Chapter 3

ROBUST TOPOLOGY ERROR IDENTIFICATION

§ 3.1 Motivation

As mentioned in Chapter 2, classic state estimation, derived by Schweppe and Wildes (1970), is based on a super-bus network modeling that relates the power and voltage measurements through a nonlinear equation involving the nodal voltage magnitudes and phase angles, termed the state variables. It is assumed that the topology of the network is completely known without error (which branches are active/inactive), and that the parameters of the lines are also known (the resistance, reactance, and capacitance). An incorrect assumed status of a line leads to a topology error, which then leads to multiple conforming bad data and incorrect state estimates. Research in topology error identification is well documented over the last two decades and numerous methods have been proposed. Most methods fall into the category of *post-processing* methods. That is, classic state estimation is carried out, and then analysis is done on the residuals. If outliers in the residuals correspond to a bus or a branch, then the topology around that area is in question. For examples of various post-processing methods, see

Clements and Davis (1988), Wu, and Liu (1989), or Costa and Leão (1993). However, due to the conforming nature of these measurements, especially if in large number, these methods may fail in detecting where the topology error(s) occurred. Thus there exists a need for a pre-processing method that is able to identify topology errors and discriminate between bad and good data.

§ 3.2 Network Modeling

The proposed model uses Kirchhoff's Current Law (KCL) as a powerful foundation. KCL states that the algebraic sum of all the current around a node (bus) is equal to zero. Equivalently, the power going in the network around a node is equal to the power leaving that node. Figure 3.1 depicts an example of KCL. Here, three lines emit from bus *j*. The circles represent measurements on the lines (flows) and at the bus (injection), and their specific values in *per unit* (*pu*). As one can see, the power entering the bus (*3.4 pu*) equals the power leaving the bus (*1.6 pu + 2.6 pu - 0.8 pu*). KCL will enable us to identify outliers, measurements that do not conform to KCL. For example, if the value of 2.6 in Figure 3.1 were changed to 7.6, that measurement would be labeled as an outlier since KCL is violated. Also, KCL will enable us to estimate power on lines that do not have a direct corresponding measurement. For example, if the value of 2.6 was removed, we can estimate the power on that line algebraically.

To use KCL, the mathematical model must encompass all lines of the network, irrespective of their assumed status. This would include short circuits that may split a bus into multiple sub buses (called substations). Modeling these types of power lines would help identify which way a bus is split. In Figure 3.2, a bus is modeled at the substation level. If the circuit breaker directly connecting bus *121* and bus *110* was closed, then we can represent this as just one super-bus, regardless of the breakers' statuses between the two buses. If this were the case,

