r}_{\mathrm{md}}^2$$

where r_{md} is the weighted residual corresponding to the mdth measurement.

Reference [50] suggests drawing a series of samples with n measurements each for which the network is minimally observable. For each sample the state is calculated and the weighted residuals of the remaining $m-n$ measurements are computed. The optimal solution corresponds to the sample that minimizes the above objective function. In [51], a power system decomposition scheme is proposed to improve both the robustness and the computing time of the LMS algorithm.

IX. External System Modeling

After state estimation and bad data processing, state estimation results are used to build a power flow model for the system of interest. There are three principal approaches to do that: 1) the power-flow-based method; 2) the one-pass state estimation method; and 3) the two-pass state estimation method. (A comprehensive bibliography on external system modeling can be found in [13].)

The first method attaches the external model to the internal model by means of boundary matching injections, which are calculated by solving a power flow for the external network while treating the inner boundary buses as swing buses (with the values of the voltage magnitudes V and voltage angles θ being obtained by state estimation and used as target, or specified, values). The model thus obtained for the interconnected network correctly reproduces the conditions of the internal system, since possible external errors are absorbed by the boundary matching injections. These errors, however,

may affect both contingency analysis and optimal power flow studies, because, although the base-case model is correct, the reactions of the external model to internal changes, such as a contingency, may be incorrect. One common criticism of this method relates to the possible accumulation of external system modeling errors in boundary injections.

The second method (the one-pass method) performs a single state estimation for the system of interest, which includes both observable and unobservable parts. Power flow variables (e.g., target values for voltages and flows) and limits (e.g., MVAr limits) are treated as pseudomeasurements or as equality/inequality constraints. If a numerically robust estimator is used, these constraints can be enforced by assigning low weights to the corresponding pseudomeasurements. Alternatively, the set of pseudo measurements can be made nonredundant, in which case even the presence of errors in the data will not corrupt the states estimated from telemetry, but special care must be taken with low impedance lines, since if neither the power flow in the low impedance branch nor the injections at its terminal nodes are "metered," these variables may become numerically indeterminate, leading to abnormally large variances. One possible solution is to assign a pseudomeasurement to each low impedance branch; normally, this can be done without affecting overall measurement redundancy.

The third method (the two-pass method) runs state estimation for the observable part of the system of interest and then attaches the external model in two passes: 1) apply the power flow method to calculate branch power flows in the unobservable network, and 2) perform state estimation using the estimated states as pseudomeasurements, for the internal system, with the power flows for the external system, run to match the two parts of the network model. Zero injections for both the internal and external network are also treated as pseudomeasurements. Whenever available, external telemetry can also be included in the model. The scheduled injections at the inner boundary nodes are used as target values, and, depending on the relative weights used for external pseudomeasurements, this will disperse the modeling errors throughout the unobservable system.

X. Dynamic Parameter Estimation

This section summarizes a dynamic state/parameter estimator based on the Kalman Bucy [77] filter and which yields estimates of a state vector augmented by the parameters to be estimated. It is also assumed that an initial value of the parameter vector $\mathbf{p}$ is given, along with its respective covariance matrix $\mathbf{R_p}$. An overview of state estimation, including Kalman filters, can be found in [4]. A review of the main developments in dynamic state estimation and hierarchical state estimation was presented in [78]. More recently, a dynamic estimator with both state and parameter prediction and including second order derivatives was suggested [32].

It is assumed that a set of N observations, or scan of measurements, is available. Each scan is modeled as follows:

$$\mathbf{z}(t_i) = \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}(t_i)) + \mathbf{e}(t_i) \quad i = 0, 1, 2, \ldots, N$$

where $\mathbf{h}(\cdot)$ is a nonlinear vector function, $\mathbf{z}(t_i)$ is the measurement vector corresponding to time t_i, $\mathbf{x}(t_i)$ is the state vector at t_i, $\mathbf{e}(t_i)$ is a vector with mean of zero and variance $\mathbf{R_z}$, and $\mathbf{p}$ is an unknown constant parameter vector.

