

Big Data Technology and its Applications

Topology Identification and
Line Parameter Estimation

张宁 ningzhang@tsinghua.edu.cn

Contents

Background

Balanced network

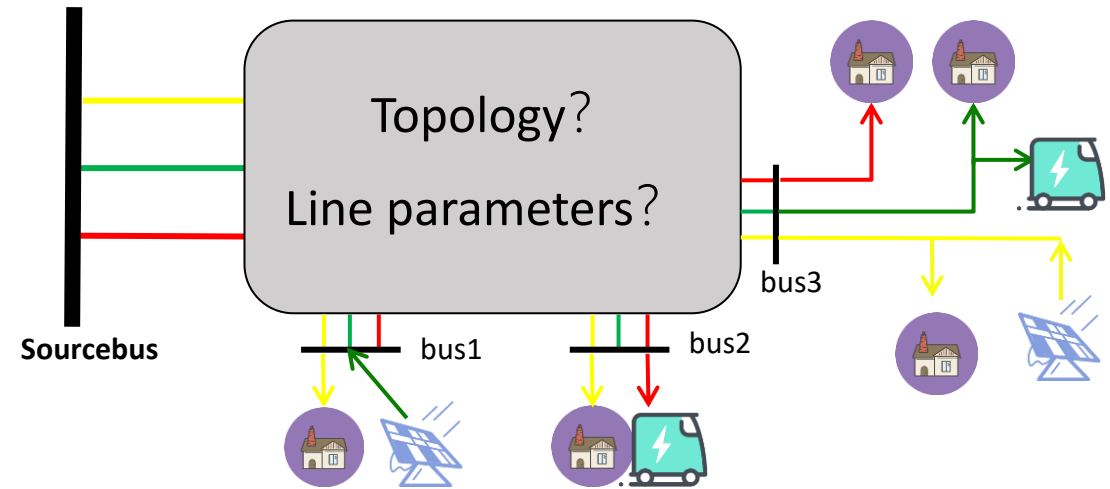
Three phase unbalanced network (Self-study)

Test Cases

Conclusions

Background

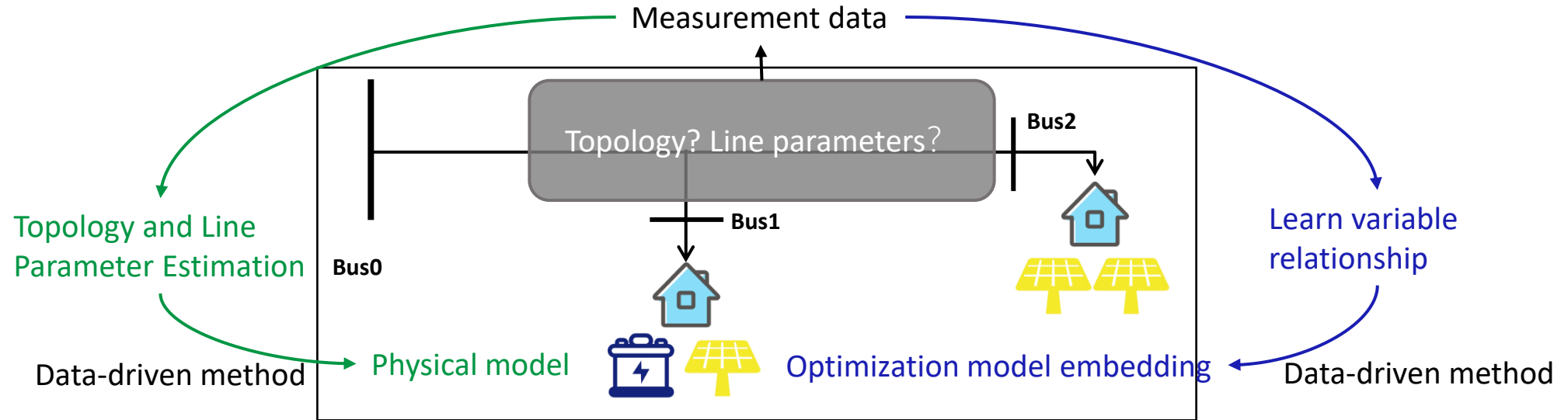
- **Background:** The operation and analysis of distribution network requires the information of system topology and line parameter.



- However, the real-time information of topology and line parameter, especially branch's conductance and susceptance **may not be available** in distribution networks, since there is fewer monitoring devices for distribution network than those in transmission network.
- **Aim:** identify the topology, estimate line parameter and recover missing voltage angle at the same time **without the measurement of voltage angle**.