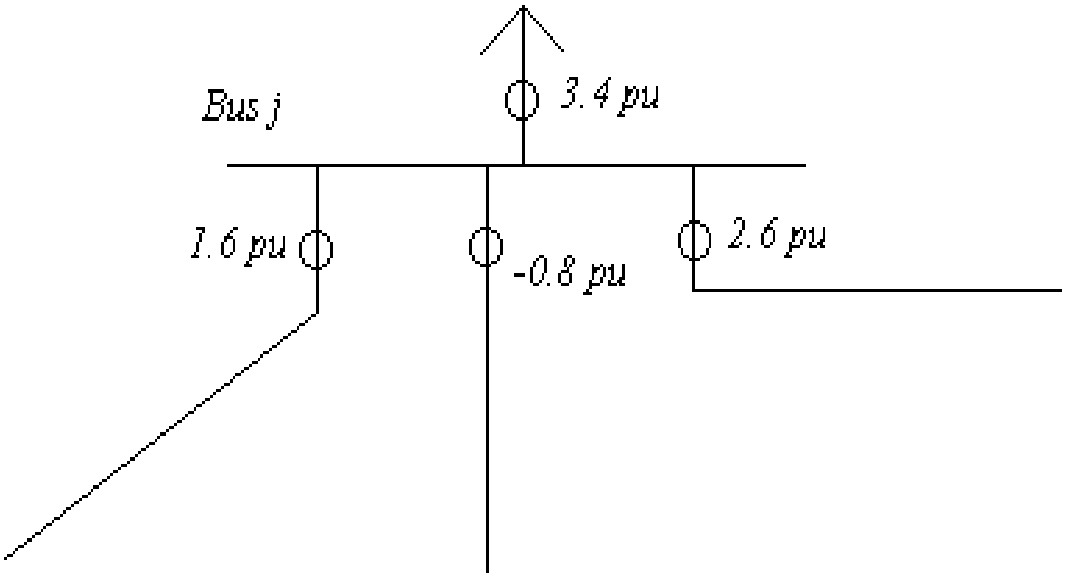

Figure 3.1. Example of Kirchhoff's Current Law

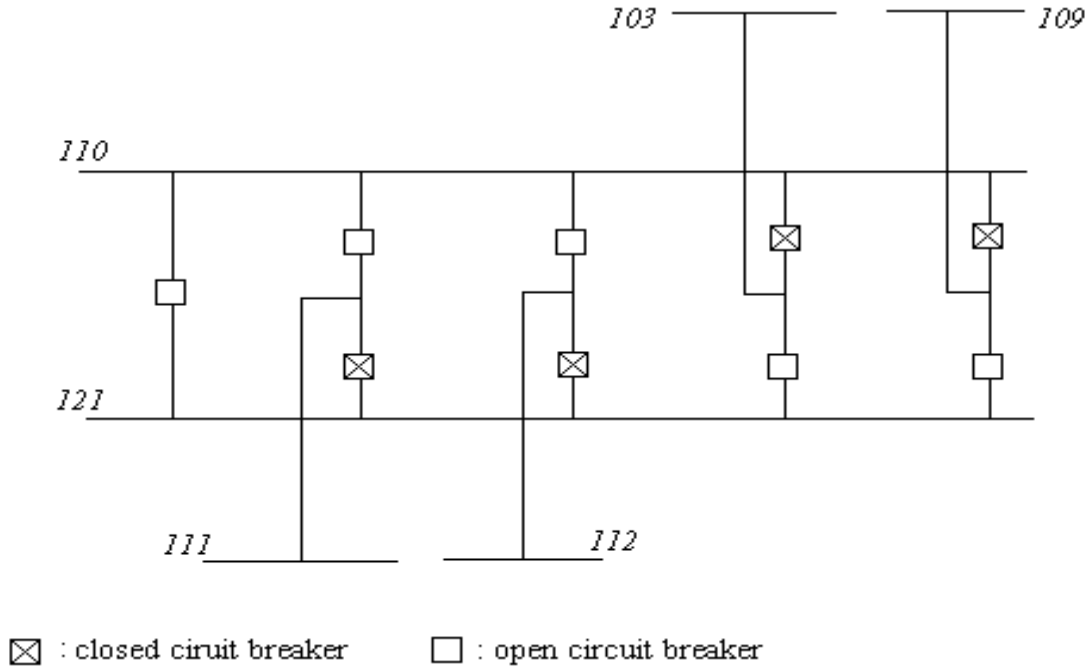

Figure 3.2. Example of a Substation

this would mean that bus *110-121* is connected to all four adjacent buses, *103, 109, 111*, and *112*. However, if we assume that the breaker is closed when in fact it is open, as depicted in Figure 3.2, then a topology error has occurred at the substation level. In actuality, bus *110* is connected to buses *103* and *109* and bus *121* is connected to buses *111* and *112*. Therefore all short circuit lines need to be modeled to find out which lines are energized, thus which way the bus is split.

§ 3.3 Modeling the Real Power

The intent is to derive a model that relates the real power measurements, $\underline{z}_P$, to the unknown state variables, $\underline{x}_p$. Each element in $\underline{x}_p$ corresponds to a line in the network, and every line in the network has a corresponding variable in $\underline{x}_p$. We start the derivation by considering the Ohmic losses that are dissipated in a branch. The power that travels through a branch leaving node *k* and entering node *l*, say, are equivalent if the line has no resistance. If there is resistance associated with line *k-l*, then some power is lost during the course of travel via friction. This power loss is called the Ohmic loss. As an example, if the power entering a branch is 3.0 *pu*, and the power leaving the branch is –2.8 *pu*, then the Ohmic loss for that branch is 0.2 *pu*. (For the real power, power that is leaving the branch is a negative value, thus the Ohmic loss for the line can be found by just adding the power on each end.) Formally, this quantity can be written as

$$P_{L_{kl}} = G_{kl}(V_k^2 + V_l^2 - 2V_kV_l(\cos(\theta_{kl})), \tag{3.1}$$

where V is the voltage magnitude, G is the known series conductance as defined in Appendix A, and $\theta_{kl} = \theta_k - \theta_l$ is the difference in voltage angles between bus k and bus l. If we let $x_{P_{kl}}$ be the power flow on the k side of the branch directed toward bus l, and assume that we have a decoupled model, then we can approximate θ_{kl} as the product of the known reactance of the line, X_{kl}, and the state variable, $x_{P_{kl}}$. That is, $\theta_{kl} \approx X_{kl} x_{P_{kl}}$. A decoupled model implies that the real power, P, and reactive power, Q, are analyzed independently of each other. Substituting into (3.1) yields

$$P_{L_{kl}} \approx G_{kl}(V_k^2 + V_l^2 - 2V_k V_l(\cos(X_{kl} x_{P_{kl}}))\,. \tag{3.2}$$