The simplified dynamic that follows has been widely used in the literature:

$$\mathbf{x}(t_i) = \mathbf{x}(t_{i-1}) + \mathbf{v}(t_i) \tag{19}$$

where $\mathbf{v}(t_i)$ has a mean of zero and a variance of

$$E\{\mathbf{v}(t_i)\mathbf{v}'(t_j)\} = \begin{cases} \mathbf{R_v}(t_i), & t_i = t_j \\ \mathbf{0}, & t_i \neq t_j \end{cases} \tag{20}$$

i.e., $\mathbf{v}(t_i)$, $i = 0, 1, 2, \ldots$ is a discrete time process with orthogonal, zero mean random variables.

The combined parameter/state estimation problem can be formulated as an unconstrained optimization problem that minimizes the following objective function:

$$\begin{aligned} J(\mathbf{x}(t_i), \mathbf{p}(t_i)) = {} & \frac{1}{2}(\bar{\mathbf{p}}(t_i) - \mathbf{p}(t_i))' \\ & \cdot \bar{\mathbf{R}}_\mathbf{p}^{-1}(\bar{\mathbf{p}}(t_i) - \mathbf{p}(t_i)) \\ & + \frac{1}{2}(\bar{\mathbf{x}}(t_i) - \mathbf{x}(t_i))'\bar{\mathbf{R}}_\mathbf{x}^{-1}(\bar{\mathbf{x}}(t_i) - \mathbf{x}(t_i)) \\ & + \frac{1}{2}(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}(t_i)))' \\ & \cdot \mathbf{R}_\mathbf{z}^{-1}(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}(t_i))). \end{aligned}$$

The normal equations for the augmented performance index can then be written as follows:

$$\begin{aligned} & \begin{pmatrix} \mathbf{H}_\mathbf{x}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{H}_\mathbf{x} + \mathbf{R}_\mathbf{x}^{-1} & \mathbf{H}_\mathbf{x}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{H}_\mathbf{p} \\ \mathbf{H}_\mathbf{p}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{H}_\mathbf{x} & \mathbf{H}_\mathbf{p}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{H}_\mathbf{p} + \mathbf{R}_\mathbf{p}^{-1} \end{pmatrix} \begin{pmatrix} \Delta\mathbf{x} \\ \Delta\mathbf{p} \end{pmatrix} \\ & = \begin{pmatrix} \mathbf{H}_\mathbf{x}' & \mathbf{0} \\ \mathbf{H}_\mathbf{p}' & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{R}_\mathbf{z}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_\mathbf{p}^{-1} \end{pmatrix} \begin{pmatrix} \Delta\mathbf{z}(t_i) \\ \mathbf{p}(t_i) - \hat{\mathbf{p}}(t_{i-1}) \end{pmatrix}. \end{aligned}$$

The covariance matrix of the augmented state estimate error, $(\mathbf{x}(t_i) - \hat{\mathbf{x}}(t_i), \mathbf{p}(t_i) - \hat{\mathbf{p}}(t_i))'$, can be decomposed as follows:

$$\hat{\mathbf{R}}(t_i) = \begin{pmatrix} \mathbf{R}_{\hat{\mathbf{x}}}(t_i) & \mathbf{R}_{\hat{\mathbf{x}},\hat{\mathbf{p}}}(t_i) \\ \mathbf{R}_{\hat{\mathbf{p}},\hat{\mathbf{x}}}(t_i) & \mathbf{R}_{\hat{\mathbf{p}}}(t_i) \end{pmatrix}.$$

The covariance matrix associated with the vector of parameter estimates, $\mathbf{R_p}(t_{i+1})$, can be obtained by applying the matrix inversion lemma to the blocked gain matrix that appears in the normal equations of the combined state/parameter estimation problem above. The result is [11]

$$\begin{aligned} \mathbf{R}_\mathbf{p}^{-1}(t_{i+1}) = {} & \mathbf{H}_\mathbf{p}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{H}_\mathbf{p} + \mathbf{R}_\mathbf{p}^{-1} \\ & - \mathbf{H}_\mathbf{p}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{H}_\mathbf{x}\left(\mathbf{H}_\mathbf{x}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{H}_\mathbf{x} + \mathbf{R}_\mathbf{x}^{-1}\right)^{-1} \\ & \cdot \mathbf{H}_\mathbf{x}'\mathbf{R}_\mathbf{z}^{-1}\mathbf{H}_\mathbf{p} \end{aligned}$$

where all matrices on the right side are computed at t_i. This expression has been used for small subnetworks without the

diagonal approximation [32]. Further approximations are normally needed, however, in dealing with larger networks. (See Remark 2.)

