Introduction

Power systems operate in one of three operating states:

- Normal state:

  Loads = Generation - Losses
  Operational constraints are NOT violated.

  If the system manages to remain in the normal state following contingencies, then it will be considered to be in a secure normal operating state.
  Requires preventive control action.

- Emergency state:

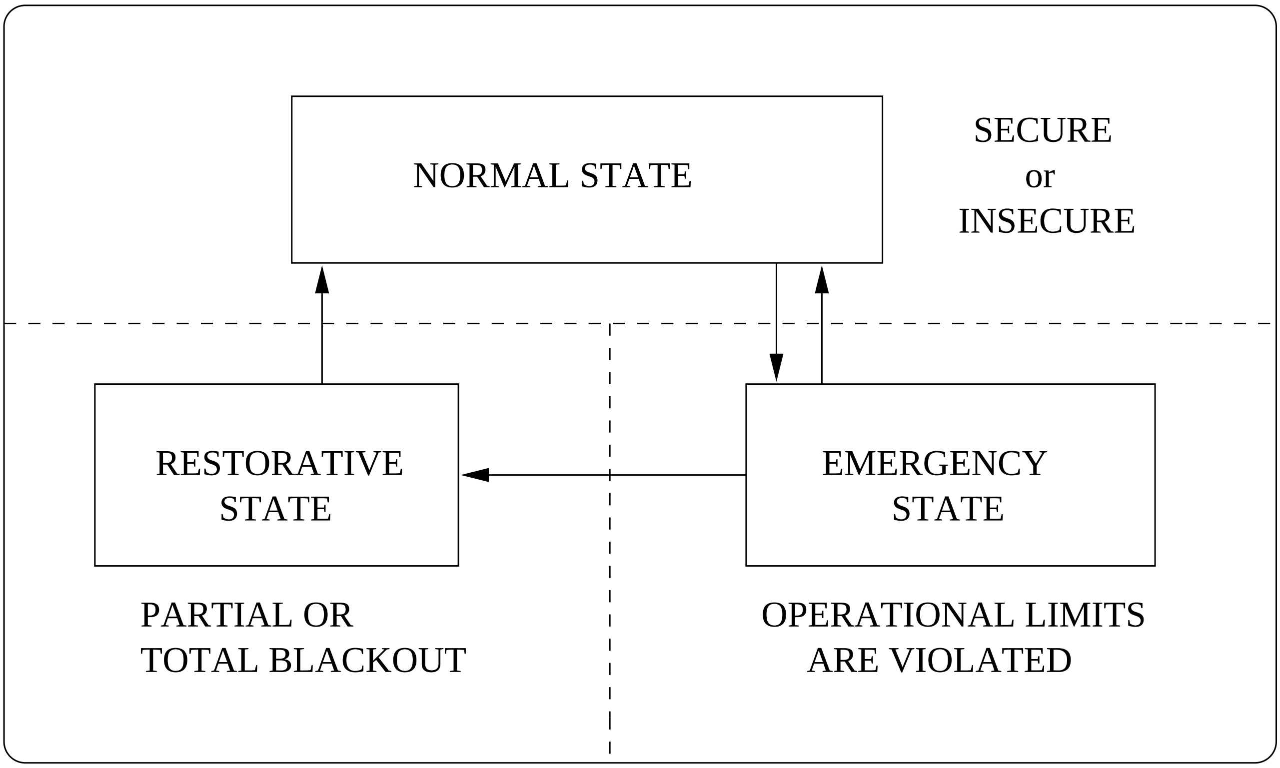
  Operating constraints are violated
  Requires immediate corrective action.

- Restorative state:

  Load versus generation balance is to be restored
  Requires restorative control actions.

System Operating States

Three operating states of power systems:



Role of SE

- Security Analysis:
  Monitoring the system, identifying its operating state, determining necessary preventive actions to make it secure.

- Monitoring involves:
  RTU's to measure and telemeter various quantities and a state estimator