Finally, in per unit, the voltage magnitudes at each bus vary slightly above and below one. So, if we set $V_k = V_l = 1$ and substitute into (3.2), we get as an approximation to the true Ohmic losses,

$$P_{L_{kl}} \approx 2G_{kl}(1 - \cos(X_{kl} x_{P_{kl}}))\,, \tag{3.3}$$

and in (3.3), the only unknown quantity is $x_{P_{kl}}$. We will incorporate this equation into our model. We start by defining the unknown state variable, $x_{P_{kl}}$ as the power associated with a particular branch. Thus, we can express the real power measurements as a function of the unknown real power. Mathematically, we have

$$P_{kl} = x_{P_{kl}} + \varepsilon_{kl}\,, \text{ and } P_{lk} = -x_{P_{lk}} + P_{L_{kl}} + \varepsilon_{kl} \approx -x_{P_{lk}} + 2G_{kl}(1 - \cos(X_{kl} x_{P_{lk}})) + \varepsilon_{lk}\,, \tag{3.4}$$

where P_{kl} is the power measurement on the line near node *k* equal to the true power on the line, and P_{lk} is the power near node *l* equal to the negative true power on the line, plus Ohmic losses. For the injections at bus *k* and bus *l*, we obtain using Kirchhoff's Current Law

$$P_k = \sum_{l \in N(k)} P_{kl} \text{ and } P_l = \sum_{k \in N(l)} P_{lk},$$

where $N(k)$ and $N(l)$ are the set of all branches connected from node *k* and node *l*, respectively. This is mathematically stating that the injection power at a node is the sum of the power on the lines leaving that particular node. As a rule for state variable assignment, if we obtain one observation for branch *k*-*l*, then that observation, P_{kl}, is associated with the state variable, $x_{P_{kl}}$. In essence, we are picking a direction of the state, or a direction of the power flow. Figure 3.3 shows this for just one measurement on the line. So, regardless of whether the measurement was taken on the sending end (where the power is coming from) or the receiving end (where the power is going to), the measurement is assigned to a "sending end", even if the measurement was taken on the true receiving end. In summary, for just one measurement on the line, $P_{kl} = x_{P_{kl}}$. For zero or two measurements on a line, the direction of the state is chosen arbitrarily.

Now, assume we have n_b branches and m_P real power measurements, $n_b < m_P$. We can relate the measurements to the nonlinear equations defined in (3.4), which can be expressed mathematically as

$$\underline{z}_P = \underline{h}_P(\underline{x}_P) + \underline{e}_P,$$

where $\underline{z}_P$ is an (m_P x 1) vector of measurements, $\underline{h}_P(\cdot)$ is an (m_P x 1) vector of equations relating the state variables to the measurements, $\underline{x}_P$ is the (n_b x 1) vector of unknowns, and $\underline{e}_P$

is an (m_P x 1) vector of random errors with mean zero and a known covariance matrix R_P = $diag(\sigma^2_{P_1},...,\sigma^2_{Pm_P})$. We can form the Jacobian matrix of partial derivatives by defining

$$H_P(\underline{x}_P) = \frac{\partial \underline{h}_P(\underline{x}_P)}{\partial \underline{x}_P},$$

and specifically, we have

$$\frac{\partial P_{kl}}{\partial x_{P_{kl}}} = 1,\ \frac{\partial P_{lk}}{\partial x_{P_{kl}}} = -1 + 2X_{kl}G_{kl}\sin(X_{kl}x_{P_{kl}}),\ \text{and}\ \frac{\partial P_k}{\partial x_{P_{kl}}} = \sum_{l\in N(k)} \frac{\partial P_{kl}}{\partial x_{P_{kl}}}.$$

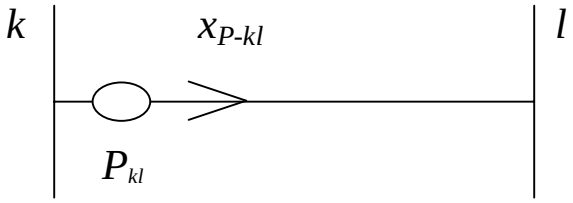

Figure 3.3. State Variable Assignment for One Observation

§ 3.4 Modeling the Reactive Power

The model for the reactive power is very similar to the model derived for the real power in Section 3.3. Again, every branch in the network has a corresponding unknown state variable, and here it is the reactive power flowing from bus *k* to bus *l*. As with the real power, we can express the reactive power on both sides of the branch as

$$Q_{kl} = x_{Q_{kl}},\ Q_{lk} = -x_{Q_{lk}} + Q_{L_{kl}},\ Q_k = \sum_{l\in N(k)} Q_{kl},\ Q_l = \sum_{k\in N(l)} Q_{lk},\qquad (3.5)$$

where $N(k)$, $N(l)$ are defined as before, and $Q_{L_{kl}}$ is the loss of reactive power for the branch *k-l*. The loss of reactive power is expressed as

