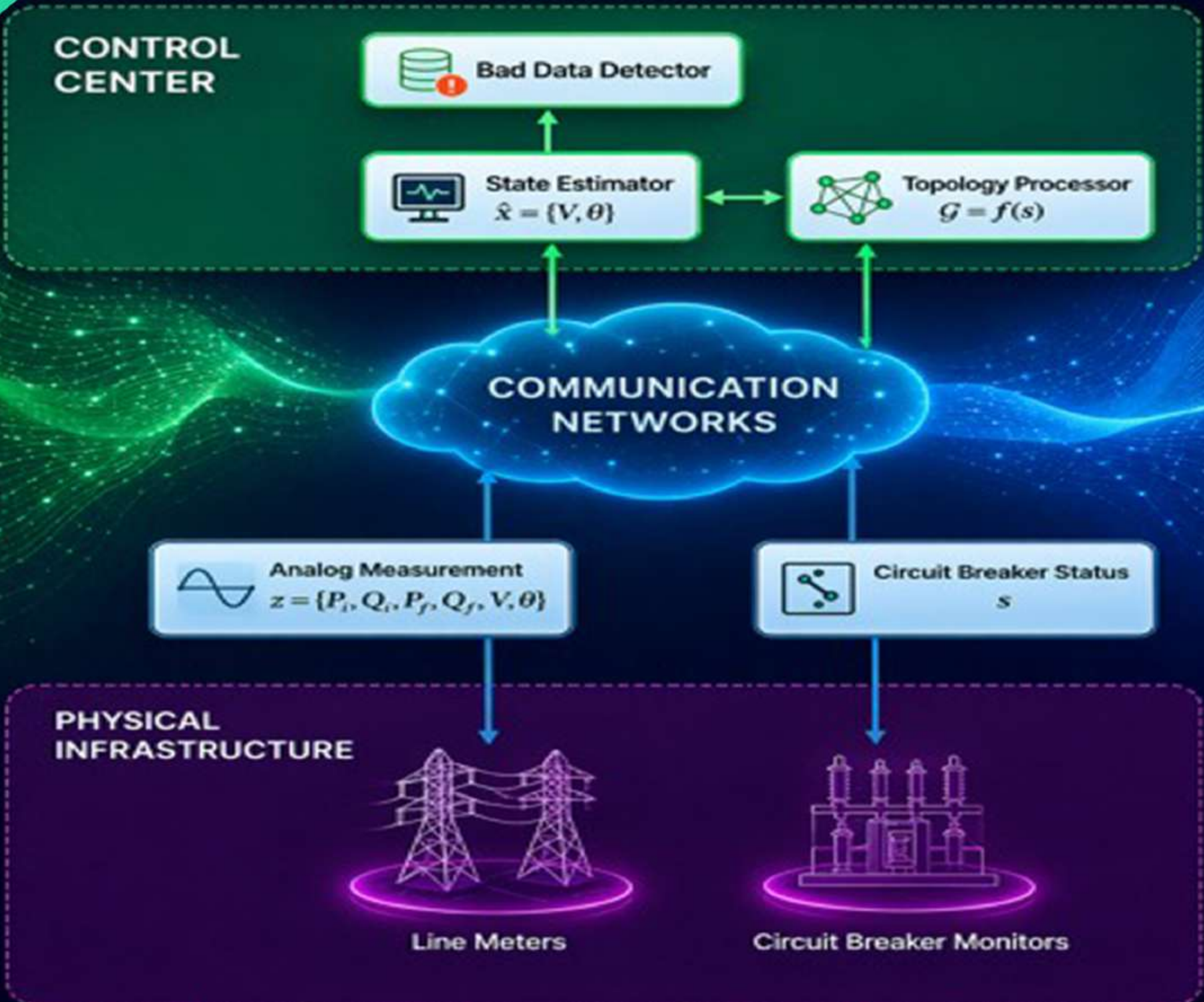
Power System State Estimation

Dr.Bishaljit Paul



Learnet Publishing
We value, we create

Power System State Estimation

Author

Dr. Bishaljit Paul

Learnet Publishing

We value, we create

Title of the Book: Power System State Estimation

Author: Dr. Bishaljit Paul

ISBN: 978-81-685619-2-2 (Digital Download and Online)

MRP: INR 799 (Hard Copy)

INR 499 (Soft Copy)

Publisher: Learnet Publishing

19/B, Kali Kumar Majumder Road, Post Office-Santoshpur Avenue, Police Station- Survey Park, Kolkata-700075, West Bengal

Email: learnetpublishing@gmail.com

info@learnetpub.co.in

Website: www.learnetpub.co.in

www.jctmg.in

Imprint: Learnet Publishing

Preface

The preface of "*Power System State Estimation*" introduces the fundamental significance of state estimation in modern power systems, emphasizing its role in ensuring reliable, secure, and efficient grid operations. It sets the stage by discussing the challenges posed by the increasing complexity of power systems, driven by the integration of renewable energy sources, dynamic loads, and evolving grid infrastructure. The authors highlight the need for accurate, real-time system monitoring to support decision-making in control centers, addressing issues like fault detection, load forecasting, and system optimization.