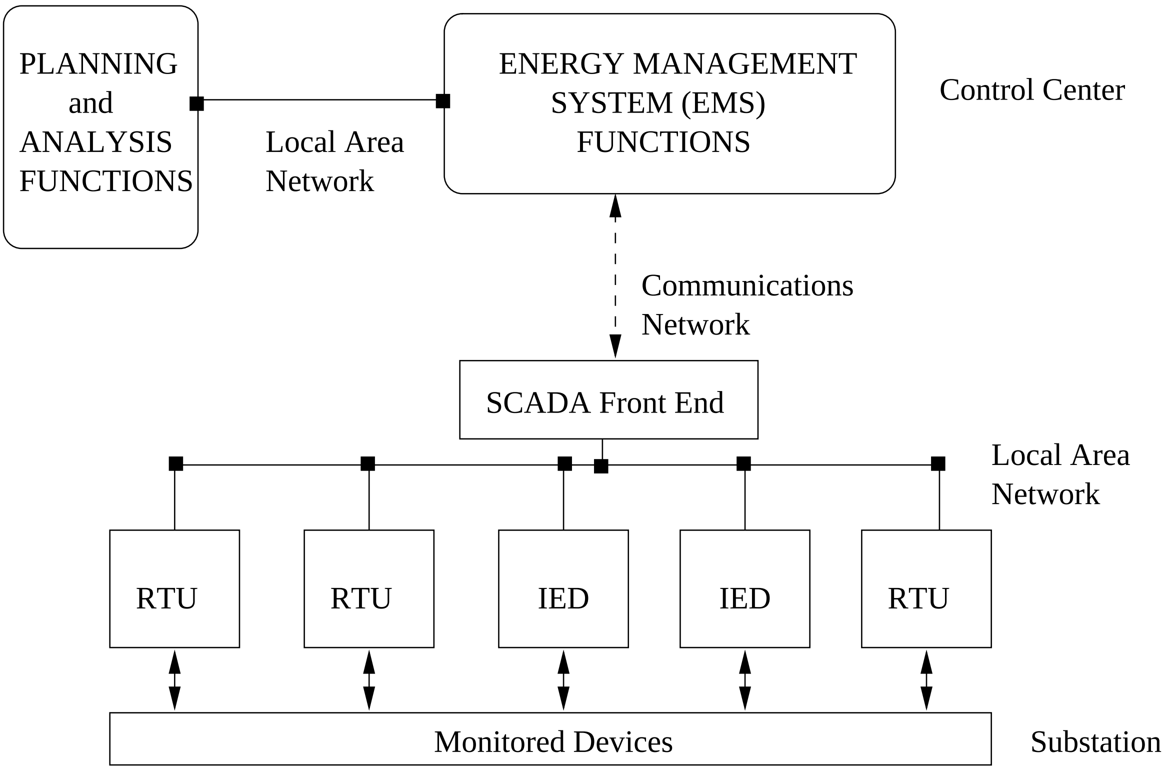
Measured quantities:

- Flows: line power flows,
- Phasor Magnitude: bus voltage and line current magnitudes,
- Phasor Angle: phase angle for bus voltage and line current,
- Injections: generator outputs and loads,
- Status: circuit breaker and switch status information, transformer tap positions, and switchable capacitor bank values.

First Introduction of the Concept

State estimation was proposed by Prof. Fred Schweppe (1934-1988) of MIT, as a tool to address both of the above issues in a computationally efficient manner. Hence, the old SCADA systems left their places to a new generation of Energy Management Systems (EMS) equipped with, among other applications, an on-line State Estimator (SE).

