

ECEN 615

Methods of Electric Power Systems Analysis

Lecture 16: Sensitivity Methods, Least Squares

Prof. Tom Overbye

Dept. of Electrical and Computer Engineering

Texas A&M University

overbye@tamu.edu



TEXAS A&M
UNIVERSITY

Announcements



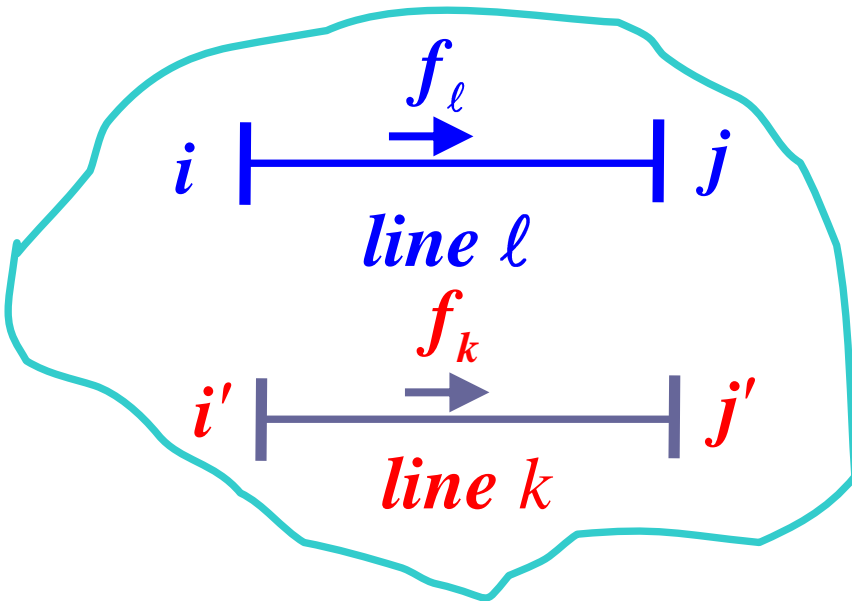
- Read Chapter 9 from the book
- Homework 4 is due on Thursday October 31.

Line Outage Distribution Factors (LODFs)

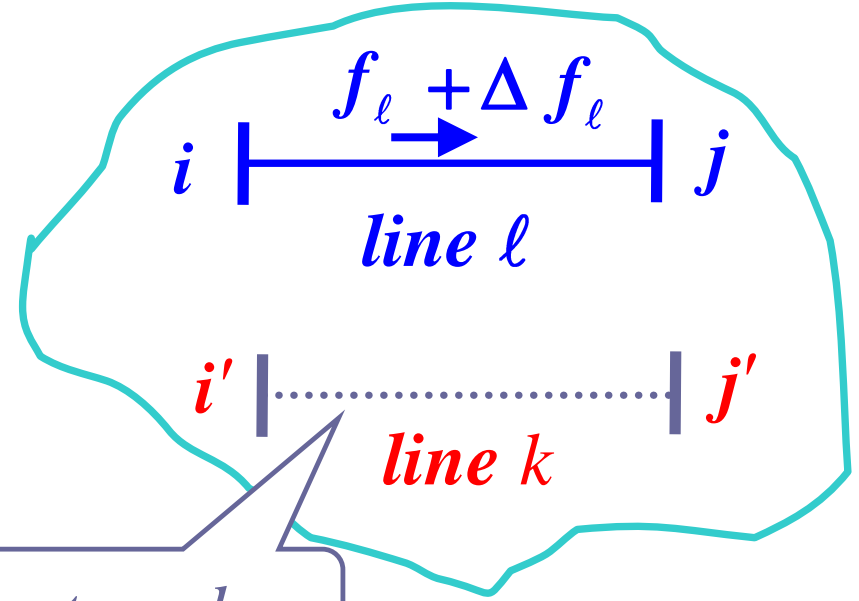


- Power system operation is practically always limited by contingencies, with line outages comprising a large number of the contingencies
- Desire is to determine the impact of a line outage (either a transmission line or a transformer) on other system real power flows without having to explicitly solve the power flow for the contingency
- These values are provided by the LODFs
- The LODF d_{ℓ}^k is the portion of the pre-outage real power line flow on line k that is redistributed to line ℓ as a result of the outage of line k

LODFs



base case



outaged

outage case

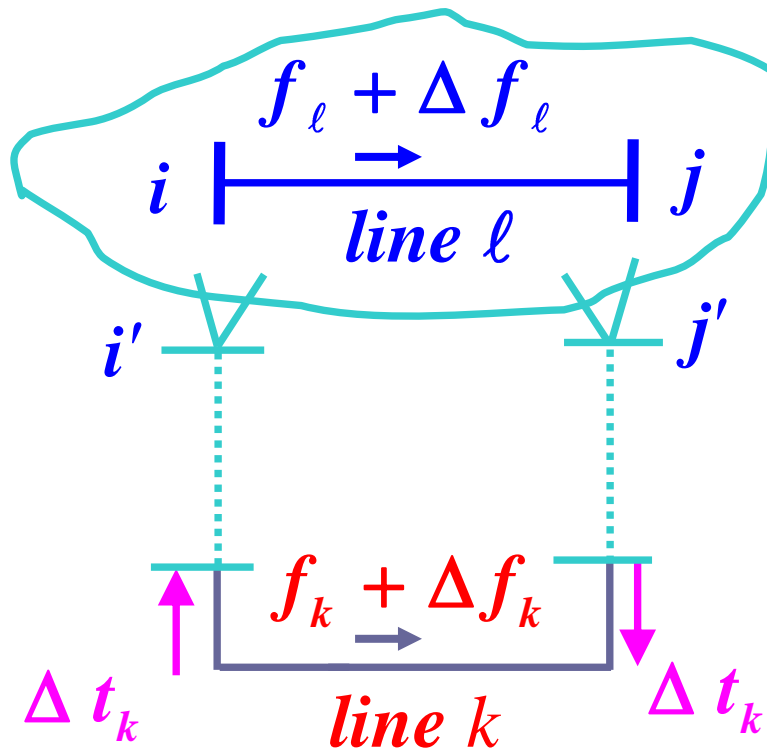
$$d_{\ell}^k = \frac{\Delta f_{\ell}}{f_k} = d_{\ell,k}$$

Best reference is Chapter 7 of the course book

LODF Evaluation



We simulate the impact of the outage of line k by adding the basic transaction $w_k = \{i', j', \Delta t_k\}$



and selecting Δt_k in such a way that the flows on the dashed lines become exactly zero

In general this Δt_k is not equal to the original line flow

LODF Evaluation



- We select Δt_k to be such that

$$f_k + \Delta f_k - \Delta t_k = 0$$

where Δf_k is the active power flow change on the line k due to the transaction w_k

- The line k flow from w_k depends on its PTDF

$$\Delta f_k = \varphi_k^{(w_k)} \Delta t_k$$

it follows that
$$\Delta t_k = \frac{f_k}{1 - \varphi_k^{(w_k)}} = \frac{f_k}{1 - (\psi_k^{i'} - \psi_k^{j'})}$$

LODF Evaluation



- For the rest of the network, the impacts of the outage of line k are the same as the impacts of the additional basic transaction w_k

$$\Rightarrow \Delta f_\ell = \varphi_\ell^{(w_k)} \Delta t_k = \frac{\varphi_\ell^{(w_k)}}{1 - \varphi_k^{(w_k)}} f_k$$