The gain matrix can then be written as

$$\hat{\mathbf{R}}^{-1}(t_i) = \begin{pmatrix} \mathbf{G_{xx}} & \mathbf{G_{px}} \\ \mathbf{G_{xp}} & \mathbf{G_{pp}} \end{pmatrix}$$

where

$$\mathbf{G_{xx}} = \mathbf{H}'_{\mathbf{x}}\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{H}_{\mathbf{x}} + \mathbf{R}_{\mathbf{x}}^{-1} - \sum_{j=1}^{m} \frac{r_j}{\sigma_j^2}\frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}^2}$$

$$\mathbf{G_{xp}} = \mathbf{H}'_{\mathbf{x}}\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{H}_{\mathbf{p}} - \sum_{j=1}^{m} \frac{r_j}{\sigma_j^2}\frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{x}\partial \mathbf{p}}$$

$$\mathbf{G_{px}} = \mathbf{H}'_{\mathbf{p}}\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{H}_{\mathbf{x}} - \sum_{j=1}^{m} \frac{r_j}{\sigma_j^2}\frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{p}\partial \mathbf{x}}$$

$$\mathbf{G_{pp}} = \mathbf{H}'_{\mathbf{p}}\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{H}_{\mathbf{p}} + \mathbf{R}_{\mathbf{p}}^{-1} - \sum_{j=1}^{m} \frac{r_j}{\sigma_j^2}\frac{\partial^2 h_j(\mathbf{x})}{\partial \mathbf{p}^2}.$$

Remark 1: The inclusion of the predicted values $\bar{\mathbf{x}}$ as additional pseudomeasurements helps in filtering bad analog data that may arise during parameter estimation; it also improves observability conditions in situations where the meter configuration may change during process the (e.g., due to temporary unavailability of certain measurements) [79]. This is especially important in real-time parameter estimation.

Remark 2: The method for parameter estimation by state augmentation suggested in [80] can be obtained from the above approach by considering $\mathbf{R}_{\mathbf{x}}^{-1} = \mathbf{0}$ and approximating the equation for updating the covariance matrix of the estimated parameters as follows:

$$\mathbf{R}_{\mathbf{p}}^{-1}(t_{i+1}) \simeq \mathbf{R}_{\mathbf{p}}^{-1}(t_i) + \mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_{\mathbf{z}}^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i).$$

This approximation is partly justified by the fact that when the measurement redundancy tends to infinity, the covariance matrix of the measurement estimate errors

$$\mathbf{R}_{\hat{\mathbf{z}}}^{-1} = \mathbf{H}_{\mathbf{x}}\left[\mathbf{H}'_{\mathbf{x}}\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{H}_{\mathbf{x}}\right]^{-1}\mathbf{H}'_{\mathbf{x}}$$

tends to a null matrix. Although the $\mathbf{R_p}$ matrix can always be initialized as a diagonal matrix, the updated matrix given by the formulae above (both the exact and approximate expressions) will turn the covariance matrix into a full matrix for the next iteration. The following approximation can be used in order to keep the covariance matrix of the parameter error estimates as a diagonal matrix throughout the estimation process:

$$\mathbf{R}_{\mathbf{p}}^{-1}(t_{i+1}) \simeq \mathbf{R}_{\mathbf{p}}^{-1}(t_i) + \mathbf{diag}\left(\mathbf{H}'_{\mathbf{p}}(t_i)\mathbf{R}_{\mathbf{z}}^{-1}(t_i)\mathbf{H}_{\mathbf{p}}(t_i)\right) + \mathbf{D}$$

where matrix $\mathbf{D}$ is a constant matrix used to guarantee that the covariance matrix remains positive definite.