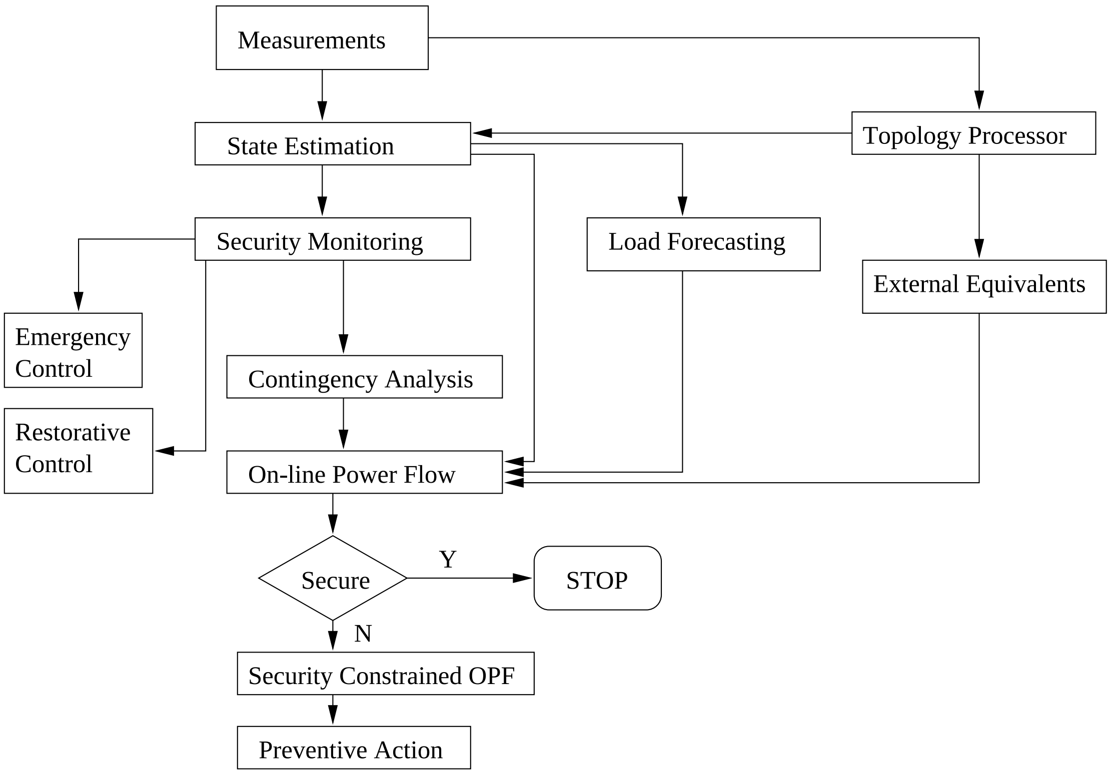
SCADA/EMS Configuration



State Estimation Functions

- Topology processor:
  Creates one-line diagram of the system using the detailed circuit breaker status information.
- Observability analysis:
  Checks to make sure that state estimation can be performed with the available set of measurements.
- State estimation:
  Estimates the system state based on the available measurements.
- Bad data processing:
  Checks for bad measurements. If detected, identifies and eliminates bad data.
- Parameter and structural error processing:
  Estimates unknown network parameters, checks for errors in circuit breaker status.

Security Analysis Functional Diagram



Weighted Least Squares State Estimation

Types of Measurements

1. Flows – Real and reactive power flows measured at the terminal buses of a transmission line or transformer.
2. Injections – Real and reactive power generation and load at system buses.
3. Voltage magnitude – At system buses.
4. Current magnitude – Ampere flows along transmission lines or transformers measured at the terminal buses.
5. Voltage phasor – At system buses.
6. Current phasor – Along transmission lines or transformers.

Gaussian (Normal) Density Function, $f(z)$

$$f(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left\{\frac{z-\mu}{\sigma}\right\}^2}$$

where z : random variable
μ : mean of $z = \mathrm{E}(z)$
σ : standard deviation of z

Standard Gaussian (Normal) Density Function, $\Phi(u)$

Let $u = \frac{z-\mu}{\sigma}$, then $\mathrm{E}(u) = 0$, $\mathrm{Var}(u) = 1.0$.

$$\Phi(u) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2}}$$

Likelihood Function, $f_m(z)$

$$f_m(z) = f(z_1) f(z_2) \cdots f(z_m)$$

Log-Likelihood Function, $\mathcal{L}$

$$\begin{aligned}
\mathcal{L} = \log f_m(z) \; &= \; \sum_{i=1}^{m} \log f(z_i) \\
&= \; -\frac{1}{2} \sum_{i=1}^{m} \left(\frac{z_i - \mu_i}{\sigma_i}\right)^2 - \frac{m}{2} \log 2\pi - \sum_{i=1}^{m} \log \sigma_i
\end{aligned}$$

Maximum Likelihood Estimator, MLE

MLE will maximize the likelihood (or log-likelihood) function for a given set of observations $z_1, z_2, \ldots, z_m$:

$$\begin{array}{rl} \text{maximize} & \log f_m(z) \\ \text{OR} & \\ \text{minimize} & \displaystyle\sum_{i=1}^{m} (\frac{z_i - \mu_i}{\sigma_i})^2 \end{array}$$

Let:

$$r_i = z_i - E(z_i)$$

$E(z_i) = h_i(x)$,
h_i is a nonlinear measurement function.

The MLE can be found by solving the following optimization problem:

$$\begin{array}{rl} \text{minimize} & \displaystyle\sum_{i=1}^{m} W_{ii} r_i^2 \\ \text{subject to} & z_i = h_i(x) + r_i, \quad i = 1, \ldots, m. \end{array}$$

where $W_{ii} = \sigma_i^{-2}$.

The solution of the above optimization problem is called the *weighted least squares (WLS) estimator* for x.

Measurement Model

Consider the set of measurements given by vector z:

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} = \begin{bmatrix} h_1(x_1, x_2, \ldots, x_n) \\ h_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ h_m(x_1, x_2, \ldots, x_n) \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$$

where:

$h^T = [h_1(x),\, h_2(x),\, \ldots,\, h_m(x)]$

$h_i(x)$ is the nonlinear function relating measurement i to the state vector x

$x^T = [x_1,\, x_2,\, \ldots,\, x_n]$ is the system state vector

$e^T = [e_1,\, e_2,\, \ldots,\, e_m]$ is the vector of measurement errors.

In compact form:

$$z = h(x) + e$$

Necessary and Sufficient Conditions

- State of a power system with N buses can be estimated if there are at least $(2N-1)$ measurements.

  Necessary condition:

$$m \geq n = (2N - 1)$$

$$\begin{aligned} n &: \text{ Number of state variables} \\ m &: \text{ Number of measurements} \end{aligned}$$

- In the set of m available measurements, there should be at least one subset of n **linearly independent** measurements.

  Sufficient condition:

  At least one set of n linearly independent measurements.

- If there are exactly n measurements that are linearly independent, then the estimated state will **exactly** satisfy all measurements, i.e. all measurement residuals will be zero.