The preface also touches on the mathematical foundation of state estimation, including techniques like weighted least squares (WLS) and Kalman filtering, while acknowledging advancements in computational methods and sensor technologies. It prepares readers for a blend of theoretical concepts and practical applications, making the book relevant for students, researchers, and professionals aiming to navigate the complexities of modern power systems.

CONTENTS

1

Energy Management Systems

In any developed country, more than one-third of energy consumed is in the form of electrical energy. The modern-day power systems are responsible for generating, transmitting, and distributing electric energy. To manage such large and complex systems, computer-based monitoring and control is essential. The management system for computer control of a power system is known as a *control center* or *energy management system*.

A modern energy management system performs many functions, such as automatic generation control, system security functions, and economic scheduling [1–3]. The functional diagram of a modern energy system is shown in Figure 1.1. As can be seen from this diagram, the functions of energy management system can be grouped into three categories:

1. Data acquisition and processing;
2. Energy/economy functions;
3. Security monitoring and control.

In this book we are mainly concerned with the role of state estimation in power system security monitoring and control.

1.1 Real-Time Control of a Power System

Modern day power systems are divided into various areas. Each of these areas is generally interconnected to its neighboring areas. The electric power systems interconnect because the interconnected systems are more reliable.

security may be explained as the probability of the system operating point remaining within given acceptable ranges under conditions of contingencies. It is a function of time and of the robustness of the system with respect to imminent disturbances. The notion of security is the basis of all real-time monitoring and control in today's power systems. The working definition of security is in terms of the system state. The system state is a compact description used to summarize key information about the system; once the system state is known, it can be used to express any variables of interest.

A state in terms of power system security is classified as being one of three or four possible conditions: emergency, restorative, or normal. A *normal* state is classified when all operating constraints and load constraints are satisfied. An *emergency* state is one in which one or more of the physical operating limits are violated (e.g., line overloads, over/under voltages, over/under frequency). A *restorative* state is one where one or more of the loads are not met (partial or total blackout), but the *partial* system is operating in a normal state. The fourth state of the system is a condition when the power system loses stability and results in total blackout. The notion of security is defined with respect to a set of *credible* contingencies. A normal state is *secure* if all postulated contingencies result in secure *normal* operations; if a disturbance transitions the state into emergency, the state is *insecure* with respect to that contingency. This definition is well aligned with the intuitive notion that a secure power system is one that has low probability of blackout or equipment damage. The principal role of power system control is to maintain a secure system state (i.e., to prevent the system state from transitioning from secure to emergency over the widest range of operating conditions).

As the demands for reliable electric power became greater, and as labor became a more significant part of the cost of providing electric power, technologies known as supervisory control and data acquisition (SCADA) systems were developed that would allow remote monitoring and even control of key system parameters. SCADA systems began to reduce and even eliminate the need for personnel to be on hand at substations.

Early SCADA systems provided remote indication and control of substation parameters using technology borrowed from automatic telephone switching systems. Data rates on these early systems were slow—data was sent in the same manner as rotary-dial telephone commands, at 10 bits per second—so only a limited amount of information could be passed using this technology. Presently a full-fledged SCADA system is made up of signal hardware for input/output, networks, control equipment, user interface (sometimes called the human-machine interface or HMI), communication equipment, and the software to go with it all. The central system of SCADA

- Short medium and long term load forecasting
- System planning;
- Unit commitment and maintenance scheduling;
- Security monitoring;
- State estimation;
- Economic dispatch;
- Load frequency control.

1.3 Security Analysis and Monitoring

The most important concern for any electric utility is the security and stability of the power system. Power system security is the ability of the power system to withstand disturbances without interrupting the power supply and compromising the quality. In security applications the disturbances are more commonly called *contingencies.* The security assessment study is performed to check the vulnerability of the system against postulated contingencies on a real- or near real-time basis. A typical power system is never in steady state in the true sense because loads and generation patterns are continuously changing. Apart from this, the power system may be subjected to major disturbances such as equipment outage, loss of transmission line, or sudden change in large load. If any contingency occurs that results in violation of the tolerance limit of any equipment, it is automatically switched out of the system by protection devices. This event may be followed by a series of further actions that may switch other equipment out. Sometimes it may result in a cascading effect and ultimately the collapse of the system, resulting in a *blackout.*

Dy Liacco in 1967 identified that a power system may be operating in three possible states. These states are defined as the normal state, emergency state, and restorative state, as described earlier. In the emergency state, it is important to relieve the stress on the components that are working beyond their normal limits. Here the economics of operation is not relevant. In the restorative state, some portion of the system has lost power and the main objective is to restore the power as quickly as possible and bring the system back to the normal state.