$$Q_{L_{kl}} = -B_{s-kl}(V_k^2 + V_l^2) - B_{kl}(V_k^2 + V_l^2 - 2V_kV_l\cos(\theta_{kl})),$$

where *B* represents the known susceptance of the line as defined in Appendix A, *V* is the voltage magnitude, and $\theta_{kl} = \theta_k - \theta_l$ is the difference in voltage angles between bus *k* and bus *l*. (The subscript *s-kl* on the first susceptance term stands for the shunt of line *k-l*. The shunt does not have any resistance, thus for the real power loss described in Section 3.3, a term G_{s-kl} does not exist). Exactly as with the real power loss, substituting θ_{kl} with $X_{kl}x_{P_{kl}}$ and setting $V_k = V_l = 1$ *pu* yields

$$Q_{L_{kl}} \approx -2B_{s-kl} - 2B_{kl}(1-\cos(X_{kl}x_{P_{kl}})). \tag{3.6}$$

Notice that this loss is dependent on the true real power flow, $x_{P_{kl}}$. We estimate the real power flow with $\hat{x}_{P_{kl}}$ and substitute it back into (3.6) to obtain $\hat{Q}_{L_{kl}} = -2B_{s-kl} - 2B_{kl}(1-\cos(X_{kl}\hat{x}_{P_{kl}}))$. Therefore, the reactive power on the *l-k* side of the line is now expressed as

$$Q_{lk} \approx -x_{Q_{lk}} + \hat{Q}_{L_{kl}}.$$

Through simulations, we were pleased to see that $\hat{Q}_{L_{kl}}$ is quite accurate at estimating the true reactive power loss $Q_{L_{kl}}$, since this is an approximation to an approximation. Because it is independent of the state variable $x_{Q_{lk}}$, it can be calculated as soon as the estimate of real power is obtained.

The derivation of the reactive model again follows suit with the real power model. We assume we have n_b branches and m_Q reactive power measurements, $n_b < m_Q$. We can relate the measurements to the nonlinear equations defined in (3.5), which can be expressed mathematically as

$$\underline{z}_Q = \underline{h}_Q(\underline{x}_Q) + \underline{e}_Q,$$

where $\underline{z}_Q$ is an (m_Q x 1) vector of measurements, $\underline{h}_Q(\cdot)$ is an (m_Q x 1) vector of equations relating the state variables to the measurements, $\underline{x}_Q$ is the (n_b x 1) vector of unknowns, and $\underline{e}_Q$ is an (m_Q x 1) vector of random errors with mean zero and a known covariance matrix R_Q = $diag(\sigma^2_{Q_1},...,\sigma^2_{Qm_Q})$. We can form the Jacobian matrix of partial derivatives by defining

$$H_Q(\underline{x}_Q) = \frac{\partial \underline{h}_Q(\underline{x}_Q)}{\partial \underline{x}_Q},$$

and specifically, we have

$$\frac{\partial Q_{kl}}{\partial x_{Q_{kl}}} = 1,\ \frac{\partial Q_{lk}}{\partial x_{Q_{kl}}} = -1, \text{ and } \frac{\partial Q_k}{\partial x_{Q_{kl}}} = \sum_{l \in N(k)} \frac{\partial Q_{kl}}{\partial x_{Q_{kl}}}.$$

It should be noted that any leverage points *do not influence* the real and reactive models derived for detecting topology errors. Leverage points are common in classical state estimation and are due to the magnitude of the reactance of a line. Whether it is very large or very small, it can have a dramatic effect on the model matrix. In our model, the entries of the model matrix are basically 1's, 0's, and –1's (in the *P* design matrix the negative entries can be slightly smaller or larger than –1).

§ 3.5 The Huber Estimator and IRLS Algorithm

Peter Huber introduced the concept of M-estimation (1964) for estimating a location parameter with non-normal errors. He developed an estimator that performs like a least squares estimator for small residuals (below a specific cutoff value, λ) and performs like a least absolute value estimator for large residuals (above λ). Thus, the Huber estimator is a two-part function of the residuals and we desire to minimize the sum of the functions of the residuals. Specifically, if we let $J(\underline{x})$ be the objective function that we would like to minimize, then the solution that makes $J(\underline{x})$ a minimum is the estimate of $\underline{x}$, $\underline{\hat{x}}$. In general, $J(\underline{x})$ is given by

$$J(\underline{x}) = \sum_{i=1}^{m} \rho(r_{wi}),$$

where precisely for the Huber estimator,

$$\rho(r_{wi}) = \begin{cases} \frac{1}{2} r_{wi}^2 & , |r_{wi}| \le \lambda \\ \lambda |r_{wi}| - \frac{c^2}{2} & , |r_{wi}| > \lambda \end{cases}.$$