Remark 3: The alternate estimation of states and parameters have also been suggested in literature [3]. To this end, the following performance index is then defined:

$$J(\mathbf{x}(t_i), \mathbf{p}) = \frac{1}{2}(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}))'\mathbf{R}_{\mathbf{z}}^{-1}(\mathbf{z}(t_i) - \mathbf{h}(\mathbf{x}(t_i), \mathbf{p}))$$

for each scan, considering a fixed value for $\mathbf{p}$. In this case the normal equations is the same as for a regular nonlinear state estimator, i.e.

$$\mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_{\mathbf{z}}^{-1}\mathbf{H}_{\mathbf{x}}(t_i)\,\Delta\mathbf{x}(t_i) = \mathbf{H}'_{\mathbf{x}}(t_i)\mathbf{R}_{\mathbf{z}}^{-1}\Delta\mathbf{z}(t_i).$$

This equation is used iteratively to minimize $J(\mathbf{x}(t_i), \mathbf{p})$ for each scan. Then a new parameter estimate is obtained by minimizing the following:

$$J(\mathbf{p}) = \sum_{i=1}^{N} J\left(\hat{\mathbf{x}}(t_i), \mathbf{p}\right).$$

The optimal state $\hat{\mathbf{p}}$ is then used to compute a new set of state estimates $\hat{\mathbf{x}}(t_i)$, $i = 1, 2, \ldots, N$. The minimization of $J(\mathbf{p})$ can be performed using the Kalman Bucy filter considering now $\mathbf{p}$ as a state vector.

XI. The Impact of the Changing Marketplace

The recent creation of Independent System Operators (ISO), with the need to control the grid, has increased the size of the networks that have to be modeled by state estimation. This trend compounds with to other phenomena: the spatial expansion stemming from the merger of several companies and the growing requirements for representing the network at lower voltage levels. As a result, supervised networks with tens of thousands of buses are becoming more and more common. The difficulties are not limited to network sizes, however.

Perhaps the most important issue in the new competitive environment is the way poorly estimated network models will affect the determination of prices of electricity. Metering quality and redundancy levels can vary significantly in a large network, measurement redundancy being more deficient at lower voltage levels. High redundancy levels will be necessary for adequate model building, and even more so for topology estimation, i.e., when parts of the network are represented at the physical level. It is envisaged that the range of measurement redundancy should evolve from today's 1.7–2.2 to 2.5–3.0. In addition to this improvement in the redundancy levels, the location of new meter should also take into consideration the need for estimating topology (statuses of switching devices).

Time skew among the measurements is normally present in most existing EMS systems, but its effect on state estimation is hardly noticeable in smaller networks. With larger networks, however, large pockets of data, sometimes involving an entire company's network, can be skewed by significant amounts from the rest of the data. This can affect both state estimation convergence and bad data processing; part of the skewed data will appear in the boundary nodes as multiple bad data (they contain errors that are conforming). Even though the dynamics of the system could be taken into consideration to correct the effect of time skew without much additional effort (e.g., using time-tagged measurements), when the protection system is activated or when the system is ramping up or down at a rapid pace in the area affected by the time skew, it is unlikely to exist a satisfactory solution to this problem in view of the scan rates that are currently used by the industry (1–2 s).

XII. Conclusion

State estimation is a key function in determining real-time models for interconnected networks as seen from EMS. In this environment, a real-time model is extracted at intervals from snapshots of real-time measurements. It is generally agreed that the emerging energy markets will demand network models more accurate and reliable than ever. This can only be achieved with state estimators that can reliably deal with both state, topology (status), and parameter estimation. With that in mind, this paper has reviewed the principal developments in state estimation and related areas such as observability analysis, bad data processing, network topology processing, topology estimation, and parameter estimation.

Acknowledgment

The author gratefully acknowledges the reviewers of this paper for their many useful suggestions and corrections.

References

[1] Z. Alaywan and J. Allen, "California electric restructuring; A broad description of the development of the California ISO," *IEEE Trans. Power Syst.*, vol. 13, pp. 1445–1452, Nov. 1998.

[2] D. Shirmohammadi, B. Wollenberg, B. Vodjani, P. Sandrin, M. V. F. Pereira, F. Schneider, and B. Stott, "Transmission dispatch and congestion management in the emerging energy market structures," *IEEE Trans. Power Syst.*, vol. 13, pp. 1466–1474, Nov. 1998.

[3] F. C. Schweppe and E. J. Handschin, "Static state estimation in electric power systems," *Proc. IEEE*, vol. 62, pp. 972–982, July 1974.

[4] E. Handschin, Ed., "Real time data processing using state estimation in electric power systems," in *Real Time Control of Electric Power Systems*. London, U.K.: Elsevier, 1972, pp. 29–57.