Assumptions

Assumptions regarding the statistical properties of the measurement errors:

- $\mathrm{E}(e_i) = 0, \quad i = 1, \ldots, m.$
- Measurement errors are independent,

  i.e. $E[e_i e_j] = 0$. Hence:

  $R = Cov(e) = E[e \cdot e^T]$

  $= \text{diag} \, \{ \, \sigma_1^2, \sigma_2^2, \cdots, \sigma_m^2 \, \}.$

The WLS estimator will minimize the following objective function:

$$
\begin{aligned}
J(x) &= \sum_{i=1}^{m} (z_i - h_i(x))^2 / R_{ii} \\
&= [z - h(x)]^T R^{-1} [z - h(x)]
\end{aligned}
$$

Solution Method

The following first order optimality conditions will have to be satisfied at the optimum point:

$$g(x) = \frac{\partial J(x)}{\partial x} \; = \; -H^T(x)R^{-1}[z - h(x)] = 0$$

$$\text{where} \quad H(x) \; = \; \left[\frac{\partial h(x)}{\partial x}\right]$$

First order Taylor approximation:

$$g(x^{k+1}) \; \approx \; g(x^k) + G(x^k) \cdot (x^{k+1} - x^k)$$

$$x^{k+1} \; = \; x^k - [G(x^k)]^{-1} \cdot g(x^k)$$

where :

$$G(x^k) \; = \; \frac{\partial g(x^k)}{\partial x} = H^T(x^k) \cdot R^{-1} \cdot H(x^k)$$

$$g(x^k) \; = \; -H^T(x^k) \cdot R^{-1} \cdot (z - h(x^k)).$$

k is the iteration index.

Properties of the WLS Estimators

Consider the linearized measurement equations:

$$\begin{aligned}
\Delta z &= H \Delta x + e \\
\text{where :} & \\
\Delta z &= z - h(x^k) \\
\Delta x &= x^{k+1} - x^k \\
H &= \frac{\partial h(x^k)}{\partial x}
\end{aligned}$$

The WLS estimate for Δx will be given by:

$$\begin{aligned}
\Delta \hat{x} &= (H^T R^{-1} H)^{-1} H^T R^{-1} \Delta z \\
&= G^{-1} H^T R^{-1} (H \Delta x + e) \\
&= \Delta x + G^{-1} H^T R^{-1} e
\end{aligned}$$

Covariance matrix for the state estimate can then be written as:

$$E[(\Delta \hat{x} - \Delta x)(\Delta \hat{x} - \Delta x)^T] = (H^T R^{-1} H)^{-1} = G^{-1}$$

Note that the WLS estimator is identical to the MLE if the errors are Gaussian distributed.

Solution Algorithm

1. Start iterations, set the iteration index $k = 0$.
2. Initialize the state vector x^k, typically at flat start.
3. Calculate the gain matrix, $G(x^k)$.
4. Calculate the right hand side:

   $$t^k = H(x^k)^T R^{-1}(z - h(x^k))$$

5. Decompose $G(x^k)$ and solve for Δx^k.
6. Test for convergence, $\max \mid \Delta x^k \mid \leq \epsilon$?
7. If no, then:

   $x^{k+1} = x^k + \Delta x^k$,
   $k = k + 1$,
   and go to step 3.

   Else, stop.

Normal Equations

The following is called the "Normal Equation" which is solved at each iteration k:

$$[G(x^k)]\Delta x^{k+1} = H^T(x^k)R^{-1}[z - h(x^k)]$$

where $\Delta x^{k+1} = x^{k+1} - x^k$.

The matrix $G(x)$ is never inverted. The inverse will in general be a full matrix, whereas $G(x)$ itself is quite sparse.

Instead, it is decomposed into its triangular factors and sparse forward/back substitutions are used to obtain the solution at each iteration k.

Computations

1. Calculation of the right hand side $H^T R^{-1}[z - h(x^k)]$.

   (a) Calculating the measurement function, $h(x^k)$.

   (b) Building the measurement jacobian, $H(x^k)$.

2. Calculation of $G(x^k)$ and solution of the Normal equations.

   (a) Building the gain matrix, $G(x^k)$.

   (b) Decomposing $G(x^k)$ into its Choleski factors.

   (c) Performing the forward/back substitutions to solve for Δx^{k+1}.

The measurement jacobian, H

The structure of H will be as follows:

$$H = \begin{bmatrix} \frac{\partial P_{inj}}{\partial \theta} & \frac{\partial P_{inj}}{\partial V} \\ \\ \frac{\partial P_{flow}}{\partial \theta} & \frac{\partial P_{flow}}{\partial V} \\ \\ \frac{\partial Q_{inj}}{\partial \theta} & \frac{\partial Q_{inj}}{\partial V} \\ \\ \frac{\partial Q_{flow}}{\partial \theta} & \frac{\partial Q_{flow}}{\partial V} \\ \\ \frac{\partial I_{mag}}{\partial \theta} & \frac{\partial I_{mag}}{\partial V} \\ \\ 0 & \frac{\partial V_{mag}}{\partial V} \end{bmatrix}$$

The gain matrix, G:

$$G(x^k) = H^T R^{-1} H$$

it has the following properties:

1. It is structurally and numerically symmetric.
2. It is sparse, yet less sparse compared to H.
3. In general it is a non-negative definite matrix, i.e. all of its eigenvalues are non-negative. It is positive definite for fully observable networks.