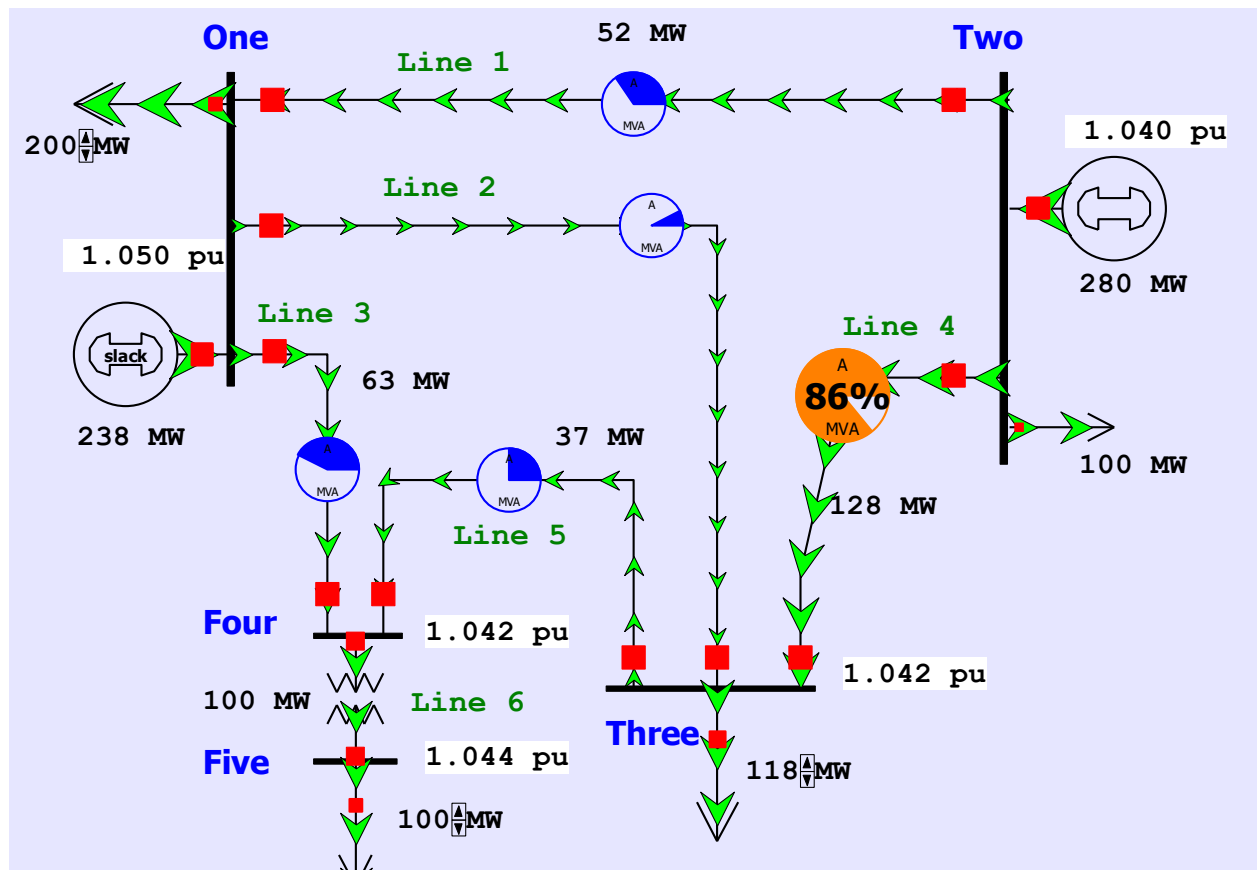
- Therefore, by definition the LODF is

$$d_\ell^k = \frac{\Delta f_\ell}{f_k} = \frac{\varphi_\ell^{(w_k)}}{1 - \varphi_k^{(w_k)}}$$

Five Bus Example



- Assume we wish to calculate the values for the outage of line 4 (between buses 2 and 3); this is line k



Say we wish to know the change in flow on the line 3 (Buses 3 to 4). PTDFs for a transaction from 2 to 3 are 0.7273 on line 4 and 0.0909 on line 3

Five Bus Example



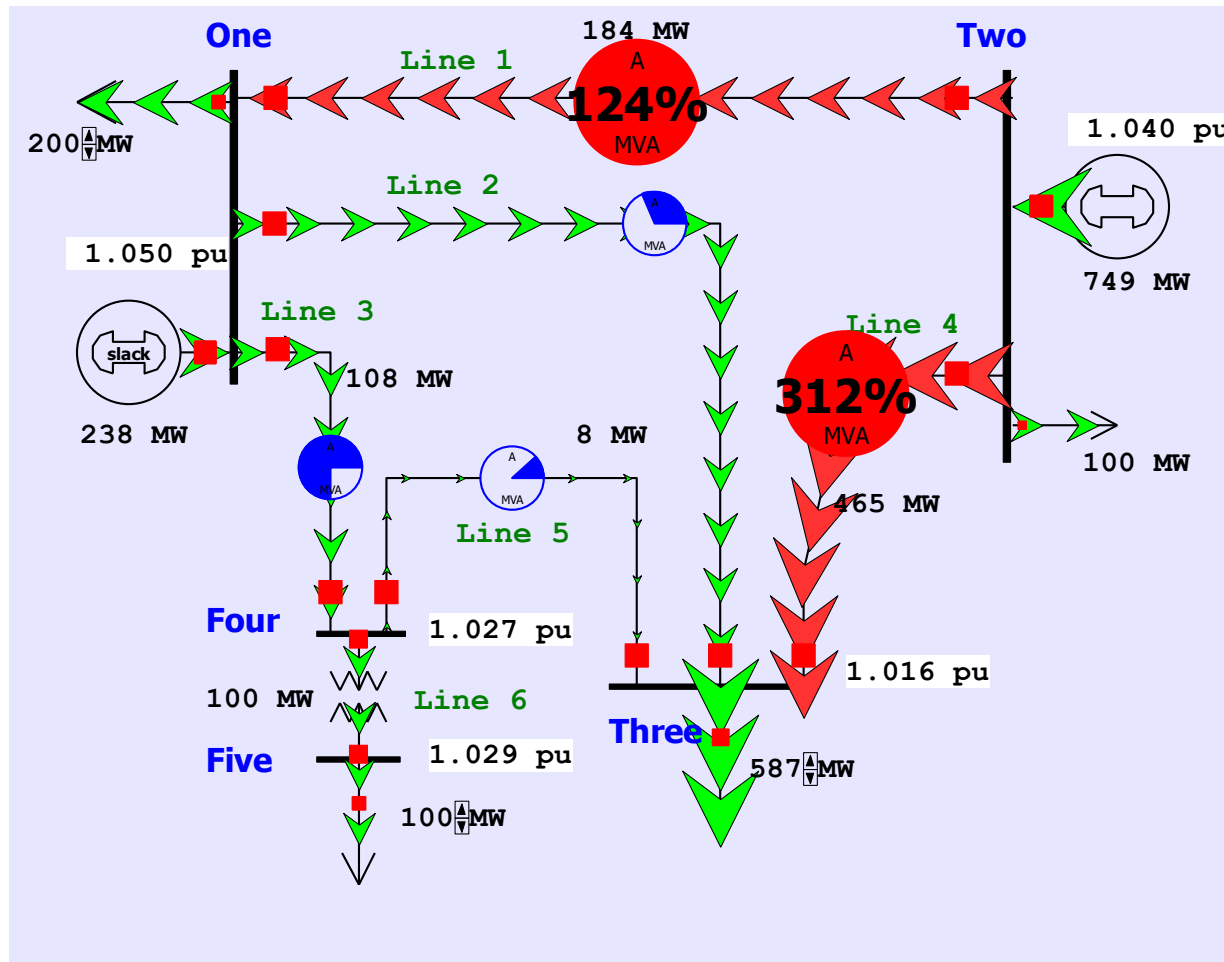
- Hence we get

$$\Delta t_k = \frac{f_k}{1 - \varphi_k^{(w_k)}} = \frac{128}{1 - 0.7273} = 469.4$$

$$d_3^4 = \frac{\Delta f_3}{f_4} = \frac{\varphi_3^{(w_4)}}{1 - \varphi_4^{(w_4)}} = \frac{0.0909}{1 - 0.7273} = 0.333$$