[5] A. Bose and K. A. Clements, "Real-time modeling of power networks," *Proc. IEEE*, vol. 75, pp. 1607–1622, Dec. 1987.

[6] F. F. Wu, "Power system state estimation: A survey," *Int. J. Elect. Power Eng. Syst.*, vol. 12, pp. 80–87, Jan. 1990.

[7] L. Holten, A. Gjelsvik, S. Aam, F. F. Wu, and W. H. E. Liu, "Comparison of different methods for state estimation," *IEEE Trans. Power Syst.*, vol. 3, pp. 1798–1806, Nov. 1988.

[8] H.-J. Koglin, Th. Neisius, G. Beißler, and K. D. Schmitt, "Bad data detection and identification," *Int. J. Elect. Power*, vol. 12, no. 2, pp. 94–103, Apr. 1990.

[9] M. B. Coutto, A. M. L. Silva, and D. M. Falcão, "Bibliography on power system state estimation (1968–1989)," *IEEE Trans. Power Syst.*, vol. 7, pp. 950–961, Aug. 1990.

[10] O. Alsaç, N. Vempati, B. Stott, and A. Monticelli, "Generalized state estimation," *IEEE Trans. Power Syst.*, vol. 13, pp. 1069–1075, Aug. 1998.

[11] A. Monticelli, *State Estimation in Electric Power Systems—A Generalized Approach*. Norwell, MA: Kluwer, 1999.

[12] F. F. Wu and A. Monticelli, "A critical review on external network modeling for on-line security analysis," *Int. J. Elect. Power Eng. Syst.*, vol. 5, pp. 222–235, Oct. 1983.

[13] K. Kato, Chairman, "External network modeling—Recent practical experience," *IEEE Trans. Power Syst.*, vol. 9, pp. 216–225, Nov. 1997.

[14] A. Monticelli and A. Garcia, "Modeling zero-impedance branches in power-system state estimation," *IEEE Trans. Power Syst.*, vol. 6, pp. 1561–1570, Nov. 1991.

[15] T. E. Dy Liacco, K. A. Ramarao, and A. W. Weiner, "Network status analysis for real-time systems," in *Proc. 8th PICA Conf.*, June 1973, pp. 356–362.

[16] A. M. Sasson, S. T. Ehrmann, P. Lynch, and L. S. van Slyck, "Automatic power system network topology determination," *IEEE Trans. Power App. Syst.*, vol. 92, pp. 610–618, Mar./Apr. 1973.

[17] M. Bertran and X. Corbella, "On the validation and analysis of a new method for power network connectivity determination," *IEEE Trans. Power App. Syst.*, vol. 101, pp. 316–324, Feb. 1982.

[18] N. Singh and H. Glavitsch, "Detection and identification of topological errors in on-line power system analysis," *IEEE Trans. Power Syst.*, vol. 6, pp. 324–331, Feb. 1991.

[19] P. Bonanomi and G. Gramberg, "Power system data validation and state calculation by network search techniques," *IEEE Trans. Power App. Syst.*, vol. 102, pp. 238–249, Jan. 1983.

[20] M. Prais and A. Bose, "A topology processor that tracks network modifications over time," *IEEE Trans. Power Syst.*, vol. 3, pp. 992–998, Aug. 1988.

[21] M. R. Irving and M. J. H. Sterling, "Substation data validation," *Proc. Inst. Elect. Eng.*, pt. C, vol. 129, no. 3, pp. 119–122, May 1982.

[22] B. C. Clewer, M. R. Irving, and M. J. H. Sterling, "Topologically independent state estimation," in *Proc. IFAC Symp. Power Systems Modelling and Control Applications*, Brussels, Belgium, Sept. 1988, pp. 17.1.1.–17.1.13.

[23] A. Monticelli, "Modeling circuit breakers in weighted least squares state estimation," *IEEE Trans. Power Syst.*, vol. 8, pp. 1143–1149, Aug. 1993.

[24] F. C. Schweppe, J. Wildes, and D. Rom, "Power system static state estimation: Parts I, II, and III," *IEEE Trans. Power App. Syst.*, vol. 89, pp. 120–135, Jan. 1970.