There are three major components of power system security assessment [5]:

1. System monitoring;

2

Power Flow Equations

2.1 Power System Representation

Power flow problems, also known as load flow, are the heart of most power system planning studies. In order to study the operational features and electrical performance, the power system is assumed to be working in a symmetrical steady state. The power flow program computes the voltage magnitude and angle at each bus in a power system, under the balanced three-phase steady state condition. Three-phase load flows are used to analyze unbalanced three-phase systems (distribution systems). However, in this chapter the power system will be assumed to be working in a balanced steady state condition. Thus only the per-phase model is considered. From these values the real and reactive power flow in the system can be calculated. Before power flow equations can be formed and their solutions obtained, it is important to derive the mathematical models of the various components of the power system.

2.1.1 Transmission Lines

Transmission lines in a power system are classified according to their lengths as short, medium, or long lines. In general the short transmission lines have lengths less than 80 km and are represented by series impedance only. The typical length of medium transmission lines is from 80 to 250 km and are represented by lumped parameters in the form of nominal π network. Lengths of more than 250 km are classified as long transmission lines; the line parameters of these lines are assumed to be distributed uniformly throughout is $y = j\,4.674 \times 10^{-6}$ siemens per phase per km. Determine the equivalent π model.

Solution

$$z = 0.0165 + j0.3306 = 0.3310 < 87.14$$

$$Z_c = \sqrt{\frac{z}{y}} = \sqrt{\frac{0.3310 < 87.10}{4.674 \times 10^{-6} < 90}} \quad 266.1 < -1.43$$

$$\gamma = \sqrt{yz} = \sqrt{(0.3310 < 87.1)4.674 \times 10^{-6} < 90} = \sqrt{1.547 \times 10^{-6}} < 177.14$$

$$\gamma l = \sqrt{1.547 \times 10^{-6}} < 177.14 \times 300$$

$$= 0.00931 + j0.3730 = 03731 < 88.57$$

$$\cosh(\gamma l) = 0.9313 + j0.0034 = 0.9313 < 0.209$$

$$\sinh(\gamma l) = 0.0087 + j0.3664 = 0.3645 < 88.63$$

$$Z' = Z_c \sinh(\gamma l) = 266.1 < -1.43 \times 0.345 < 88.63 = 91.80 < 87.17$$

$$\frac{Y'}{2} = \frac{\tanh\left(\gamma l/2\right)}{Z_c} = \frac{\cosh(\gamma l) - 1}{(\gamma l/2 \sinh(\gamma l) Z_c} = 7.095 \times 10^{-4} < 89.97$$

The equivalent model is shown in Figure 2.1.

2.1.2 Power Transformer

In a power system, three-phase power is transmitted over transmission lines and three-phase transformers are used to convert voltage from one level to another. The primary and secondary windings of three-phase transformers can be connected in either wye (Y) or delta (Δ) configurations. This results in four possible configurations: *Y-Y*, Y-Δ, Δ-Δ, Δ-Y as shown in Figure 2.2. The three-phase transformers are represented by their per-phase model. As well, in power system quantities such as voltage, current, power, and impedance are often represented in per-unit or percentage of specified base values. One advantage of using per-unit values is that by properly specifying base quantities, the transformer equivalent circuit can be simplified. Since the transformers are having different values of voltages and currents at the primary and secondary sides, by using per-unit values the voltages, currents, and impedances do not change when referred from one side of a transformer to the other side. This is a significant advantage in a power system where hundreds of transformers are involved. The per-unit quantities are calculated as follows:

of a transformer is selected to be the same as the ratio of the transformer voltage ratings. In this way the per-unit impedance remains unchanged when referred from one side of transformer to the other side.

For the three-phase system instead of obtaining the per-phase values, the per a unit values can be obtained directly by using three-phase base quantities.

Selecting three-phase MVA as base = MVA_{base}, line to line base voltage as KV_{base}.