Here, $r_{wi} = r_i / \sigma_i$ is the i^{th} rescaled residual, $r_i = z_i - h_i(\underline{x})$ is the i^{th} residual, and λ is an arbitrary cutoff value which can range from 1.0 to 3.0. At the Gaussian distribution, the lower the cutoff value, the more robust the estimator is in the presence of outlying data; the higher the cutoff value, the more efficient the estimator is with regard to the variance of the estimates. Also, as the cutoff value decreases, the more the estimator becomes akin to the least absolute value estimator; and as the cutoff value increases, the more the estimator becomes akin to the least squares estimator. The estimate $\underline{\hat{x}}$ is the solution to $\partial J(\underline{x}) / \partial \underline{x} = \underline{0}$, and in particular,

$$\frac{\partial J(\underline{x})}{\partial \underline{x}} = \frac{\sum_{i=1}^{m} \partial \rho(r_{wi})}{\partial \underline{x}} = \frac{\sum_{i=1}^{m} \partial \rho(r_{wi})}{\partial r_{wi}} \frac{\partial r_{wi}}{\partial \underline{x}} = \sum_{i=1}^{m} \psi(r_{wi}) \frac{\partial r_{wi}}{\partial \underline{x}} =$$

$$\sum_{i=1}^{m} \psi(r_{wi}) \frac{\partial (z_i - h_i(\underline{x}))}{\partial \underline{x}} = \sum_{i=1}^{m} \psi(r_{wi}) \frac{1}{\sigma_i} \frac{\partial (z_i - h_i(\underline{x}))}{\partial \underline{x}} =$$

$$-\sum_{i=1}^{m} \psi(r_{wi}) \frac{1}{\sigma_i} H_i(\underline{x}) = -\sum_{i=1}^{m} \underline{\ell}_i \psi(r_{wi}) = \underline{0}.$$

We can multiply each side by (–1) to get rid of the negative sign. Finally, we have,

$$\sum_{i=1}^{m} \underline{\ell}_i \psi(r_{wi}) = \underline{0}, \tag{3.7}$$

where $\underline{\ell}^T$ is the i^{th} row of the weighted Jacobian matrix, $\sqrt{R^{-1}}H$, and $\psi(r_{wi})$ is Huber's ψ-function, the derivative of $\rho(r_{wi})$ with respect to r_{wi}. This is expressed as

$$\psi(r_{wi}) = \begin{cases} r_{wi}, & |r_{wi}| \le \lambda \\ \lambda, & |r_{wi}| > \lambda \end{cases}.$$

One method to solve (3.7) is to use the iterated reweighted least squares algorithm (IRLS). To set the foundation for this algorithm, we first multiply and divide (3.7) by r_{wi}. This yields

$$\sum_{i=1}^{m} \underline{\ell}_i \psi(r_{wi}) \frac{r_{wi}}{r_{wi}} = \sum_{i=1}^{m} \underline{\ell}_i q(r_{wi}) \frac{r_i}{\sigma_i} = \underline{0},$$

where $q(r_{wi}) = \psi(r_{wi}) / r_{wi}$ is a weight function bounded between 0 and 1. We can rewrite this in matrix notation as

$$H^T R^{-1} Q \underline{r} = \underline{0}, \tag{3.8}$$

where $\underline{r} = \underline{z} - \underline{h}(\underline{x})$ and Q is the (m x m) diagonal matrix of weights, $q(r_{wi})$. We can approximate $\underline{h}(\underline{x})$ by a first order Taylor series expansion about the point $\underline{x}^{(k)}$ by

$$\underline{h}(\underline{x}) = \underline{h}(\underline{x}^{(k)}) + H(\underline{x}^{(k)})(\underline{x} - \underline{x}^{(k)}),$$

where $\underline{x}^{(k)}$ is the estimate of $\underline{x}$ at the k^{th} iteration. Substituting into (3.8) gives, at the $(k+1)^{st}$ iteration,

$$H(\underline{x}^{(k)})^T R^{-1} Q(\underline{x}^{(k)})(\underline{z} - \underline{h}(\underline{x}^{(k)}) - H(\underline{x}^{(k)})(\underline{x}^{(k+1)} - \underline{x}^{(k)})) = \underline{0}, \text{ or}$$

$$H^{(k)T} R^{-1} Q^{(k)}(\underline{z} - \underline{h}(\underline{x}^{(k)})) - H^{(k)T} R^{-1} Q^{(k)} H^{(k)}(\underline{x}^{(k+1)} - \underline{x}^{(k)}) = \underline{0},$$