$$\Delta f_3 = (0.333) f_4 = 0.333 \times 128 = 42.66 \text{ MW}$$

Five Bus Example Compensated

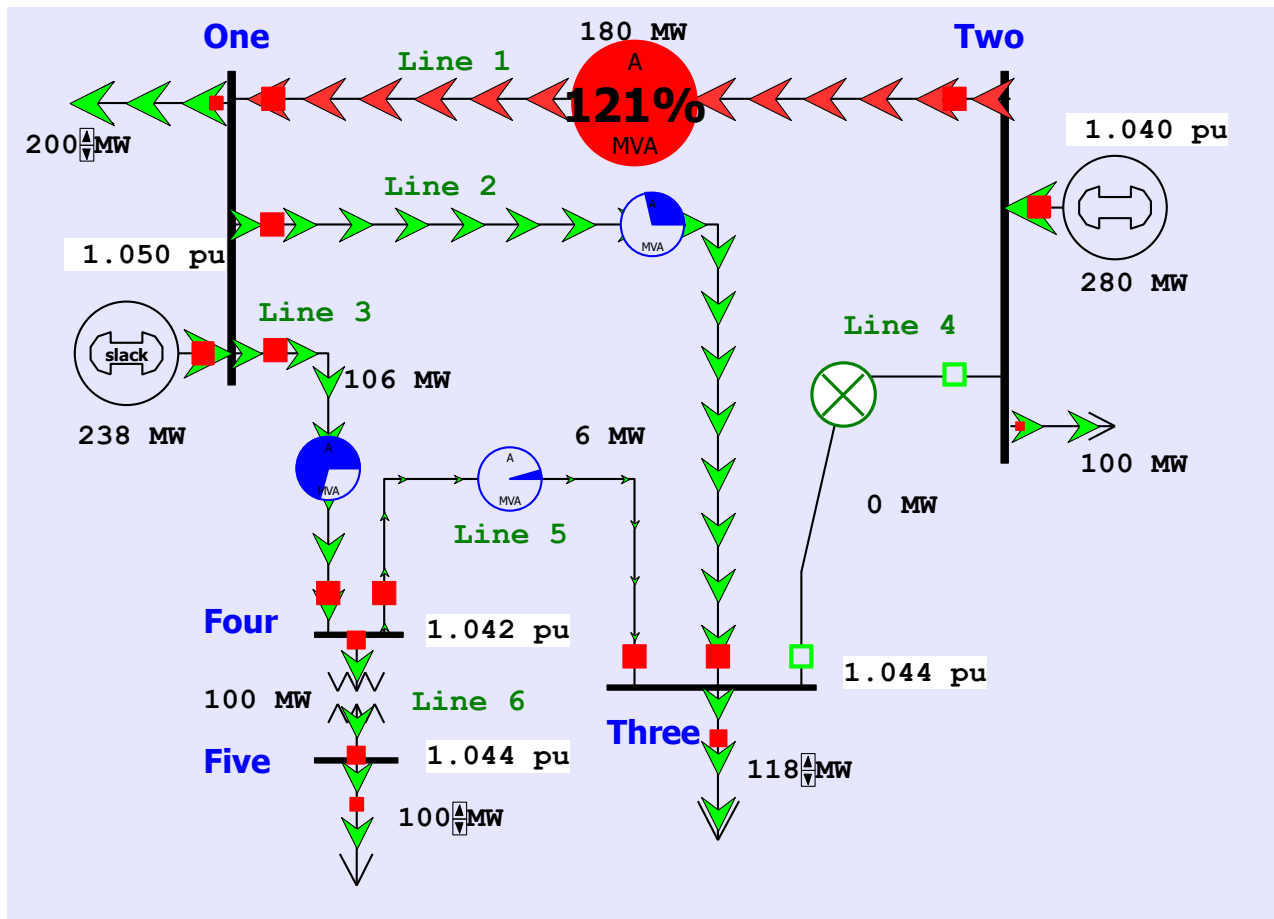


Here is the system with the compensation added to bus 2 and removed at bus 3; we are canceling the impact of the line 4 flow for the reset of the network.

Five Bus Example



- Below we see the network with the line actually outaged



The line 3 flow changed from 63 MW to 106 MW, an increase of 43 MW, matching the LODF value

Developing a Critical Eye



- In looking at the below formula you need to be thinking about what conditions will cause the formula to fail

$$\Rightarrow \Delta f_\ell = \varphi_\ell^{(w_k)} \Delta t_k = \frac{\varphi_\ell^{(w_k)}}{1 - \varphi_k^{(w_k)}} f_k$$

Here the obvious situation is when the denominator is zero

- That corresponds to a situation in which the contingency causes system islanding
 - An example is line 6 (between buses 4 and 5)
 - Impact modeled by injections at the buses within each viable island

Calculating LODFs in PowerWorld



- Select Tools, Sensitivities, Line Outage Distribution Factors
 - Select the Line using dialogs on right, and click Calculate LODFS; below example shows values for line 4

The screenshot shows the PowerWorld software interface for calculating Line Outage Distribution Factors (LODFs). The 'Tools' menu is open, and the 'Sensitivities' option is selected. The 'Line Outage Distribution Factors (LODFs) - Case: B5_DistFact_PTDF.PWB' dialog box is displayed, showing a list of lines with their respective voltages. Line 4 (Four) [138 kV] is selected. The 'Calculate LODFs' button is visible.

The 'LODFs' window shows the following data table:

	From Number	From Name	To Number	To Name	Circuit	% LODF	MW From	MW To	CTG MW From	CTG MW To
1	2	Two	1	One	1	100.0	51.6	-51.6	180.0	-180.0
2	1	One	3	Three	1	66.7	26.3	-26.3	111.9	-111.9
3	1	One	4	Four	1	33.3	63.3	-63.3	106.1	-106.1
4	2	Two	3	Three	1	-100.0	128.4	-128.4	0.0	0.0
5	4	Four	3	Three	1	33.3	-36.7	36.7	6.1	-6.1
6	5	Five	4	Four	1	0.0	-100.0	100.0	-100.0	100.0

Blackout Case LODFs



- One of the issues associated with the 8/14/03 blackout was the LODF associated with the loss of the Hanna-Juniper 345 kV line (21350-22163) that was being used in a flow gate calculation was not correct because the Chamberlin-Harding 345 kV line outage was missed
 - With the Chamberlin-Harding line assumed in-service the value was 0.362
 - With this line assumed out-of-service (which indeed it was) the value increased to 0.464

2000 Bus LODF Example



Line Outage Distribution Factors (LODFs) - Case: ECE615_2000.PWB Status: Initialized | Simulator 20

File Case Information Draw Onlines Tools Options Add Ons Window

Run Mode Log Script Power Flow Tools Simulator Options... Solve Restore Contingency Analysis... RAS + CTG Case Info Sensitivities Line Loading Replicator... Limit Monitoring... Difference Case... Scale Case... Model Explorer... Connections Other Equivalencing... Modify Case... Renumber... Edit Mode

Output Option: Single LODF LODF Matrix

Linear Calculation Method: Linearized AC Lossless DC Lossless DC With Phase Shifters

Action: Outage Sensitivities Closure Sensitivities

Line Closure Options: Calculate based on post-closure flow (LCDF) Calculate based on pre-closure flow (MLCDF)

Calculate LODFs Advanced LODF Calculation DC Model Options...

Sort by Name Number

3048 Search For Near Bus Select Far Bus, CKT

3041 (SILVER 0) [230.0 kV]	1079 (ODESSA 1 8) [500.0 kV] CKT 1
3042 (SILVER 1) [115.0 kV]	1079 (ODESSA 1 8) [500.0 kV] CKT 2
3043 (SILVER 2) [13.80 kV]	3046 (ROSCOE 5 0) [230.0 kV] CKT 1
3044 (SILVER 3) [13.80 kV]	3046 (ROSCOE 5 0) [230.0 kV] CKT 2
3045 (SILVER 4) [13.80 kV]	5045 (STEPHENVILLE 0) [500.0 kV] CKT 1
3046 (ROSCOE 5 0) [230.0 kV]	5045 (STEPHENVILLE 0) [500.0 kV] CKT 2
3047 (ROSCOE 5 1) [115.0 kV]	5120 (BROWNWOOD 0) [500.0 kV] CKT 1
3048 (ROSCOE 5 2) [500.0 kV]	5394 (ALBANY 1 0) [500.0 kV] CKT 1
3049 (ANSON 0) [115.0 kV]	
3050 (DEL RIO 0) [230.0 kV]	
3051 (DEL RIO 1) [115.0 kV]	
3052 (HUNT 0) [115.0 kV]	
3053 (WINGATE 0) [230.0 kV]	
3054 (WINGATE 1) [115.0 kV]	