For a star connection the base current is

$$I_{base} = 1000 \times \frac{MVA_{base}}{\sqrt{3} KV_{base}} A$$

and

$$\text{Base Impedance } Z_{base} = \frac{(KV_{base})^2}{MVA_{base}}$$

The transformers in power flow problems are represented by π network to make them compatible with transmission line models. Since the core losses and magnetization current for power transformers are of the order of 1% of maximum ratings, the shunt branch of transformer equivalent circuit is represented by only a leakage reactance and resistance representing winding and core losses for power flow calculations.

2.1.2.1 Fixed Tap Setting Transformer

Transformers with provision of tap changing are used to regulate the voltages at the bus. A transformer with a fixed tap setting is represented by its impedance Z_{ij} or admittance Y_{ij} in series with ideal transformer, if the transformer is connected at bus i in line i–j (Figure 2.3).

If the turn ratio of ideal transformer is a:

$$I_i = \left(V_i - aV_j\right)\frac{Y_{ij}}{a^2} = -I_j \tag{2.6}$$

The equivalent circuit of this transformer is shown in Figure 2.4.

3

Weighted Least Square Estimation

3.1 Introduction

A power system essentially consists of generation, transmission, and distribution systems. A transmission system contains a large number of substations that are interconnected by transmission lines, transformers, and other switching devices for system control and protection. The efficient and optimum economic operation and planning along with security of electric power systems have always occupied an important position in the power industry. In order to achieve these objectives, it is essential for power engineers to accurately monitor the power system operating states. An essential tool for real time monitoring of the power system is state estimation (SE). It determines the best estimates of the actual power system state based on available supervisory control and data acquisition (SCADA) measurements, power system model and other data. The weighted least square (WLS) estimation method is the most commonly used technique in state estimation of power system.

The idea of least square estimation has been used since the early part of the nineteenth century. The linear least square problem deals with an overdetermined system of linear equations (a system with more known equations than unknown). In a power system, the state variables are voltage magnitudes and relative phase angles at the system nodes. Static state estimation refers to the procedure of obtaining voltage magnitude and phase angles at all the nodes of a power system. The earliest application of state estimation in power system was given by Schweppe, et al. [1, 2] in the 1970s. Simultaneous measurements of data are required at all the nodes, and collection of this data at can then be written in terms of these unknown parameters. This function is called the likelihood function. The estimation of the state is selected based on the method that maximizes this probability. The likelihood function will be maximum when the unknown parameters are closest to their actual values. Some of the measured quantities as obtained from various measuring devices are assumed to have errors. Thus the measured quantity differs from its actual value by an unknown random error.

$$z_{meas} = z_{act} + \eta \tag{3.1}$$

If the network is in steady state and the measurements performed a large number of times, the error would average to zero and $z_{meas} = z_{act}$. The measurement errors are normally assumed to have a Gaussian distribution. The parameters for Gaussian distribution are the mean μ and its variance σ^2. The problem of maximum likelihood estimation is therefore solved for these parameters. The probability density function of number z is assumed to have normal Gaussian distribution.

The Gaussian probability density function for the random number z is given by

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp - {(z-\mu)^2}/{2\sigma^2} \tag{3.2}$$

Where σ is called the standard deviation, σ^2 is called the variance of the random number, and μ is mean or expected value of z. The f(z) describes the behavior of the random number. If $(z-\mu)/\sigma = \eta$, then the plot of f(η) versus σ, which is known as Gaussian probability density function and is shown in Figure 3.1. The value of σ gives an indication of random measurement error. If σ is large the measurement is inaccurate (poor measuring instrument), whereas a small value of σ indicates the good quality measuring device. The normal distribution is commonly used for modeling measurement errors as these errors are caused by many factors.

3.3.1 Likelihood Function

The likelihood function for a sample of m independently and identically distributed observations can be obtained by finding the joint probability density function (pdf). If there are m independent measurements and each has the same Gaussian pdf, and each measurement is assumed to be independent of

$$\sum_{i=1}^{m} \log f(z_i) = \frac{1}{2}\sum_{i=1}^{m}\left(\frac{z_i - \mu_i}{\sigma_i}\right)^2 - \frac{m}{2}\log 2\pi - \sum_{i=1}^{m}\sigma_i \tag{3.5}$$

The maximum likelihood estimate is obtained by finding the maximum of log $f_m(z)$, which can also be obtained by finding the minimum of

$$\min J(z) = \sum_{i=1}^{m}\left(\frac{z_i - \mu_i}{\sigma_i}\right)^2 \tag{3.6}$$

Equation (3.6) is known as *weighted least squares* estimator. It is equivalent to a maximum likelihood estimator if the measurement errors are represented by random numbers having a normal distribution.