It can be efficiently built as:

$$H = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_m \end{bmatrix}, R = \begin{bmatrix} R_{11} & 0 & \cdots & 0 \\ 0 & R_{22} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & R_{mm} \end{bmatrix}$$

$$G = \sum_{i=1}^{m} H_i^T R_{ii}^{-1} H_i$$

Solution of Sparse Linear Equations

Consider the linear equation given below:

$$Ax = b$$

representing:

$$[G(x^k)]\Delta x^{k+1} = H^T(x^k)R^{-1}[z - h(x^k)]$$

Direct solution of x:

1. Triangular decomposition of A:

$$\begin{aligned} A &= LDU \\ &\text{OR} \\ A &= LL^T \quad \text{if symmetric and p.d.} \end{aligned}$$

2. Forward and back substitutions:

$$\begin{aligned} y &= D^{-1}L^{-1}b \\ x &= U^{-1}y \end{aligned}$$

OR if symmetric and p.d.

$$\begin{aligned} y &= L^{-1}b \\ x &= (L^T)^{-1}y \end{aligned}$$

Factorization of Sparse Symmetric Matrices

$$
\begin{aligned}
A = A_0 \; &= \begin{bmatrix} d_1 & a_1^T \\ a_1 & B_1' \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{d_1} & 0 \\ \frac{a_1}{\sqrt{d_1}} & I_{n-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \underbrace{B_1' - \frac{a_1 a_1^T}{d_1}}_{B_1} \end{bmatrix} \begin{bmatrix} \sqrt{d_1} & \frac{a_1^T}{\sqrt{d_1}} \\ 0 & I_{n-1} \end{bmatrix} \\
&= L_1 \cdot A_1 \cdot L_1^T
\end{aligned}
$$

The above decomposition can now be applied to the submatrix B_1 as below:

$$
\begin{aligned}
A_1 \; &= \begin{bmatrix} 1 & 0 \\ 0 & B_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & d_2 & a_2^T \\ 0 & a_2 & B_2' \end{bmatrix} \\
&= \begin{bmatrix} 1 & & \\ 0 & \sqrt{d_2} & \\ 0 & \frac{a_2}{\sqrt{d_2}} & I_{n-2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & B_2' - \frac{a_2 a_2^T}{d_2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ & \sqrt{d_2} & \frac{a_2^T}{\sqrt{d_2}} \\ & & I_{n-2} \end{bmatrix} \\
&= L_2 \cdot A_2 \cdot L_2^T
\end{aligned}
$$

Continuing with this process, A can be written as the following product of elementary matrices:

$$
\begin{aligned}
A \; &= \underbrace{L_1 \cdot L_2 \cdots L_{n-1} \cdot L_n}_{L} \cdot \underbrace{L_n^T \cdot L_{n-1}^T \cdots L_2^T \cdot L_1^T}_{L^T} \\
&= L \cdot L^T
\end{aligned}
$$

Ordering Sparse Symmetric Matrices

Cholesky factorization scheme proceeds one pivot at a time:

$$B' - \frac{aa^T}{d}$$

The matrix aa^T will have as many non-zeros as the square of the number of nonzeros in a.

Tinney-2 scheme:

1. Choose the row with the minimum number of nonzeros.
2. Using the chosen row/column as the pivot, carry out one elimination step using Cholesky method, i.e. $B' - aa^T/d$.
3. Go back to step 1 and repeat the procedure until all pivots are processed.

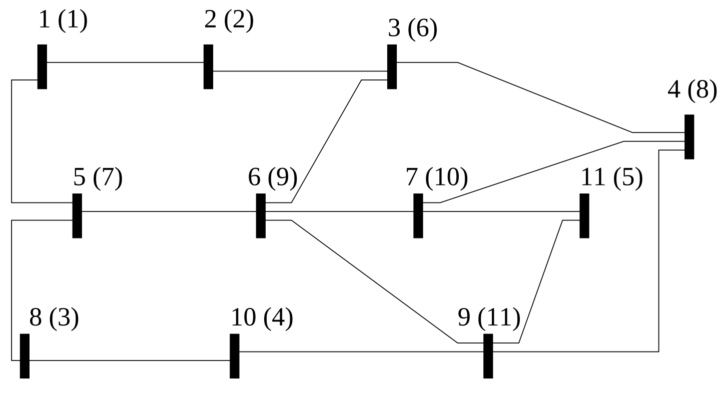
Example

Let A be a sparse symmetric matrix as given below:

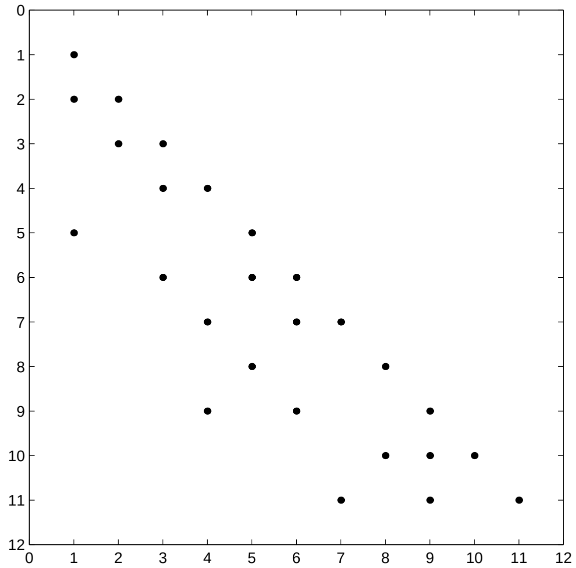
$$A = \begin{bmatrix} 12 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 12 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 12 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 12 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 12 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 12 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 12 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 12 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 12 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 12 \end{bmatrix}$$

Applying the Tinney-2 ordering algorithm to matrix A yields the following row/column order:

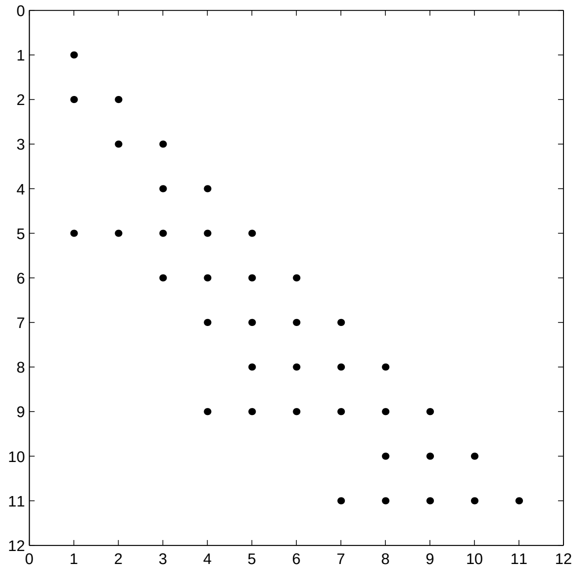
$$\begin{bmatrix} 1 & 2 & 8 & 10 & 11 & 3 & 5 & 4 & 6 & 7 & 9 \end{bmatrix}$$



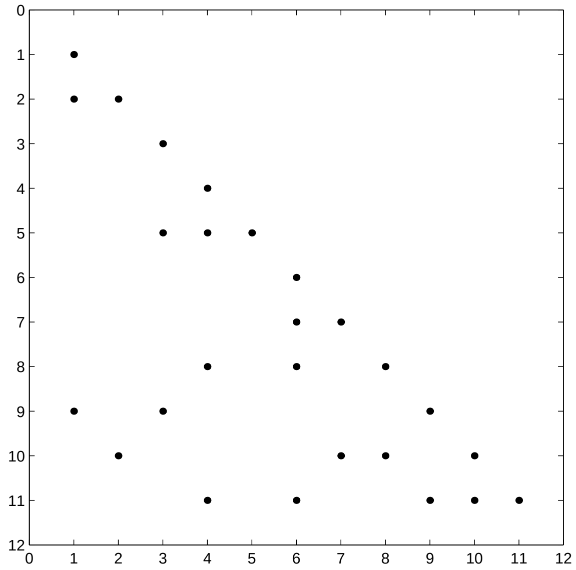
Sparsity Structure of Lower Triangular Part of A NZ=26



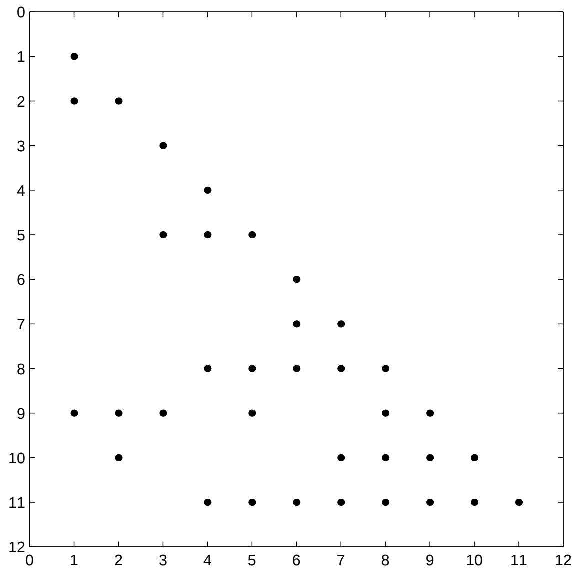
Sparsity Structure of Lower Triangular Factor of A (Not ordered), NZ=38