[25] A. G. Phadke, J. S. Thorp, and K. J. Karimi, "State estimation with phasor measurements," *IEEE Trans. Power Syst.*, vol. 1, pp. 233–241, Feb. 1986.

[26] F. C. Aschmoneit, N. M. Peterson, and E. C. Adrian, "State estimation with equality constraints," in *Proc. PICA'77*, May 1977, pp. 427–430.

[27] A. Simões-Costa and V. H. Quintana, "An orthogonal row processing algorithm for power system sequential state estimation," *IEEE Trans. Power App. Syst.*, vol. 100, pp. 3791–3800, Aug. 1981.

[28] Contribution to power system state estimation and transient stability analysis, Contract DOE/ET/29362-1, Feb. 1984. Prepared by ESCA Corporation for the U.S. Department of Energy.

[29] N. Vempati, I. W. Slutsker, and W. F. Tinney, "Enhancements to givens rotations for power system state estimation," *IEEE Trans. Power Syst.*, vol. 6, pp. 842–849, May 1991.

[30] N. Vempati, I. W. Slutsker, and W. F. Tinney, "Orthogonal sparse vector methods," *IEEE Trans. Power Syst.*, vol. 7, pp. 926–932, May 1992.

[31] K. A. Clements, P. W. Davis, and K. D. Frey, "Treatment of inequality constraints in power-system state estimation," *IEEE Trans. Power Syst.*, vol. 10, pp. 567–573, May 1995.

[32] I. W. Slutsker, S. Mokhtari, and K. A. Clements, "Real time recursive parameter estimation in energy management systems," *IEEE Trans. Power Syst.*, vol. 11, pp. 1393–1399, Aug. 1998.