(using for shorthand notation, $A^{(k)} = A(\underline{x}^{(k)})$, for a matrix A dependent on $\underline{x}^{(k)}$),

$$H^{(k)T} R^{-1} Q^{(k)}(\underline{z} - \underline{h}(\underline{x}^{(k)})) = H^{(k)T} R^{-1} Q^{(k)} H^{(k)}(\underline{x}^{(k+1)} - \underline{x}^{(k)}),$$

$$\left(H^{(k)T} R^{-1} Q^{(k)} H^{(k)}\right)^{-1} H^{(k)T} R^{-1} Q^{(k)} \underline{r}^{(k)} = \Delta \underline{x}^{(k)}. \qquad (3.9)$$

In (3.9), $\Delta \underline{x}^{(k)} = \underline{x}^{(k+1)} - \underline{x}^{(k)}$, and $\underline{r}^{(k)} = \underline{z} - \underline{h}(\underline{x}^{(k)})$. Convergence is reached when all of the $|\Delta x_i|$, $i = 1,...,n_b$, are below some critical threshold or tolerance, ε. The IRLS algorithm proved numerically stable and showed good convergence rates in all simulations performed.

Once the state vector is estimated, the extreme outliers are identified, deleted from the measurement set and one execution of the IRLS algorithm is performed from the previous step. We delete the outlying measurements to cancel out any influence on the estimates of the state. Due to possible low redundancy, only those measurements that have a weighted residual with absolute amplitude of, say, at least 6.0, are completely deleted from the measurement set.

§ 3.6 Testing the Estimates

Once the power flows are estimated, a statistical test needs to be applied in order to verify a line's status. We are not interested in testing the individual elements in the estimated state

vector, per se, but rather the amplitudes of power on the line. To be more specific, we desire to test the power at the *sending end*. For example, the power could leave node *k*, the sending end, and enter node *l*, the receiving end, but we only have an observation at the receiving end. Thus the unknown state variable is associated with that measurement on the receiving end, as described in Section 3.3. If the line *k-l* has a resistance and/or reactance associated with it, then power has been lost over the course of travel through the line due to friction. As a consequence, our estimates of the real and reactive power for that line may be underestimating the true values. However, after obtaining the estimates of real power flow, $\hat{\underline{x}}_P$, we can estimate any loss of power on the line, both real and reactive. Once we obtain this estimated loss, we add it back to the state estimate of power for branch *k-l*, *if necessary*, thereby obtaining an estimate of power from the sending end. (The state variable for branch *k-l* may already be associated with the true sending end, as described in Section 3.3. If this is the case, then we do not need to add back possible losses). In summary, we desire to estimate an *observation* on the sending end, where the power just enters the line, whether a measurement was taken there or not. By adding back any possible losses that have accumulated, we increase the statistical power of the statistical test, which in turn decreases the number of cases where the line is wrongly tagged as disconnected. In essence, we assume nothing about the topology of the network and let the measurements estimate the topology. To continue, let $\hat{\tilde{\underline{z}}}$ be the vector of estimated *amplitudes* of power flow, the sending end estimates, and let the standardized estimates of amplitude for branch *j*, say, be $\hat{\tilde{z}}_j / \hat{s}_{z_j}$, $j = 1,\ldots, m$, where $\hat{s}_{z_j}$ is the standard error for the flow estimate, $\hat{\tilde{z}}_j$. To derive the standard error, we first write $\hat{\tilde{\underline{z}}}$ as a function of the state vector $\hat{\underline{x}}$. That is, $\hat{\tilde{\underline{z}}} = \tilde{\underline{h}}(\hat{\underline{x}})$, and $\tilde{\underline{h}}(\hat{\underline{x}})$

is the corresponding nonlinear vector function for each of the values of $\underline{\hat{\tilde{z}}}$. If we perform a first order Taylor series expansion of $\underline{\tilde{h}}(\underline{x})$ about $\underline{\hat{x}}$, we get

$$\underline{\tilde{h}}(\underline{x}) = \underline{\tilde{h}}(\underline{\hat{x}}) + \tilde{H}(\underline{\hat{x}})(x - \hat{x}), \text{ or}$$

$$\underline{\tilde{h}}(\underline{\hat{x}}) = \underline{\tilde{h}}(\underline{x}) - \tilde{H}(\underline{\hat{x}})(x - \hat{x}),$$

where $\tilde{H}(\underline{\hat{x}}) = \partial \underline{h}(\underline{x}) / \partial \underline{x}$ evaluated at $\underline{x} = \underline{\hat{x}}$. Replacing this expression for $\underline{\hat{\tilde{z}}} = \underline{\tilde{h}}(\underline{\hat{x}})$, we get

$$\underline{\hat{\tilde{z}}} = \underline{\tilde{h}}(\underline{x}) - \tilde{H}(\underline{\hat{x}})(x - \hat{x}).$$