The minimization problem can also be written in the form of measurement residuals. If r_i is residual of measurement i, then

$$r_i = z_i - \mu_i \tag{3.7}$$

Where the mean μ_i, which is the expected value of z_i, can be expressed as $h_i(x)$, a nonlinear function relating the system state vector **x** to the ith measurement. The square of residual r_i weighted by W_i is equal to inverse of error variance σ^2. Hence the minimization problem of (3.7) can also be written as

$$\text{minimize} \quad \sum_{i=1}^{m} W_i r_i^2 \tag{3.8}$$

subject to

$$z_i = h_i(x) + r_i \tag{3.9}$$

3.4 Matrix Formulation and Measurement Model

If we have a sample of m measurements given by vector **z**, and n states (voltage magnitude and angle), $n < m$, and e is m dimensional vector of measurement errors. Here h is the m dimensional nonlinear function vector relating

4

Network Observability and Pseudomeasurements

4.1 Network Graphs and Matrices

The objective of a state estimation is to obtain a computer model that accurately represents the current conditions in a power system. It is a data processing scheme that computes the state of a system from the following three pieces of information:

1. Measurement of system variables;
2. Mathematical model of the system (includes the system topology);
3. Prior knowledge of system inputs and outputs known as pseudomeasurements.

When set of available measurements are sufficient to calculate the state vector of the system, it is termed as an observable system [1–4]. For real-time monitoring of a power system, an observability test should be carried out prior to state estimation. If the system is observable the state estimation may be carried out straight away. However, in an unobservable system, there may be observable islands and unobservable regions within the network. The states in unobservable regions of the network are estimated by adding additional measurements using pseudomeasurements. Pseudomeasurements are not obtained from meters but are typically calculated using historical data or umns is known as *node to branch* incidence *matrix*. The matrix elements are given by:

$A_{ij} = 1$, if the *j*th element is incident to but directed away from the node *i*

$A_{ij} = -1$ if the *j*th element is incident to and directed towards node *i*

$A_{ij} = 0$, if the *j*th element is not incident to node *i*

- *Reduced node to branch incidence matrix:* Any node of the network graph can be selected as reference node and the matrix obtained by deleting the row corresponding to reference node is known as *reduced node to bus incidence matrix* (A_r) for a power system network. Here the nodes are termed as buses and elements as branches. A loop-free subgraph F of G is called a forest.

Example 4.1

A single diagram of a power system is shown in Figure 4.1. The graph of the network is shown in Figure 4.2. Here, $n = 4$ and $l = 9$. The incidence matrix **A** is shown in Table 4.1.

Taking node zero as a reference node, the reduced node to branch incidence matrix is shown in Table 4.2.

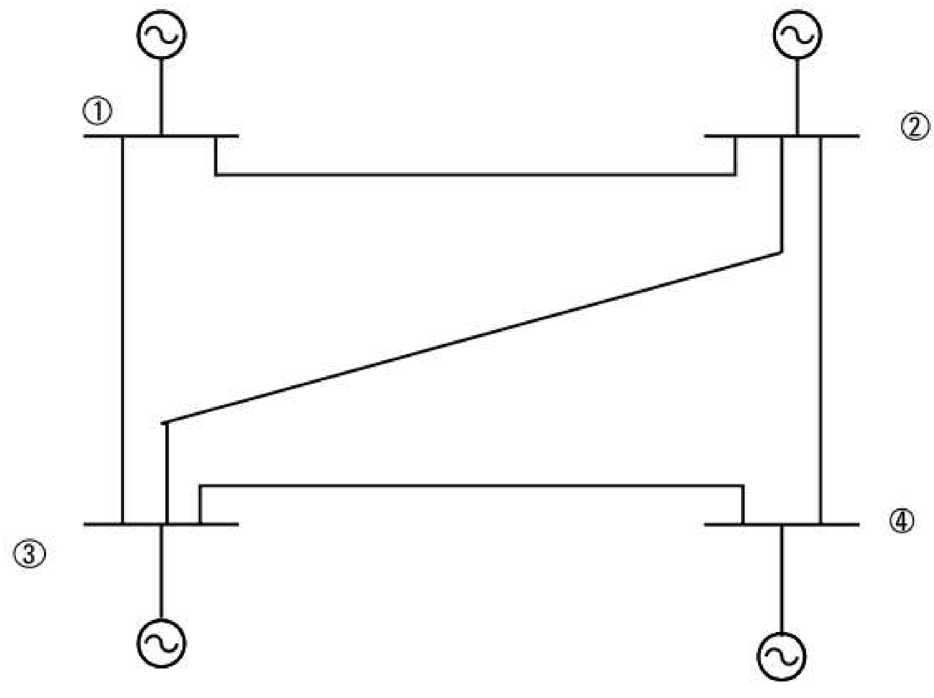


Figure 4.1 One-line diagram of a power system.