[33] R. A. M. Van Amerongen, "On convergence analysis and convergence enhancement of power-system least-squares state estimators," *IEEE Trans. Power Syst.*, vol. 10, pp. 2038–2044, Nov. 1995.
[34] A. Monticelli, C. A. F. Murari, and F. F. Wu, "A hybrid state estimator: Solving normal equations by orthogonal transformations," *IEEE Trans. Power Syst.*, vol. 105, pp. 3460–3468, Dec. 1985.
[35] Å. Björck, "Stability analysis of the method of seminormal equations for linear least squares problems," *Linear Algebra and Its Applicat.*, vol. 88–89, pp. 31–48, 1987.
[36] R. A. M. Van Amerongen, "On the exact incorporation of virtual measurements in orthogonal-transformation based state estimation procedures," *Int. J. Elect. Power*, vol. 13, no. 3, pp. 167–174, June 1991.
[37] F. L. Alvarado and W. F. Tinney, "State estimation augmented blocked matrices," *IEEE Trans. Power Syst.*, vol. 5, no. 3, pp. 911–921, Aug. 1990.
[38] R. R. Nucera and M. L. Gilles, "A blocked sparse-matrix formulation for the solution of equality-constrained state estimation," *IEEE Trans. Power Syst.*, vol. 6, pp. 214–224, Feb. 1991.
[39] A. Gjelsvik, S. Aam, and L. Holten, "Hachtel's augmented matrix method: A rapid method for improving numerical stability in power system state estimation," *IEEE Trans. Power App. Syst.*, vol. 104, pp. 2987–2993, Nov. 1985.
[40] P. A. Machado, G. P. Azevedo, and A. Monticelli, "A mixed pivoting approach to the factorization of indefinite matrices in power system state estimation," *IEEE Trans. Power Syst.*, vol. 6, pp. 676–682, May 1991.
[41] F. F. Wu, E. H. E. Liu, and S. M. Lun, "Observability analysis and bad data-processing for state estimation with equality constraints," *IEEE Trans. Power Syst.*, vol. 3, pp. 541–578, May 1988.
[42] R. R. Nucera, V. Brandwajn, and M. L. Gilles, "Observability analysis and bad data—Analysis using augmented blocked matrices," *IEEE Trans. Power Syst.*, vol. 8, pp. 426–433, May 1993.
[43] H. M. Merril and F. C. Schweppe, "Bad data suppression in power system state estimation," *IEEE Trans. Power App. Syst.*, vol. 90, pp. 2718–2725, Nov./Dec. 1971.
[44] M. R. Irving, R. C. Owen, and M. Sterling, "Power system state estimation using linear programming," *Proc. Inst Elect. Eng.*, vol. 125, pp. 978–885, Sept. 1978.
[45] W. W. Kotiuga and M. Vidyasagar, "Bad data rejection properties of weighted least absolute value techniques applied to static state estimation," *IEEE Trans. Power App. Syst.*, vol. 101, pp. 844–853, Apr. 1982.
[46] D. Falcão and S. M. Assis, "Linear-programming state estimation—Error analysis and gross error identification," *IEEE Trans. Power Syst.*, vol. 3, pp. 809–815, Aug. 1988.
[47] H. Singh and F. L. Alvarado, "Network topology determination using least absolute value state estimation," *IEEE Trans. Power Syst.*, vol. 10, pp. 1159–1165, Aug. 1995.
[48] A. Abur, H. Kim, and M. K. Celik, "Identifying the unknown circuit-breaker statuses in power network," *IEEE Trans. Power Syst.*, vol. 10, pp. 2029–2037, Nov. 1995.
[49] A. Abur, F. H. Magnano, and F. L. Alvarado, "Elimination of leverage measurements via matrix stretching," *Int. J. ElecT. Power and En. Syst.*, vol. 19, pp. 557–562, Nov. 1997.
[50] L. Mili, V. Phaniraj, and P. J. Rousseeuw, "Least median of squares estimation in power-systems," *IEEE Trans. Power App. Syst.*, vol. 6, pp. 511–523, May 1991.
[51] M. G. Cheniae, M. G. Mili, and P. J. Rousseeuw, "Identification of multiple interacting bad data via power system decomposition," *IEEE Trans. Power App. Syst.*, vol. 11, pp. 1555–1563, Aug. 1996.
[52] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York: Wiley, 1993.
[53] L. Mili, M. G. Cheniae, N. S. Vichare, and P. J. Rousseeuw, "Robust state estimation based on projection statistics," *IEEE Trans. Power App. Syst.*, vol. 11, pp. 1118–1127, May 1996.
[54] R. C. Pires, A. Simões-Costa, and L. Mili, "Iteratively reweighted least squares state estimation through Givens rotations," in *Proc. 1998 IEEE/PES Meeting*, Paper PE-406-Power Syst.-0-06, San Diego, CA.
[55] K. A. Clements and B. F. Wollenberg, "An algorithm for observability determination in power system state estimation," in *Proc. IEEE/PES Summer Meeting*, Paper A75 447-3, San Francisco, CA, July 1975.
[56] G. R. Krumpholz, K. A. Clements, and P. W. Davis, "Power system observability: A practical algorithm using network topology," *IEEE Trans. Power App. Syst.*, vol. 99, pp. 1534–1542, July/Aug. 1980.
[57] J. J. Allemong, G. D. Irisarri, and A. M. Sasson, "An examination of solvability for state estimation algorithms," in *Proc. IEEE PES Winter Meeting*, Paper A80-008-3, N.Y., 1980.
[58] A. Monticelli and F. F. Wu, "Network observability: Theory," *IEEE Trans. Power App. Syst.*, vol. 104, pp. 1035–1041, May 1985.
[59] A. Monticelli and F. F. Wu, "Network observability: Identification of observable islands and measurement placement," *IEEE Trans. Power App. Syst.*, vol. 104, pp. 1042–1048, May 1985.
[60] A. Monticelli, "The impact of modeling short circuit branches in state estimation," *IEEE Trans. Power Syst.*, vol. 8, pp. 364–370, Feb. 1993.
[61] K. A. Clements, "Observability methods and optimal meter placement," *Int. J. Elect. Power Eng. Syst.*, vol. 12, no. 2, pp. 88–93, Apr. 1990.
[62] A. Abur and A. G. Exposito, "Algorithm for determining phase-angle observability in the presence of line-current-magnitude measurements," *Inst. Elect. Eng. Proc. Trans. Dist.*, pt. D, vol. 142, pp. 453–458, Sept. 1995.
[63] A. Abur and A. G. Exposito, "Detecting multiple solutions in state estimation in the presence of current magnitude measurements," *IEEE Trans. Power Syst.*, vol. 12, pp. 370–375, Feb. 1997.
[64] A. Monticelli, A. Garcia, and I. Slutsker, "Handling discardable measurements in power system state estimation," *IEEE Trans. Power Syst.*, vol. 7, pp. 1341–1349, Aug. 1992.
[65] A. Monticelli and F. F. Wu, "Observability analysis for orthogonal transformation based state estimation," *IEEE Trans. Power Syst.*, vol. 1, pp. 201–206, Feb. 1986.
[66] G. C. Contaxis and G. N. Korres, "A reduced model for power system observability analysis and restoration," *IEEE Trans. Power Syst.*, vol. 3, no. 4, pp. 1411–1417, Nov. 1988.
[67] E. Handschin, F. C. Schweppe, J. Kohlas, and A. Fiechter, "Bad data analysis for power systems state estimation," *IEEE Trans. Power App. Syst.*, vol. 94, pp. 329–337, Mar./Apr. 1975.
[68] K. A. Clements and A. Simões-Costa, "Topology error identification using normalized Lagrange multipliers," *IEEE Trans. Power Syst.*, vol. 13, pp. 347–353, May 1998.
[69] A. Monticelli, "Testing equality constraint hypotheses in weighted least squares state estimators," in *Proc. PICA'99*, Santa Clara, CA, May 1999.
[70] H. Duran, "Surrogate measurements make faster state estimation optimal and general," in *IEEE PES Summer Meeting*, Paper A 77 598-6, Mexico City, 1977.
[71] F. Aboytes and B. J. Cory, "Identification of measurement and configuration errors in static estimation," in *Proc. 9th Power Industry Computer Application Conf.*, New Orleans, LA, June 1975, pp. 298–302.
[72] A. Garcia, A. Monticelli, and P. Abreu, "Fast decoupled state estimation and bad data processing," *IEEE Trans. Power App. Syst.*, vol. 98, pp. 1645–1652, Sept./Oct. 1979.
[73] N.-D. Xiang, S.-Y. Wang, and E.-K. Yu, "Estimation and identification of multiple bad data in power system state estimation," in *Proc. 7th Power System Computation Conf., PSCC*, Lausanne, Switzerland, July 1981.
[74] N.-D. Xiang, S.-Y. Wang, and E.-K. Yu, "A new approach for detection and identification of multiple bad data in power system state estimation," *IEEE Trans. Power App. Syst.*, vol. 101, pp. 454–462, Feb. 1982.
[75] L. Mili, Th. Van Cutsem, and M. Ribbens-Pavella, "Hypothesis testing identification: A new method for bad data analysis in power system state estimation," *IEEE Trans. Power App. Syst.*, vol. 103, pp. 2139–3252, Nov. 1984.
[76] A. Monticelli, F. F. Wu, and M. Yen, "Multiple bad data identification for state estimation by combinatorial optimization," *IEEE Trans. Power Delivery*, vol. PWRD-1, pp. 361–369, July 1986.