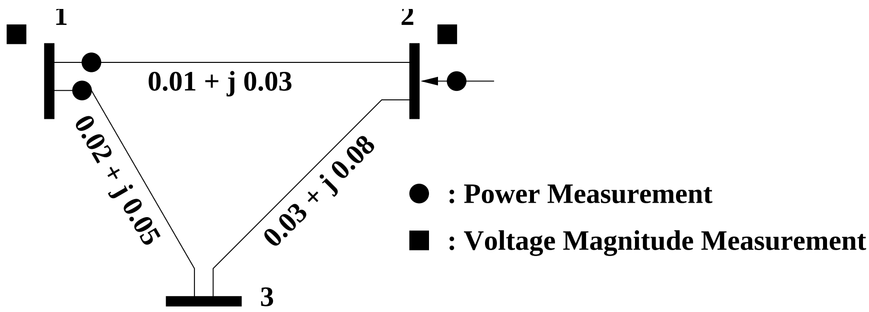
Sparsity Structure of Lower Triangular Part of A after Tinney-2 Ordering, NZ=26



Sparsity Structure of Lower Triangular Factor of A after Tinney-2 Ordering, NZ=33



Example 1:



Results of the 1st iteration

$$
H(x^0) = \begin{array}{c} \\ \partial p_{12} \\ \partial p_{13} \\ \partial q_{12} \\ \partial q_{13} \\ \partial p_2 \\ \partial q_2 \\ \partial V_1 \\ \partial V_2 \end{array}
\begin{array}{c} \begin{array}{ccccc} \partial\theta_2 & \partial\theta_3 & \partial V_1 & \partial V_2 & \partial V_3 \end{array} \\
\left[\begin{array}{rrrrr}
-30.0 & & 10.0 & -10.0 & \\
 & -17.2 & 6.9 & & -6.9 \\
10.0 & & 30.0 & -30.0 & \\
 & 6.9 & 17.2 & & -17.2 \\
40.9 & -10.9 & -10.0 & 14.1 & -4.1 \\
-14.1 & 4.1 & -30.0 & 40.9 & -10.9 \\
 & & 1.0 & & \\
 & & & 1.0 &
\end{array} \right] \end{array}
$$

$$
x^0 = \begin{array}{c} \theta_2 \\ \theta_3 \\ V_1 \\ V_2 \\ V_3 \end{array} \left[\begin{array}{c} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{array} \right]
$$

$$
\begin{aligned}
z^T &= [p_{12}, p_{13}, q_{12}, q_{13}, p_2, q_2, V_1, V_2] \\
&= [0.888, 1.173, 0.568, 0.663, -0.501, -0.286, 1.006, 0.968]
\end{aligned}
$$

$$
R^{-1} = diag[15625, 15625, 15625, 15625, 10000, 10000, 62500, 62500]
$$

Convergence Summary

Iterations, k	1	2	3
$\Delta\theta_2^k$	-2.1e-2	-5.52e-4	0
$\Delta\theta_2^k$	-4.52e-2	-2.69e-3	3.0e-6
ΔV_1^k	1.67e-4	-1.09e-4	-2.0e-6
ΔV_2^k	-2.53e-2	-1.06e-4	-2.0e-6
ΔV_3^k	-5.68e-2	1.15e-3	2.0e-6
Objective Function $J(x^k)$	49,144.3	59.1	9.1

Fast Decoupled Formulation

Consider the following **A**ctive/**R**eactive partitioning:

$$
\begin{aligned}
z^T &= [z_A^T \;\; z_R^T] \\
H &= \begin{bmatrix} H_{AA} & H_{AR} \\ H_{RA} & H_{RR} \end{bmatrix} \\
R &= \begin{bmatrix} R_A & 0 \\ 0 & R_R \end{bmatrix}
\end{aligned}
$$

The decoupling assumptions are:

1. Constant, decoupled gain matrix evaluated at flat start:

$$
\begin{aligned}
G &= \begin{bmatrix} G_{AA} & 0 \\ 0 & G_{RR} \end{bmatrix} \\
G_{AA} &= H_{AA}^T R_A^{-1} H_{AA} \\
G_{RR} &= H_{RR}^T R_R^{-1} H_{RR}
\end{aligned}
$$

2. Approximated right hand side vector:

$$
T = \begin{bmatrix} H_{AA}^T R_A^{-1} \Delta z'_A \\ H_{RR}^T R_R^{-1} \Delta z'_R \end{bmatrix} = \begin{bmatrix} T_A \\ T_R \end{bmatrix}
$$

$$
\begin{aligned}
\text{where :} \quad \Delta z'_A &= \Delta z_A / V \\
\Delta z'_R &= \Delta z_R / V
\end{aligned}
$$

Decoupled Solution Algorithm

1. Build and factorize G_{AA} and G_{RR}.
2. Calculate T_A.
3. Solve $G_{AA}\,\Delta\theta = T_A$.
4. Update $\theta^{k+1} = \theta^k + \Delta\theta$.
5. Calculate T_R.
6. Solve $G_{RR}\,\Delta V \;=\; T_R$.
7. Update $V^{k+1} = V^k + \Delta V$.
8. Go to 2.

Repeat steps 2 – 8 until both $\mid \Delta\theta \mid$ and $\mid \Delta V \mid$ become less than the convergence tolerance.