LODFs Interface LODFs

	From Number	From Name	To Number	To Name	Circuit	% LODF	MW From	MW To	CTG MW From	CTG MW To
1	3048	ROSCOE 5 2	5120	BROWNWOOD 1	1	-100.0	519.5	-516.1	0.0	3.4
2	5045	STEPHENVILLE	5120	BROWNWOOD 1	1	61.6	-403.1	405.1	-83.3	85.3
3	3048	ROSCOE 5 2	5045	STEPHENVILLE 1	1	37.6	1070.1	-1057.9	1265.6	-1253.4
4	3048	ROSCOE 5 2	5045	STEPHENVILLE 2	1	37.6	1070.1	-1057.9	1265.6	-1253.4
5	5120	BROWNWOOD 0	5239	GOLDTHWAITE 1	1	-34.0	82.4	-82.3	-94.2	94.2
6	5451	COPPERAS CO	5239	GOLDTHWAITE 1	1	21.2	-907.1	912.2	-797.0	802.1
7	3048	ROSCOE 5 2	5394	ALBANY 1 0	1	14.6	-152.9	153.2	-76.8	77.1
8	5137	WACO 1 0	5388	WACO 2 0	1	-12.3	426.7	-426.3	362.9	-362.6
9	5236	OLNEY 1 0	5394	ALBANY 1 0	1	-12.2	-720.5	726.8	-784.1	790.4
10	5137	WACO 1 0	5451	COPPERAS CO 1	1	12.0	-674.5	679.6	-611.9	617.0
11	5260	GLEN ROSE 1 0	5045	STEPHENVILLE 1	1	-11.5	-1590.6	1603.3	-1650.6	1663.3
12	5239	GOLDTHWAITE	6210	MARBLE FALLS 1	1	-10.5	-808.9	816.3	-863.2	870.7
13	5358	RIESEL 1 0	5179	CORSICANA 2 1	1	-10.1	1275.3	-1266.9	1222.6	-1214.3
14	5388	WACO 2 0	5317	GRANBURY 1 0	1	-9.6	-8.1	8.1	-58.2	58.2
15	5279	TEMPLE 1 0	5358	RIESEL 1 0	1	-7.6	334.3	-333.4	294.9	-294.1
16	5410	KILLEEN 3 0	5451	COPPERAS CO 1	1	7.6	19.4	-19.4	58.6	-58.6
17	5317	GRANBURY 1 0	5260	GLEN ROSE 1 0	1	-7.5	-2609.4	2612.1	-2648.4	2651.1
18	5131	BELTON 0	5279	TEMPLE 1 0	1	-7.2	594.2	-593.4	556.8	-556.1
19	5018	JACKSBORO 1 1	5413	PALO PINTO 1 1	1	6.9	729.1	-727.3	764.9	-763.1
20	5380	ENNIS 0	5384	DALLAS 3 0	1	-6.8	1015.6	-1013.0	980.0	-977.4
21	5131	BELTON 0	5410	KILLEEN 3 0	1	6.8	313.3	-313.1	348.6	-348.4
22	5179	CORSICANA 2 1	5380	ENNIS 0	1	-6.7	911.4	-910.0	876.5	-875.0
23	5018	JACKSBORO 1 1	5236	OLNEY 1 0	2	-6.1	-691.2	693.0	-722.7	724.4
24	5018	JACKSBORO 1 1	5236	OLNEY 1 0	1	-6.1	-691.2	693.0	-722.7	724.4
25	5047	MANSFIELD 0	5179	CORSICANA 2 1	1	5.7	-313.3	313.9	-283.4	284.0
26	5055	GRAHAM 0	5018	JACKSBORO 1 1	1	-5.0	-621.5	622.8	-647.3	648.6
27	5021	ALEDO 1 0	5413	PALO PINTO 1 1	1	-4.8	-1285.7	1295.0	-1310.8	1320.2
28	5055	GRAHAM 0	5196	BRYSON 1 0	1	4.8	811.4	-809.1	836.4	-834.2
29	5196	BRYSON 1 0	5204	POOLVILLE 0	1	4.7	823.3	-819.4	847.8	-843.9
30	5361	BRIDGEPORT 0	5015	KELLER 2 0	1	4.5	1947.4	-1930.7	1970.8	-1954.1
31	5484	CROSS PLAINS	5073	BANGS 0	1	4.5	57.4	-57.0	80.6	-80.2
32	5334	CLYDE 0	5484	CROSS PLAINS	1	4.5	64.1	-63.7	87.3	-86.8
33	5121	BROWNWOOD 0	5073	BANGS 0	1	-4.5	-48.1	48.2	-71.2	71.3
34	5121	BROWNWOOD 0	5120	BROWNWOOD 0	1	4.5	-28.6	28.6	-5.4	5.4
35	5361	BRIDGEPORT 0	5204	POOLVILLE 0	1	-4.4	-757.1	757.7	-779.8	780.4
36	6107	ROCKDALE 1 0	8082	FRANKLIN 0	1	-4.3	612.7	-608.6	590.1	-586.1

LODF is for line between 3048 and 5120; values will be proportional to the PTDF values

Clear LODF Matrix Results

Run Mode

Solution Animation Stopped

AC

Viewing Present

Multiple Line LODFs



- LODFs can also be used to represent multiple device contingencies, but it is usually more involved than just adding the effects of the single device LODFs
- Assume a simultaneous outage of lines k_1 and k_2
- Now setup two transactions, w_{k_1} (with value Δt_{k_1}) and w_{k_2} (with value Δt_{k_2}) so

$$f_{k_1} + \Delta f_{k_1} + \Delta f_{k_2} - \Delta t_{k_1} = 0$$

$$f_{k_2} + \Delta f_{k_1} + \Delta f_{k_2} - \Delta t_{k_2} = 0$$

$$f_{k_1} + \varphi_{k_1}^{(w_{k_1})} \Delta t_{k_1} + \varphi_{k_1}^{(w_{k_2})} \Delta t_{k_2} - \Delta t_{k_1} = 0$$

$$f_{k_2} + \varphi_{k_2}^{(w_{k_1})} \Delta t_{k_1} + \varphi_{k_2}^{(w_{k_2})} \Delta t_{k_2} - \Delta t_{k_2} = 0$$

Multiple Line LODFs



- Hence we can calculate the simultaneous impact of multiple outages; details for the derivation are given in C.Davis, T.J. Overbye, "Linear Analysis of Multiple Outage Interaction," *Proc. 42nd HICSS*, 2009
- Equation for the change in flow on line ℓ for the outage of lines k_1 and k_2 is

$$\Delta f_{\ell} = \begin{bmatrix} d_{\ell}^{k1} & d_{\ell}^{k2} \end{bmatrix} \begin{bmatrix} \mathbf{1} & -d_{k1}^{k2} \\ -d_{k2}^{k1} & \mathbf{1} \end{bmatrix}^{-1} \begin{bmatrix} f_{k1} \\ f_{k2} \end{bmatrix}$$