Table 4.2
Node to Branch Incidence Matrix

Elements	Nodes			
	1	2	3	4
1	1	0	0	0
2	0	1	0	0
3	0	0	1	0
4	0	0	0	1
$A_r =$ 5	1	–1	0	0
6	1	0	–1	0
7	0	1	–1	0
8	0	1	0	–1
9	0	0	1	–1

an admittance or impedance network. The impedance form of a primitive network is shown in Figure 4.3(a).

The performance equation of a primitive network in impedance form is:

$$v_{ij} + e_{ij} = Z_{ij} i_{ij} \tag{4.1}$$

where $v_{ij} = v_i - v_j$ and v_i and v_j are the node voltages.

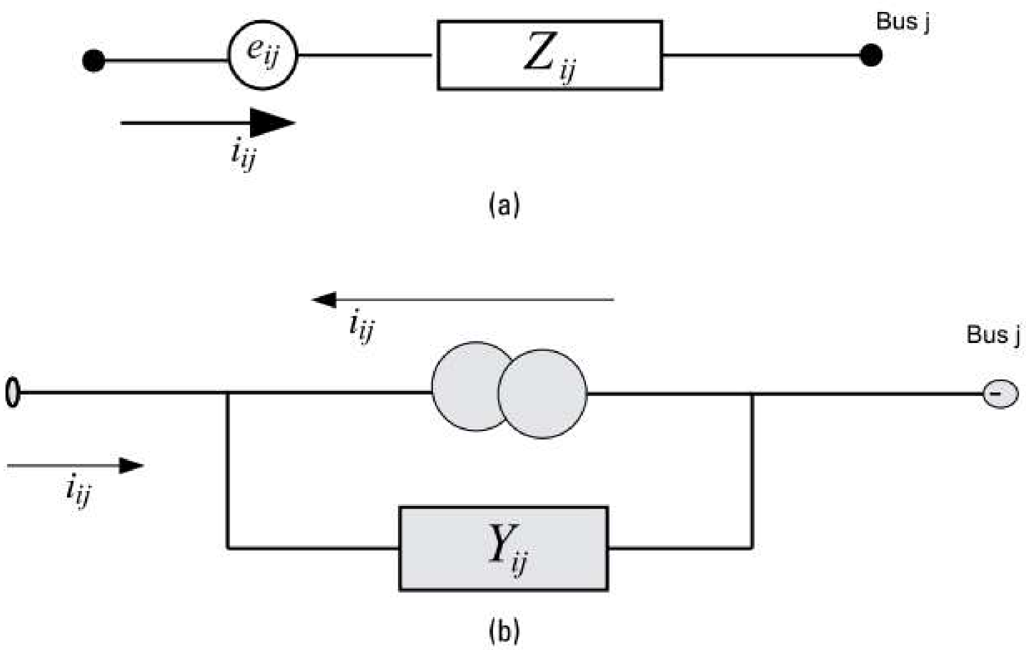


Figure 4.3 Primitive network: (a) impedance and (b) admittance.

5

Bad Data Detection

5.1 Bad Data Detection in WLS Method

State estimation algorithms apply measured data obtained from the power system to a mathematical model to provide a reliable data base for monitoring, security assessment, and control functions. If the estimates obtained are not compatible with the standard deviations of quantities estimated, the possibility is that either the measured quantities are contaminated or the model is not accurate, or both. These algorithms may encounter four types of errors: measurement errors, wrong data, structural errors in the model, and parameter uncertainty in model parameters. One of the essential features of a state estimator is to detect measurement errors and identify and rectify them. Measurement errors may occur for many reasons. Random errors in measurement exist due to the finite accuracy of meters and communication medium. These errors can be filtered out by the state estimator, provided there is sufficient redundancy in measurements. The nature of filtering action depends on the specific method of estimation. Wrong data may be available due to defects in the meters, telecommunication failures, or noise caused by interference.

It is possible to identify some of the bad data easily and eliminated before filtering it by the estimator. For example, negative voltages, readings that are many times more or less than normal values, or large difference in incoming and outgoing values of currents at the nodes may be easily identified as bad data and are not considered in estimation. However, not all types of errors can be easily identified, and state estimators must be equipped with techniques to identify all types of bad data and filter them out.

tive to bad data. Detection of bad data refers to the determination whether the measurement set contains a bad data. Estimation in the presence of bad data can be performed indirectly in one of several ways. These methods can be classified as mathematical methods and intelligent methods. Among the mathematical methods are the chi-square distribution test, the largest normalized residual test, and hypothesis testing identification. Intelligent methods either require intensive training under different conditions or methods that do not require training but require high computation. These methods are therefore not very useful. Before removal of bad data it is important to detect the existence of bad data in the measurement set. Once the presence of bad data is detected, it is identified and removed or corrected.