Thus, the covariance matrix of $\underline{\hat{\tilde{z}}}$ can be found by

$$Cov(\underline{\hat{\tilde{z}}}) = Cov(\underline{\tilde{h}}(\underline{x}) - \tilde{H}(\underline{\hat{x}})(x - \hat{x})),$$

$$= Cov(-\tilde{H}(\underline{\hat{x}})(x - \hat{x})),$$

$$= \tilde{H} Cov(x - \hat{x}) \tilde{H}^T = \tilde{H} Cov(\hat{x}) \tilde{H}^T = S. \qquad (3.10)$$

For M-estimates, Huber (1973) derived the asymptotic covariance matrix of $\underline{\hat{x}}$ and it is given by

$$Cov(\hat{x}) = \hat{s}_x^2 (H^T R^{-1} H)^{-1},$$

and $\hat{s}_x$ is a scale estimate based on the weighted residuals and is calculated by

$$\hat{s}_x^2 = \frac{\sum_{i=1}^{m} \psi^2(r_{wi})}{\left(\sum_{i=1}^{m} \psi'(r_{wi}) \right)^2},$$

where ψ^2 is the square of the ψ-function, and ψ' is the first derivative with respect to the residual, r_{wi}. So, $\hat{s}_{z_j}^2$ is the j^{th} diagonal element of S in (3.10). The standardized flow estimates found by $\hat{\tilde{z}}_j / \hat{s}_{z_j}$ are then compared to a critical value which can vary between 2.5 and 3.0. If the estimate is smaller than the critical value, the line is labeled as off, and if the estimate is

larger than the critical value, the line is labeled as on. Monte Carlo simulations showed that the scale estimate ranged anywhere from 1.1 – 1.4 for different systems, but for practical purposes we set $\hat{s}_x^2 = 1.0$.

§ 3.7 Zero-Injections

One problem that occurs in practice deals with the concept of *redundancy*. As mentioned in Chapter 2, *redundancy* is the ratio of the number of observations relative to the number of state variables (parameters) to be estimated. In power systems, the redundancy can be quite low for the classical state estimation model when not accounting for any zero-injections. Low redundancy can cause problems in *observing* the network (whether all state variables can be estimated), as well as problems with adequate estimation of error (for statistical diagnostics). For the robust topology error identification method just described, the number of unknown state variables is about 1.5 times greater than the number of variables in classical state estimation. This, in turn, decreases the redundancy for the topology error method by about 33%. For some utility companies, extra measurements necessary to sufficiently estimate the topology may be very hard to acquire.

One technique to increase the number of measurements is to use zero-injections. For our methodology, zero injections are an excellent addition to the measurement set. A zero injection is the addition of one power measurement, assigned the value of zero, to all branches in the network that we *know* are not energized. This approach aids with estimation and assists with the convergence of the algorithm via additional observations. Since these are perfect measurements (i.e. no error), Kirchhoff's Current Law is now strengthened around that area of the network for the identification of outliers as well as for accurate estimation of power on branches not

associated with any measurements. This is due to the additional weight given in the matrix *R* for those zero measurements, although this is not the only way to handle these types of measurements.

§ 3.8 Simulation Results

We applied the topology error identification to a 120-bus system that was derived from the standard IEEE 118-bus system. Specifically, we altered bus *11* and bus *24* by making them two buses that were connected by a single bus coupler. So, bus *119* is coupled with bus *11* and bus *120* is coupled with bus *24*. These buses only split in one way (i.e. the coupler is on or off, thus the buses are connected or not). We model the short circuit branches, the bus couplers, as actual lines in our model, and model all other buses as *super-nodes*. Restating, a super-node is a bus that is actually a conglomerate of smaller buses. In Figure 3.2, if we considered bus *110* and bus *121* to be connected, then it would constitute a super-node. The measurement configuration consisted of 240 pairs of *P* and *Q* measurements along with 19 zero injections for a total of 259 pairs of observations. The difference between a regular measurement and a zero injection is that the zero injection is a perfect measurement (i.e. no error) and consequently is given a much higher weight in *R*, the diagonal matrix of known measurement variances. The network has 181 branches, implying there are 181 pairs of power flows that need to be estimated, the state variables for the *P* and *Q* models.

Several cases of topology errors and gross measurements were introduced to the system, and we will go into specific detail with one particular case. We introduced 9 topology errors and 8 bad measurements in the system (4 bad *P*-measurements and 4 bad *Q*-measurements). For the topology errors, we assumed that lines *27-115*, *28-9*, and *62-67* are connected when in fact they

are disconnected. In addition we assume that lines *15-19*, *49-54*, and *69-70* are disconnected when in fact they are connected. We also assumed that the bus coupler between bus *11* and bus *119* is closed while in fact it is open and the bus coupler between *24-120* is open while in fact it is closed. Finally, we assumed that the coupler between buses *110* and *121*, as depicted in Figure 3.2, is closed (it was modeled as a super-node). The injection measurement at bus *110* is only associated with the power on lines connected to bus *103* and bus *109*. Now, since bus *110* is represented as a super-node, thus the injection is the sum of power associated with lines to buses *103*, *109*, *111*, and *112*, we expect that the Huber estimator rejects the real and reactive power injection measurements at that bus as bad data. This is due to Kirchhoff's Current Law. It does that effectively by providing these injection measurements with rescaled residuals of -7.61 and 5.56, respectively. In the second step of the procedure, the substation associated with bus 110 is represented in detail as shown in Figure 3.4.