Multiple Line LODFs

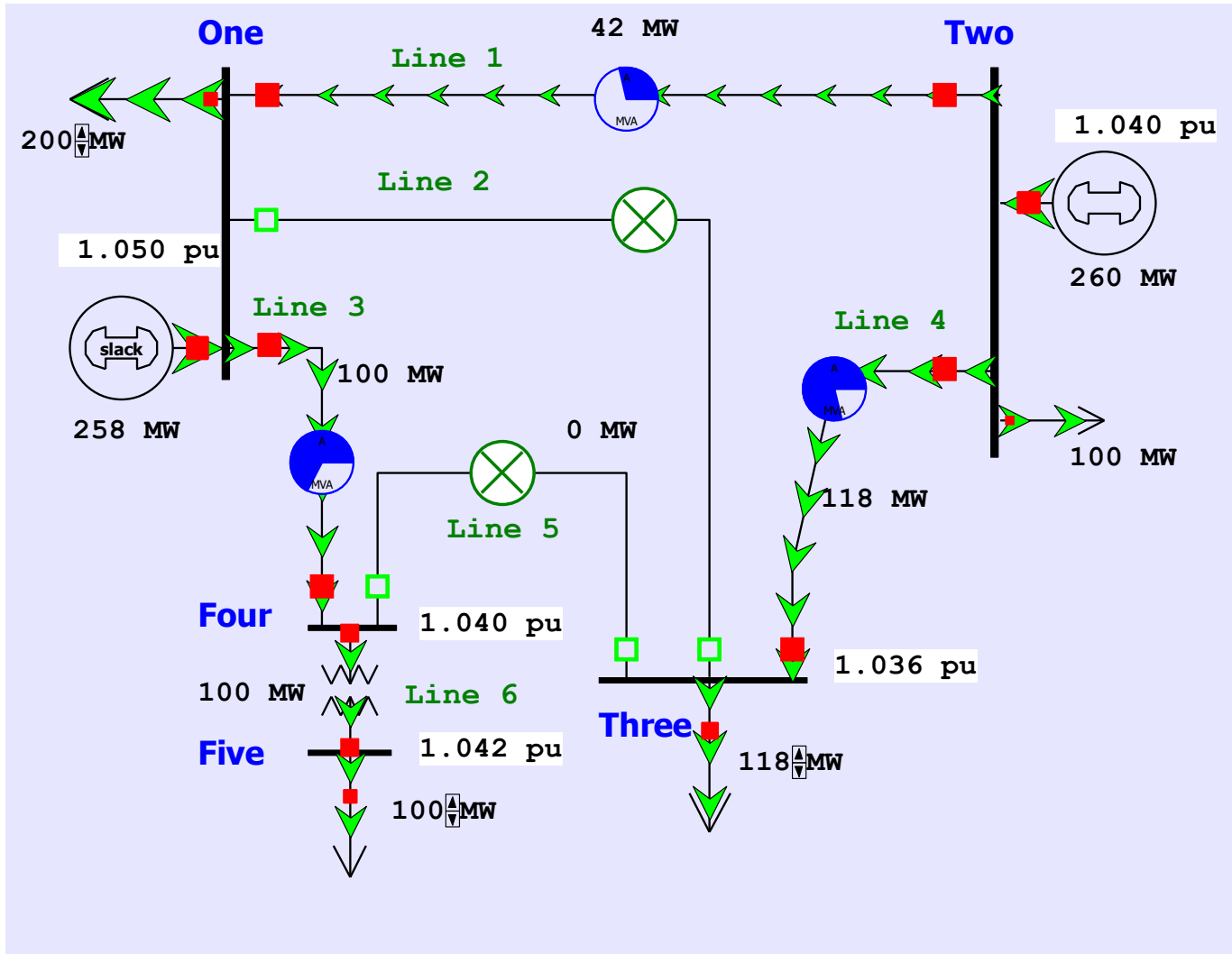


- Example: Five bus case, outage of lines 2 and 5 to flow on line 4.

$$\Delta f_\ell = \begin{bmatrix} d_\ell^{k1} & d_\ell^{k2} \end{bmatrix} \begin{bmatrix} 1 & -d_{k1}^{k2} \\ -d_{k2}^{k1} & 1 \end{bmatrix}^{-1} \begin{bmatrix} f_{k1} \\ f_{k2} \end{bmatrix}$$

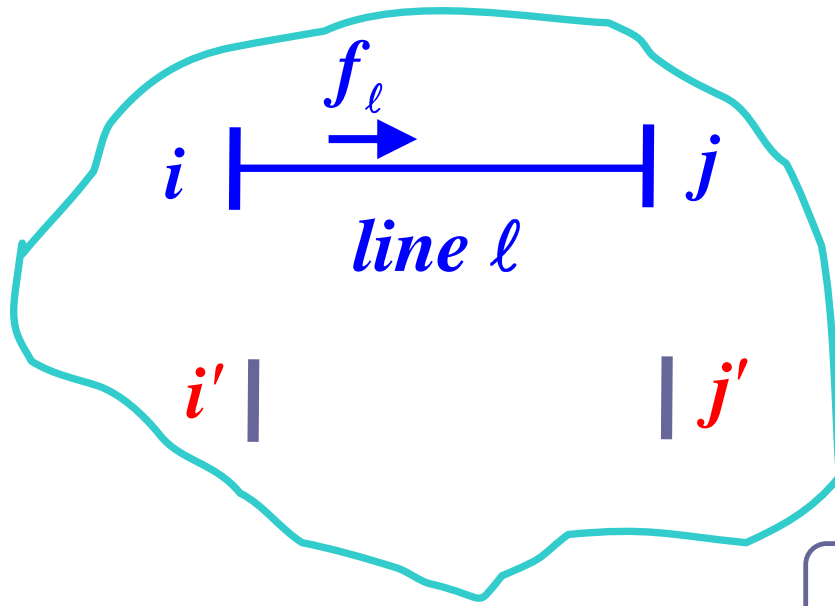
$$\Delta f_\ell = \begin{bmatrix} 0.4 & 0.25 \end{bmatrix} \begin{bmatrix} 1 & -0.75 \\ -0.6 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0.336 \\ -0.331 \end{bmatrix} = 0.005$$

Multiple Line LODFs

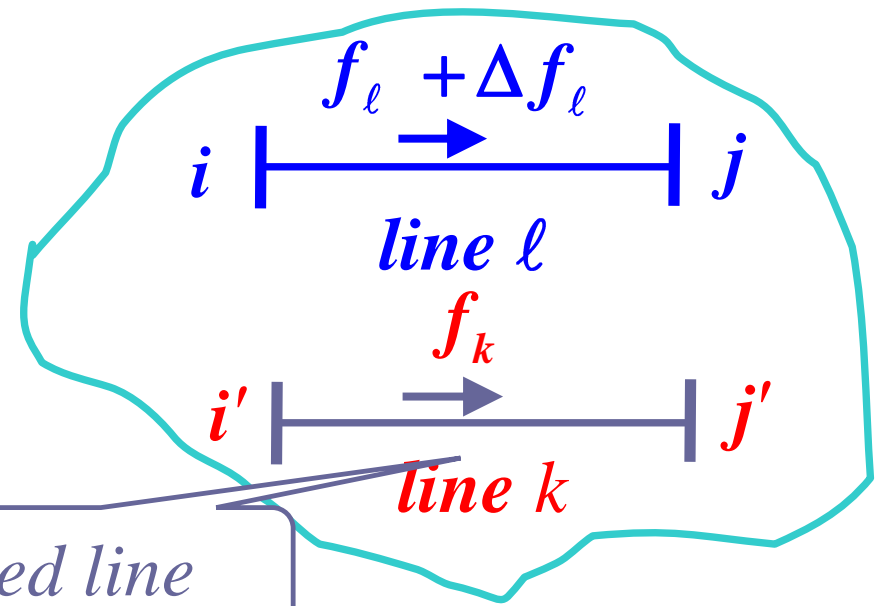


Flow goes from 117.5 to 118.0

Line Closure Distribution Factors (LCDFs)



base case



line k addition case

$$LCDF_{\ell}^k = \frac{\Delta f_{\ell}}{f_k} = LCDF_{\ell,k}$$

LCDF Definition

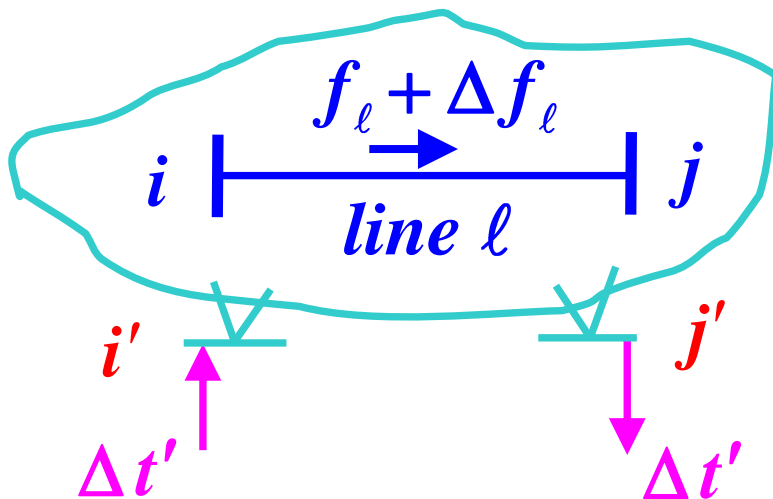


- The line closure distribution factor (LCDF), $\text{LCDF}_{\ell,k}$, for the closure of line k (or its addition if it does not already exist) is the portion of the line active power flow on line k that is distributed to line ℓ due to the closure of line k
- Since line k is currently open, the obvious question is, "what flow on line k ?"
- Answer (in a dc power flow sense) is the flow that will occur when the line is closed (which we do not know)

LCDF Evaluation

- We simulate the impact of the closure of line k by imposing the additional basic transaction

$$w_k = \{i', j', \Delta t_k\}$$



on the base case network and we select Δt_k so that

$$\Delta t_k = -f_k$$

LCDF Evaluation



- For the other parts of the network, the impacts of the addition of line k are the same as the impacts of adding the basic transaction w_k

$$\Delta f_\ell = \varphi_\ell^{(w_k)} \Delta t_k = -\varphi_\ell^{(w_k)} f_k$$

- Therefore, the definition is

$$LCDF_{\ell,k} = \frac{\Delta f_\ell}{f_k} = -\varphi_\ell^{(w_k)}$$

- The post-closure flow f_k is determined (in a dc power flow sense) as the flow that would occur from the angle difference divided by $(1 + \varphi_k^{(w_k)})$