5.2.1 Chi-Squares Test

The chi-square test is a convenient technique to detect the presence of bad data in measurement vector z. The objective function in the WLS state estimator is

$$\min J(X) = \sum_{i=1}^{m} \left(z_i - h_i(x)\right)^2 / R_{ii} \tag{5.1}$$

$$J(X) \sum_{i=1}^{m} (e_i)^2 / R_{ii} \tag{5.2}$$

$$= \sum_{i=1}^{m} \left(\frac{e_i}{\sqrt{R_{ii}}} \right)^2 \tag{5.3}$$

where e_i is the ith measurement error, R_{ii} is the diagonal entry of measurement error covariance matrix, and m is the number of measurements.

It is of the form:

$$Y = \sum_{i=1}^{N} \varkappa_i^2 \tag{5.4}$$

which corresponds to chi-squared $\varkappa^2(N)$ distribution, with N degree of freedom. In the power system state estimation the performance index J has

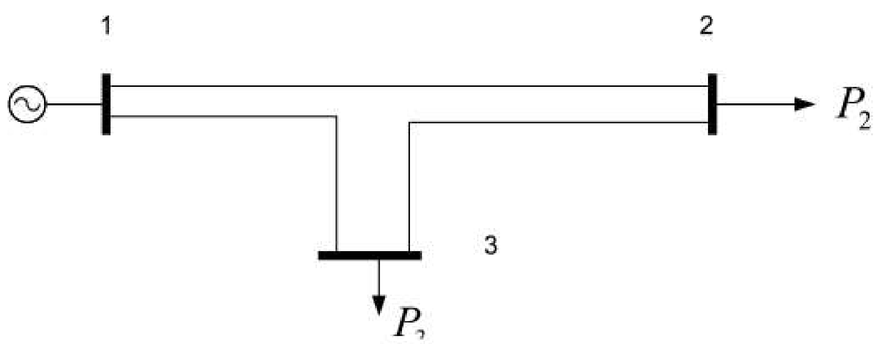


Figure 5.2 Three-bus system.

$$\sigma_{P_2}^2 = 0.004\ p.u.,\ \sigma_{P_{31}}^2 = 0.002\ p.u.,\ \sigma_{P_{13}}^2 = 0.002\ p.u.$$

Solution

From (5.2):

$$J(x) = \sum_{i=1}^{m} \left(\frac{e_i}{\sqrt{R_{ii}}} \right)^2$$

where R_{ii} is the diagonal entry of measurement error covariance matrix.

$$R_{ii} = \begin{bmatrix} \sigma_{P_2}^2 & & \\ & \sigma_{P_{13}}^2 & \\ & & \sigma_{P_{31}}^2 \end{bmatrix} = \begin{bmatrix} 0.004 & & \\ & 0.002 & \\ & & 0.002 \end{bmatrix}$$

The measurement vector estimate is given by $e_i = \left(z_i^{meas} - z_i^{est} \right)$.
The objective function is obtained as

$$J(\hat{x}) = \sum_{i=1}^{m} \frac{\left(z_i - h_i(\hat{x}) \right)^2}{\sigma_i^2}$$

The estimated values are given by

$$\hat{x} = R_{z^{est}}\ W\ z$$

Here $R_{z^{est}}$ = covariance matrix, which is obtained as follows: First, the Jacobian matrix is obtained:

6

Robust State Estimation

6.1 Basic Formulation

In a power system, estimators are used to provide a reliable estimate of unknown states in a given mathematical model from available redundant measurements. The measurement set consists of

- *Telemetered measurements*, which are online telemetered bus voltage magnitudes, active and reactive power flows, active and reactive injections (subjected to noise or error in metering, communication system, and so forth).
- *Pseudomeasurements* or zero injections, which contain no error.

The state estimation programs are also built with an ability to clean the erroneous data that may be contaminated due to various factors. If the measurements have normally distributed errors, WLS method, which is the most common method used in the industry for state estimation, provides an optimum solution to the state estimation problem. However, if some of the measurements have very large errors (statistical outliers), then the WLS method becomes highly unreliable. Since the works of Adibi and Thorne [1] and Adibi and Stoval [2], there is a growing concern about the poor quality of measurements provided by the SCADA system. The measurements in a conventional SCADA system have two main sources of errors that degrade the quality of data. The first one of these is error due to calibration of instruments, as these was formulated by Donoho and Huber [5]. It has been called the exact fit point to indicate that the majority of measurements are free of errors, and lie exactly on a hyperplane when linear regression models are considered. In statistics, linear regression is an approach to model the relationship between two variables by fitting a linear equation to observed data. One of these variables is a dependent variable and the other is an explanatory variable. If there is only one explanatory variable it is called simple regression. If more than one explanatory variable is present it is known as multiple regression.