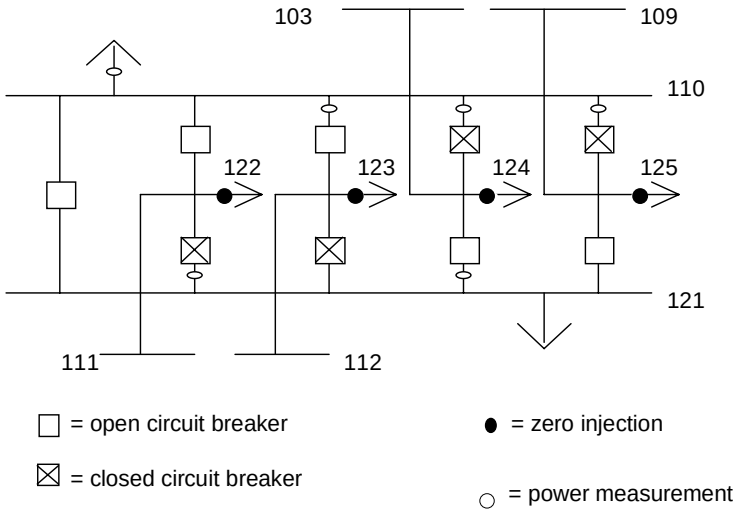

Figure 3.4. Detailed Substation

We model the circuit breakers in the substation based on the pioneering scheme proposed by Irving and Sterling (1982) coupled with the work by Monticelli (1993) where the circuit breakers are modeled as lines connected to zero injection buses. This now leads to a system with 125 buses and 190 branches. The Huber estimator is then executed. The measurements with absolute weighted residuals larger than 3 are displayed in Table 3.1. We observe in Table 3.2 that all of the 8 bad data have been properly rejected. The extreme bad data with $|r_i/\sigma_i| > 6.0$ are then deleted and one iteration of the IRLS algorithm was carried out. The results of the statistical test applied to the flow estimates are summarized in Table 3.1. We see that the method identifies correctly which branches are connected and which are not. For instance, it identifies that branches *110-121*, *110-122*, *110-123*, *121-124*, and *121-125* are off, revealing that bus *110* splits in two separate buses.

Table 3.1. Power Rescaled Residuals for the 125-Bus System.

Measurement	*Exact value*	*Est. value*	*Weight. Resid.*
PFL 8-30	0.72	0.66	5.25
PFL 30-8	-0.72	-0.66	30.52
PFL 49-69	-0.48	-0.59	26.82
PFL 69-49	0.51	0.62	-6.36
PFL 63-59	1.55	1.49	-42.13
PIN 115	-0.22	-0.07	-13.80
QFL 30-8	-0.73	-0.66	20.88
QFL 49-69	0.11	-0.07	-17.54
QFL 63-59	-0.34	-0.33	30.23
QIN 9	0.01	-0.20	-4.73
QIN 115	-0.07	0.08	-11.09

Note: **Bold** represents the outlying data introduced into the system

Table 3.2. Standardized Flows and Branch Statuses

Branch	$\hat{P}_{kl}$	$\hat{P}_{kl} / \hat{s}_{zi}$	$\hat{Q}_{kl}$	$\hat{Q}_{kl} / \hat{s}_{zi}$	*Status*
15-19	0.14	6.60	-0.17	-4.52	ON
49-54	0.71	23.07	-0.27	-5.37	ON
70-69	1.02	77.72	0.16	6.75	ON
27-115	0.01	0.59	-0.02	-1.06	OFF
28-29	0.01	1.17	-0.02	-1.00	OFF
62-67	0.01	0.57	-0.02	-1.30	OFF
11-119	0.28	22.75	-0.34	-14.32	ON
24-120	0.02	1.31	-0.01	-0.52	OFF
110-121	0.11	1.23	0.05	0.55	OFF
110-122	0.01	0.15	0.00	0.08	OFF
121-122	0.36	28.16	0.00	0.06	ON
110-123	0.04	0.66	-0.02	-0.29	OFF
121-123	0.68	54.23	-0.30	-11.90	ON
110-124	0.58	56.69	0.03	1.207	ON
121-124	0.01	0.86	-0.01	-0.41	OFF
110-125	0.13	10.63	0.10	4.11	ON
121-125	0.01	0.21	0.01	0.16	OFF