Outage Transfer Distribution Factor

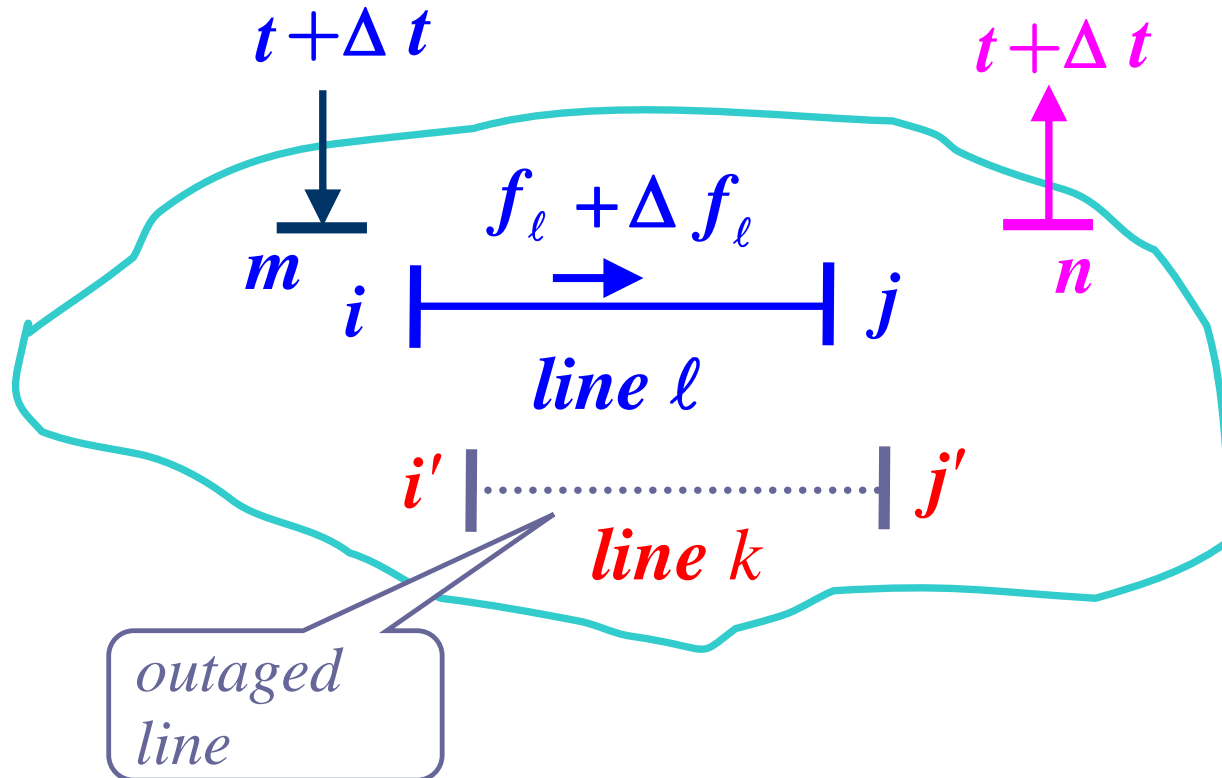


- The outage transfer distribution factor (OTDF) is defined as the PTDF with the line k outaged
- The OTDF applies only to the post-contingency configuration of the system since its evaluation explicitly considers the line k outage

$$\left(\phi_{\ell}^{(w)} \right)^k$$

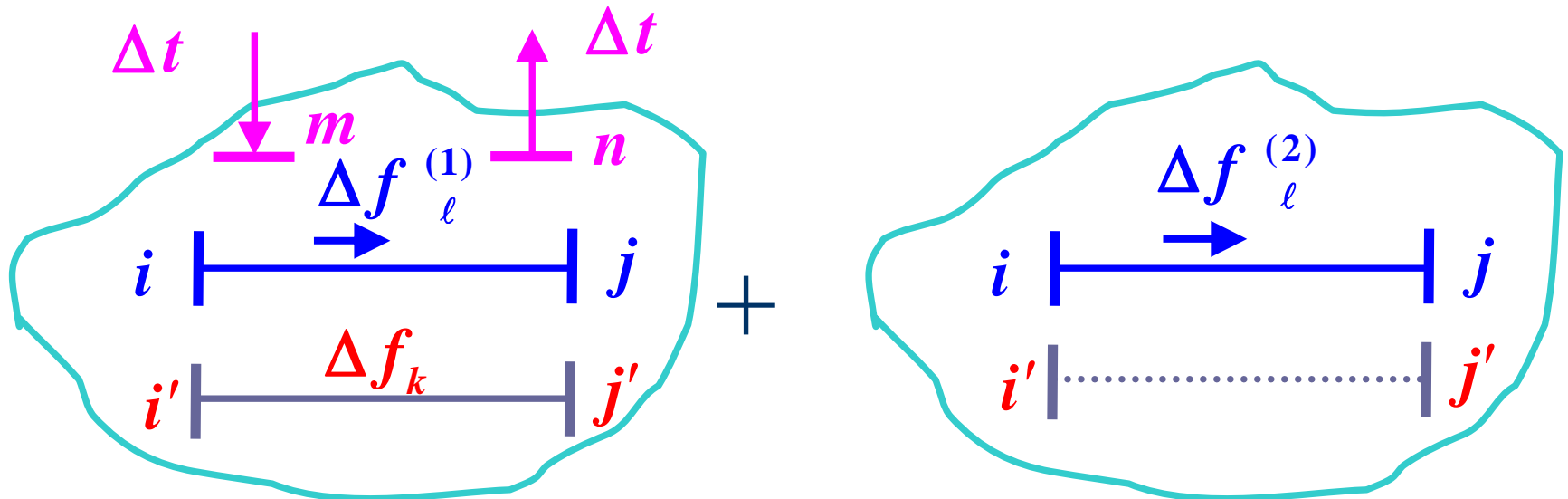
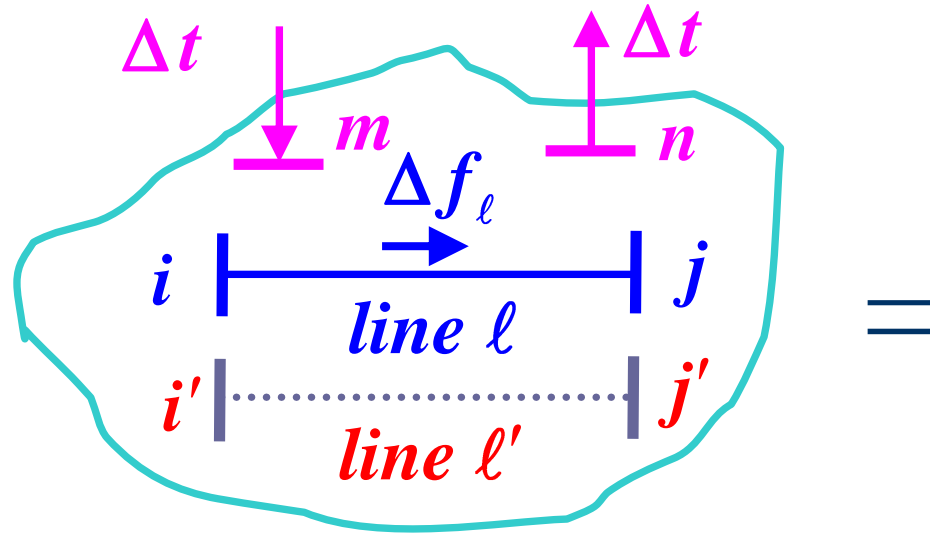
- This is a quite important value since power system operation is usually contingency constrained

Outage Transfer Distribution Factor (OTDF)



$$\left(\varphi_l^{(w)} \right)^k \triangleq \frac{\Delta f_l}{\Delta t} \Big|_{k \text{ outaged}}$$