In multiple linear regression, each of m measurements z_i (dependent variable) are represented in terms of n unknown state variables x_i and n explanatory variables $l_{i1} \ldots l_{in}$

In equation form it can be written as

$$z_i = l_{i1}x_1 + \ldots + l_{in}x_n + e_i \ldots i = 1,2,\ldots m \tag{6.1}$$

or

$$z = Hx + e \tag{6.2}$$

Here H is the $m \times n$ matrix, known as the design matrix of regression.

$$H = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ l_{21} & l_{22} & \cdots & l_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ l_{n1} & \cdots & \cdots & l_{nn} \end{bmatrix} \tag{6.3}$$

The row vector of H, $l_i^T = l_{i1} \ldots l_{in}$, defines a point in n-dimensional subspace, called the factor space of regression.

In a power system, the linear regression model is obtained with DC approximations. It is derived by linearizing the n-bus system model about the flat voltage profile and neglecting the series resistances and shunt capacitances of the branches. Two linear models as proposed in [2] can be written. One model relates the real power measurements to the (n-1) voltage phase angles θ_i. The other model relates to the reactive power measurements P_i to the n nodal voltages magnitudes. Then H is the P, θ or Q, V Jacobian submatrix, respectively.

The robustness of an estimator can be quantified by a finite sample breakdown point. Consider a set Z of m good measurement

$$\hat{Z} = \{0.98 + 1.00 + 1.05 + 1.07 + 1.10\}/5 = 1.04$$

If one of the measurements is an outlier, say measurement 1.10 is replaced by an arbitrarily large value of ∞, then the estimated value becomes ∞. This means that even one abnormally large value will ruin the estimator, or the breakdown point for this estimator is zero.

If a median estimator is used instead of a mean estimator, then the estimated value is 1.05. Now if value 1.10 is replaced by ∞, the estimated value is still 1.05. If two measurements 1.07, 1.10 are replaced by ∞ each, the estimated value is still 1.05. However, if three measurements 1.05, 1.07, and 1.10 are replaced by ∞, the estimated value also becomes ∞ (unbounded). Thus the ratio 2/5 = 0.4 is the breakdown point and the median estimator is more robust than the mean estimator.

6.2.1 Leverage Points

Since the presence of even a single bad data will distort the WLS estimate, improvement in bad data detection and estimation methods have been developed for power system state estimators. Among these methods, the *weighted least absolute value estimator* (WLAV) that minimizes the weighted sum of absolute errors has been found to be simpler. The reason is that WLAV estimation problem can be solved as linear programming problem (Appendix 6A). WLAV estimators successfully reject bad measurement data as long as none of these measurements are leverage points [6]. Leverage points are defined as the points of a regression that are far away from the bulk of the data points in the factor space (Figure 6.1).

A common practice is to flag an observation as having a large leverage value, if its leverage value is three times larger than the mean leverage value [7]. Although the characterization of a leverage point depends only on the independent variable x, their classification as good or bad data also depends on the measured values z as well as on the corresponding variances. The location of points in *x-space* is important in determining the property of a regression model. Remote points in *x-space* have a large impact on parameter estimates, standard error, and model summary. There are two approaches to solve the problem when a measurement data set has leverage points:

1. To identify the leverage points and eliminate them before estimating.
2. To use estimators that are not sensitive to leverage points.

Dr. Bishaljit Paul

presently is working as Assistant Professor of EE department at Narula Institute of Technology, Agarpara. His research interests are Electricity Power Markets, Power Congestion Management Techniques, and Power System Optimization, Uncertainty Decision Making and Stochastic Programming.

Narula Institute of Technology is a premier educational institute in Kolkata. It is a part of JIS Educational Initiatives. Today, it ranks among the top private engineering college in Kolkata, West Bengal. Most of its eligible programs are NBA accredited. It has received several accolades and rankings from industry leaders. Some of them are NIRF, CAREERS360, India Today, The Week, and Times of India. Narula Institute of Technology stands among the best private engineering colleges in Kolkata, West Bengal through its near-perfect adherence to global quality standards.

Published by:

LEARNET PUBLISHING
19/b, Kali Kumar Majumder Road,
P.O.-Santoshpur Avenue, P.S.- Survey Park,
Kolkata-700075, West Bengal
Email: learnetpublishing@gmail.com
Website: www.learnrtpub.co.in
www.jctmg.in

ISBN 978-81-685619-2-2
9 788168 561922