Decoupled versus Full Solution

Benefits of the Decoupled Solution:

- Requires less memory.
- Requires less computation time per iteration. Matrices are smaller and they are constant. They are factorized only once at the first iteration.

Limitations of the Decoupled Solution:

- Network parameters or operating conditions violating the decoupling assumptions, may lead to numerical instability.
- Current magnitude (ampere) or phase measurements can not be readily decoupled. They should be excluded from the measurement set when using decoupled solution.

Computer Exercises [DAY 1]

Exercise 1: [Use the file *exer01.pet* for this exercise]

Given the network diagram and its associated measurement configuration, use the ©P.E.T program to perform the following:

1. Run base case power flow. Create the measurements by choosing the *Use Power Flow Analysis Results* option under the *Measurement* menu. Run the WLS estimator. Can you estimate the state?

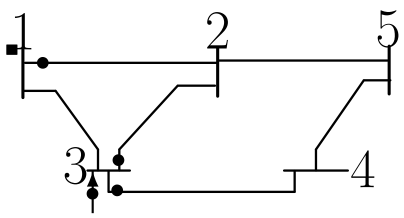
2. Calculate the number of state variables n and the total number of measurements m. Is the necessary condition ($m \geq n$) satisfied?

3. Delete the injection measurement at bus 3 and add a new injection measurement at bus 2. Make sure you update the measurements. Run the WLS estimator. Can you estimate the state now? What are n and m for this case? Discuss why the estimation is possible for this case.

4. While still in *State Estimation* mode, go to *options/program settings/simulation parameters* and choose *none* for noise. Run the WLS estimator. Go to the directory where PET is located and search for the file *wlsoutput.dat*. Look at the residuals for the measurements and comment.

5. This time choose *Gaussian* for noise and repeat the previous item. Do you see a change in the residuals for the measurements?

6. Add flow measurements on lines 2-5 and 4-5 and repeat the previous item. Do you see a change in the residuals? Discuss the results of the last three items.



- Power Measurement
- ▪ Voltage Magnitude Measurement

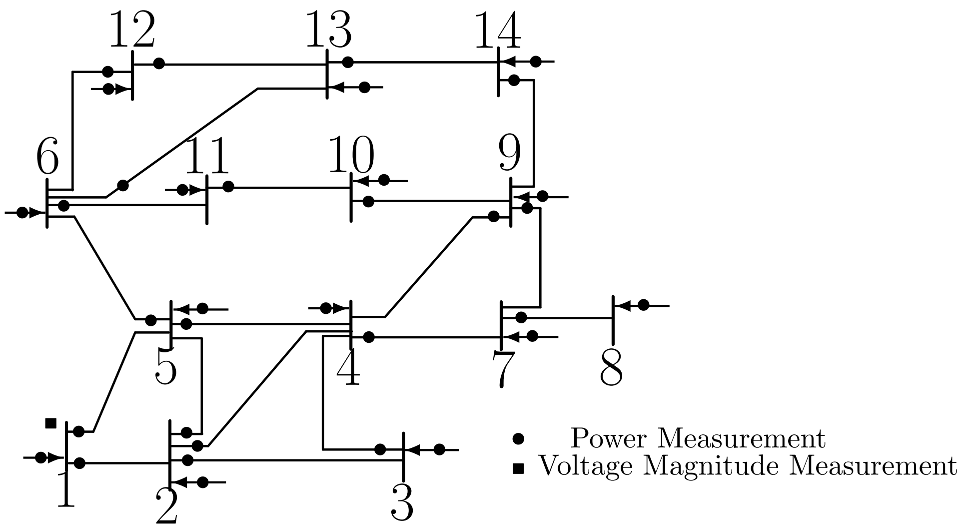
Exercise 2: [Use the file *exer02.pet* for this exercise]

This exercise illustrates one way of enforcing equality constraints (e.g. zero injections at buses) while estimating the states.

Given the network diagram and its associated measurement configuration, use the ©P.E.T program to perform the following:

1. Run base case power flow. Create the measurements and run the WLS estimator with Gaussian noise in the measurements. Edit the file *wlsoutput.dat* created by P.E.T and record the residuals for the real and reactive injection measurements at bus 7.

2. Now, set the standard deviations σ_i of the real and reactive injection measurements at bus 7 equal to 1.0×10^{-5}. Run the WLS estimator, edit the newly created output file *wlsoutput.dat* and record the residuals for the real and reactive injection measurements at bus 7. Compare them with your results from part 1.



Exercise 3: [Use the file *exer03.pet* for this exercise]

This exercise illustrates the effect of parameter errors on the state estimate and measurement residuals.

Given the IEEE 14-bus system and its measurement configuration below, do the following:

- Click on the transformer 4-9 and change the tap value to 0.9. Run the WLS estimator.
- Edit the output file *wlsoutput.dat* and look at the measurement residuals and the cost function $J(x)$ at the converged solution.

Can you tell where the parameter error is by simply looking at the output of the WLS estimator? We will discuss some possible ways of accomplishing that in the coming lectures.