OTDF Evaluation



OTDF Evaluation



- Since $\Delta f_{\ell}^{(1)} = \varphi_{\ell}^{(w)} \Delta t$

and $\Delta f_k = \varphi_k^{(w)} \Delta t$

then $\Delta f_{\ell}^{(2)} = d_{\ell}^k \Delta f_k = d_{\ell}^k \varphi_k^{(w)} \Delta t$

so that

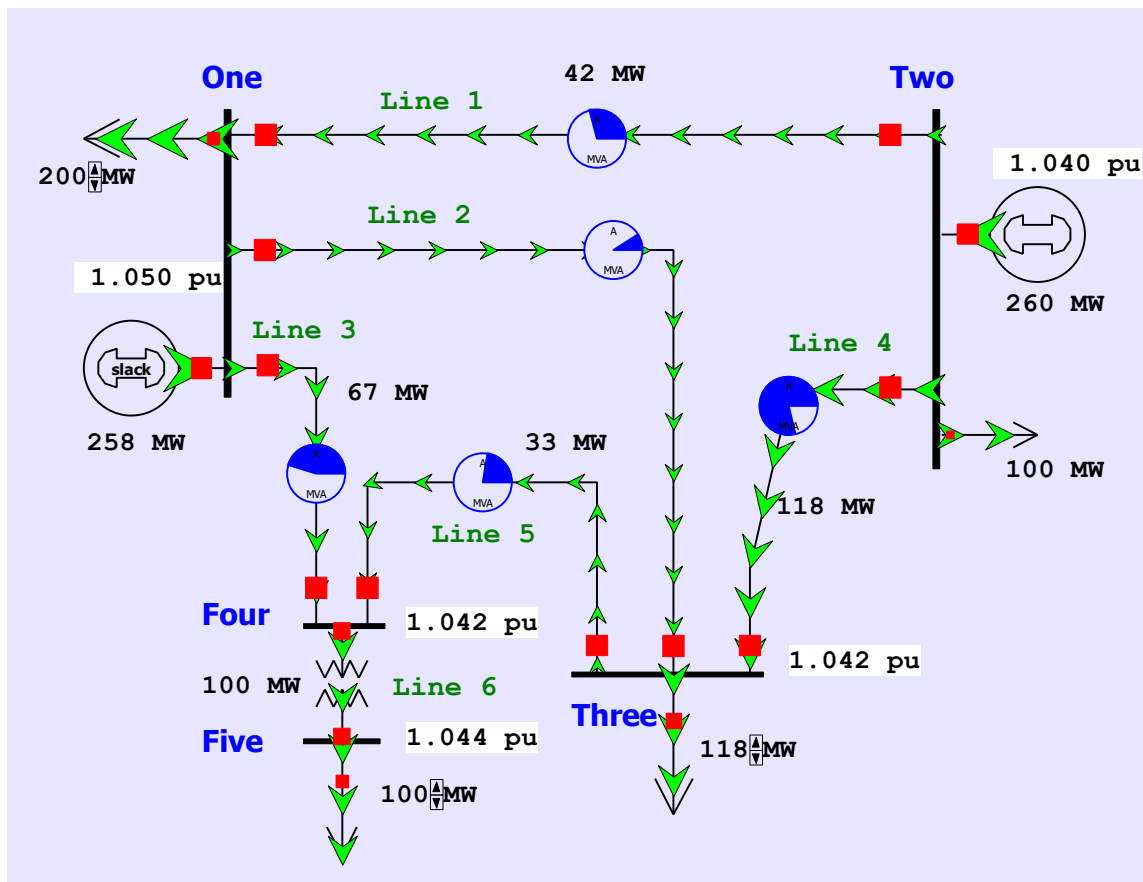
$$\Delta f_{\ell} = \Delta f_{\ell}^{(1)} + \Delta f_{\ell}^{(2)} = \left[\varphi_{\ell}^{(w)} + d_{\ell}^k \varphi_k^{(w)} \right] \Delta t$$

$$\left(\varphi_{\ell}^{(w)} \right)^k = \varphi_{\ell}^{(w)} + d_{\ell}^k \varphi_k^{(w)}$$

Five Bus Example



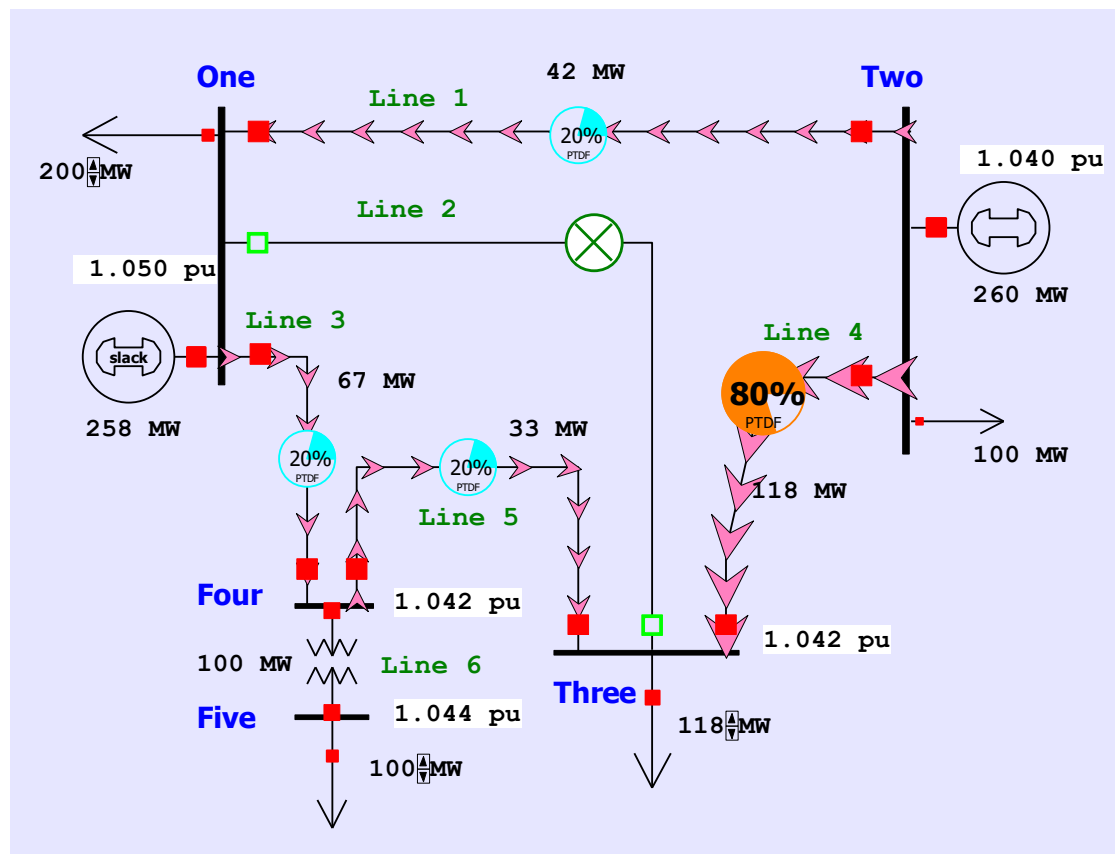
- Say we would like to know the PTDF on line 1 for a transaction between buses 2 and 3 with line 2 out



Five Bus Example



- Hence we want to calculate these values without having to explicitly outage line 2



Hence the value we are looking for is 0.2 (20%)

Five Bus Example



- Evaluating: the PTDF for the bus 2 to 3 transaction on line 1 is 0.2727; it is 0.1818 on line 2 (from buses 1 to 3); the LODF is on line 1 for the outage of line 2 is -0.4

- Hence
$$\left(\varphi_{\ell}^{(w)}\right)^k = \varphi_{\ell}^{(w)} + d_{\ell}^k \varphi_k^{(w)}$$

$$\mathbf{0.2727} + (-\mathbf{0.4}) \times (\mathbf{0.1818}) = \mathbf{0.200}$$

- For line 4 (buses 2 to 3) the value is

$$\mathbf{0.7273} + (\mathbf{0.4}) \times (\mathbf{0.1818}) = \mathbf{0.800}$$

August 14, 2003 OTDF Example



- Flowgate 2264 monitored the flow on Star-Juniper 345 kV line for contingent loss of Hanna-Juniper 345 kV normally the LODF for this flowgate is 0.361
 - flowgate had a limit of 1080 MW
 - at 15:05 EDT the flow as 517 MW on Star-Juniper, 1004 MW on Hanna-Juniper, giving a flowgate value of $520 + 0.361 * 1007 = 884$ (82%)
 - Chamberlin-Harding 345 opened at 15:05, but was missed
 - At 15:06 EDT (after loss of Chamberlin-Harding 345) #2265 had an incorrect value because its LODF was not updated.
 - Value should be $633 + 0.463 * 1174 = 1176$ (109%)
 - Value was $633 + 0.361 * 1174 = 1057$ (98%)

UTC Revisited



- We can now revisit the uncommitted transfer capability (UTC) calculation using PTDFs and LODFs
- Recall trying to determine maximum transfer between two areas (or buses in our example)
- For base case maximums are quickly determined with PTDFs

$$u_{m,n}^{(0)} = \min_{\varphi_{\ell}^{(w)} > 0} \left\{ \frac{f_{\ell}^{max} - f_{\ell}^{(0)}}{\varphi_{\ell}^{(w)}} \right\}$$