[77] R. S. Bucy, *Lectures on Discrete Time Filtering*. Berlin, Germany: Springer-Verlag, 1994.
[78] P. Rousseaux, Th. Van Cutsem, and T. E. Dy Liacco, "Whither dynamic state estimation," *Int. J. Elect. Power*, vol. 12, no. 2, pp. 105–116, Apr. 1990.
[79] A. M. L. Silva, M. B. Coutto, and J. M. C. Cantera, "An efficient dynamic state estimation algorithm including bad data processing," *IEEE Trans. Power Syst.*., vol. 2, pp. 1050–1058, Nov. 1987.
[80] A. S. Debs, "Estimation of steady-state power system model parameters," *IEEE Trans. Power App. Syst.*, vol. 93, pp. 1260–1268, Sept./Oct. 1974.

A. Monticelli (Fellow, IEEE) received the B.Sc. degree in electronic engineering from Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos, Brazil, in 1970, the M.S. degree from Universidade Federal da Paraíba (UFPb), Campina Grande, Brazil, in 1972, and the Ph.D. degree from Universidade de Campinas (Unicamp), Campinas, Brazil, in 1975.

From 1982 to 1985 he was with the University of California, Berkeley. From 1991 to 1992 he was with Mitsubishi Electric Corporation, Japan. Currently he is a Professor of Electrical Engineering at UNICAMP, Campinas, Brazil.

Dr. Monticelli is a member of the Brazilian Academy of Sciences.