Note we are ignoring zero (or small) PTDFs; would also need to consider flow reversal

UTC Revisited



- For the contingencies we use

$$u_{m,n}^{(1)} = \min_{(\varphi_\ell^{(w)})^k > 0} \left\{ \frac{f_\ell^{max} - f_\ell^{(0)} - d_\ell^k f_k^{(0)}}{(\varphi_\ell^{(w)})^k} \right\}$$

- Then as before $u_{m,n} = \min \{ u_{m,n}^{(0)}, u_{m,n}^{(1)} \}$

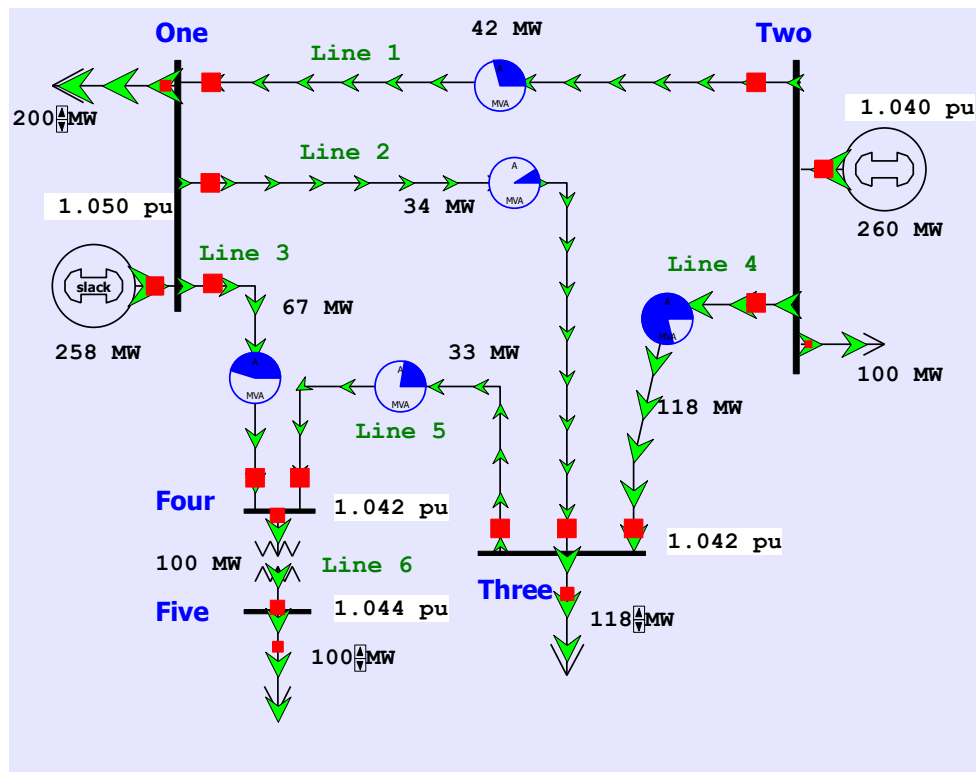
We would need to check all contingencies! Also, this is just a linear estimate and is not considering voltage violations.

Five Bus Example



$$w = \{2, 3, \Delta t\} \quad \mathbf{f}^{(0)} = [42, 34, 67, 118, 33, 100]^T$$

$$\mathbf{f}^{max} = [150, 400, 150, 150, 150, 1,000]^T$$



Five Bus Example



Therefore, for the base case

$$\begin{aligned} u_{2,2}^{(0)} &= \min_{\varphi_l^{(w)} > 0} \left\{ \frac{f_l^{max} - f_l^{(0)}}{\varphi_l^{(w)}} \right\} \\ &= \min \left\{ \frac{150 - 42}{0.2727}, \frac{400 - 34}{0.1818}, \frac{150 - 67}{0.0909}, \frac{150 - 118}{0.7273}, \frac{150 - 33}{0.0909} \right\} \\ &= 44.0 \end{aligned}$$

Five Bus Example



- For the contingency case corresponding to the outage of the line 2

$$u_{2,3}^{(1)} = \min_{(\varphi_{\ell}^{(w)})^2 > 0} \left\{ \frac{f_{\ell}^{max} - f_{\ell}^{(0)} - d_{\ell}^2 f_2^{(0)}}{(\varphi_{\ell}^{(w)})^2} \right\}$$

The limiting value is line 4

$$\frac{f_{\ell}^{max} - f_{\ell}^{(0)} - d_{\ell}^2 f_2^{(0)}}{(\varphi_{\ell}^{(w)})^2} = \frac{150 - 118 - 0.4 \times 34}{0.8}$$

Hence the UTC is limited by the contingency to 23.0

Additional Comments



- Distribution factors are defined as small signal sensitivities, but in practice, they are also used for simulating large signal cases
- Distribution factors are widely used in the operation of the electricity markets where the rapid evaluation of the impacts of each transaction on the line flows is required
- Applications to actual system show that the distribution factors provide satisfactory results in terms of accuracy
- For multiple applications that require fast turn around time, distribution factors are used very widely, particularly, in the market environment
- They do not work well with reactive power!

Least Squares



- So far we have considered the solution of $\mathbf{Ax} = \mathbf{b}$ in which \mathbf{A} is a square matrix; as long as \mathbf{A} is nonsingular there is a single solution
 - That is, we have the same number of equations (m) as unknowns (n)
- Many problems are overdetermined in which there are more equations than unknowns ($m > n$)
 - Overdetermined systems are usually inconsistent, in which no value of \mathbf{x} exactly solves all the equations
- Underdetermined systems have more unknowns than equations ($m < n$); they never have a unique solution but are usually consistent

Method of Least Squares



- The least squares method is a solution approach for determining an approximate solution for an overdetermined system
- If the system is inconsistent, then not all of the equations can be exactly satisfied
- The difference for each equation between its exact solution and the estimated solution is known as the error
- Least squares seeks to minimize the sum of the squares of the errors
- Weighted least squares allows different weights for the equations

Least Squares Solution History



- The method of least squares developed from trying to estimate actual values from a number of measurements
- Several persons in the 1700's, starting with Roger Cotes in 1722, presented methods for trying to decrease model errors from using multiple measurements
- Legendre presented a formal description of the method in 1805; evidently Gauss claimed he did it in 1795
- Method is widely used in power systems, with state estimation the best known application, dating from Fred Schweppe's work in 1970

Least Squares and Sparsity



- In many contexts least squares is applied to problems that are not sparse. For example, using a number of measurements to optimally determine a few values
 - Regression analysis is a common example, in which a line or other curve is fit to potentially many points)
 - Each measurement impacts each model value
- In the classic power system application of state estimation the system is sparse, with measurements only directly influencing a few states
 - Power system analysis classes have tended to focus on solution methods aimed at sparse systems; we'll consider both sparse and nonsparse solution methods

Least Squares Problem



- Consider $\mathbf{Ax} = \mathbf{b}$ $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$

or

$$\begin{bmatrix} (\mathbf{a}^1)^T \\ (\mathbf{a}^2)^T \\ \vdots \\ (\mathbf{a}^m)^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Least Squares Solution



- We write $(\mathbf{a}^i)^T$ for the row i of \mathbf{A} and \mathbf{a}^i is a column vector
- Here, $m \geq n$ and the solution we are seeking is that which minimizes $\|\mathbf{Ax} - \mathbf{b}\|_p$, where p denotes some norm
- Since usually an overdetermined system has no exact solution, the best we can do is determine an \mathbf{x} that minimizes the desired norm.

Choice of p



- We discuss the choice of p in terms of a specific example
- Consider the equation $\mathbf{Ax} = \mathbf{b}$ with

$$\mathbf{A} = \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{with } b_1 \geq b_2 \geq b_3 \geq 0$$

(hence three equations and one unknown)

- We consider three possible choices for p :

Choice of p



(i) $p = 1$

$\|\mathbf{Ax} - \mathbf{b}\|_1$ is minimized by $x^* = b_2$

(ii) $p = 2$

$\|\mathbf{Ax} - \mathbf{b}\|_2$ is minimized by $x^* = \frac{b_1 + b_2 + b_3}{3}$

(iii) $p = \infty$

$\|\mathbf{Ax} - \mathbf{b}\|_\infty$ is minimized by $x^* = \frac{b_1 + b_3}{2}$