Background

- Data-driven method based on measurement helps to realize the fine regulation of distribution network



Smart meter, PMU, DTU, TTU... Widely connected to the distribution network

Topology and Line
Parameter Estimation

$$\hat{P}_i^t = \hat{V}_i^t \sum_{j=1}^N \hat{V}_j^t \left(G_{ij} \cos \hat{\theta}_{ij}^t + B_{ij} \sin \hat{\theta}_{ij}^t \right)$$

Learn variable relationship from data

Background

- Comparing with power flow calculation and state estimation, we can better understand *Topology and Line Parameter Estimation*

Known variables

$$\hat{P}_i^t = \hat{V}_i^t \sum_{j=1}^N \hat{V}_j^t \left(G_{ij} \cos \hat{\theta}_{ij}^t + B_{ij} \sin \hat{\theta}_{ij}^t \right)$$

Power flow calculation

Unknown variables

$$\hat{P}_i^t = \hat{V}_i^t \sum_{j=1}^N \hat{V}_j^t \left(G_{ij} \cos \hat{\theta}_{ij}^t + B_{ij} \sin \hat{\theta}_{ij}^t \right)$$

State estimation

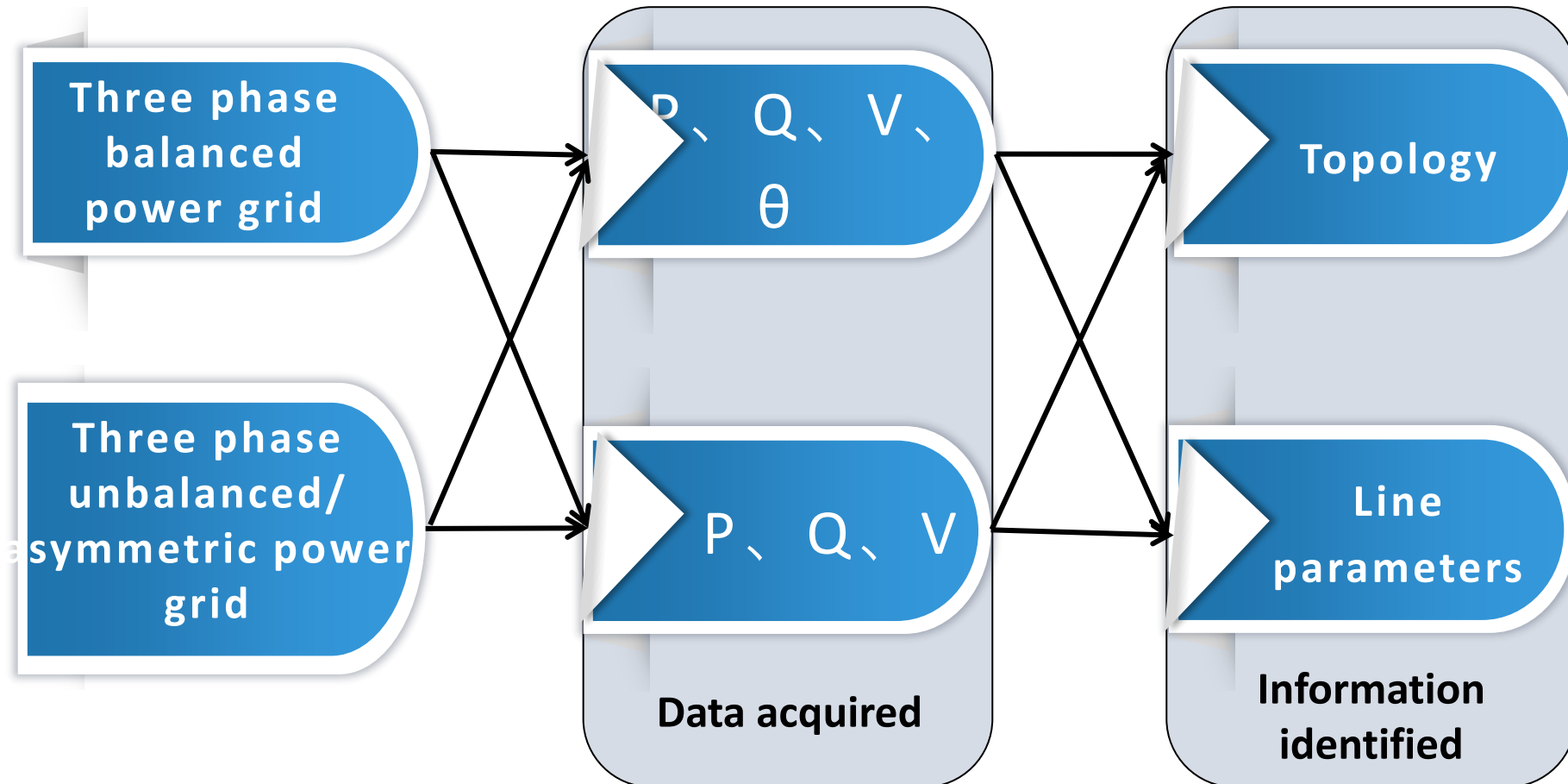
$$\hat{P}_i^t = \hat{V}_i^t \sum_{j=1}^N \hat{V}_j^t \left(G_{ij} \cos \hat{\theta}_{ij}^t + B_{ij} \sin \hat{\theta}_{ij}^t \right)$$

Topology and Line
Parameter Estimation

- Estimation of the distribution network's topology and line parameter is equivalent to estimate matrix G_{ij} , B_{ij} with various power flow and voltage data.

Background

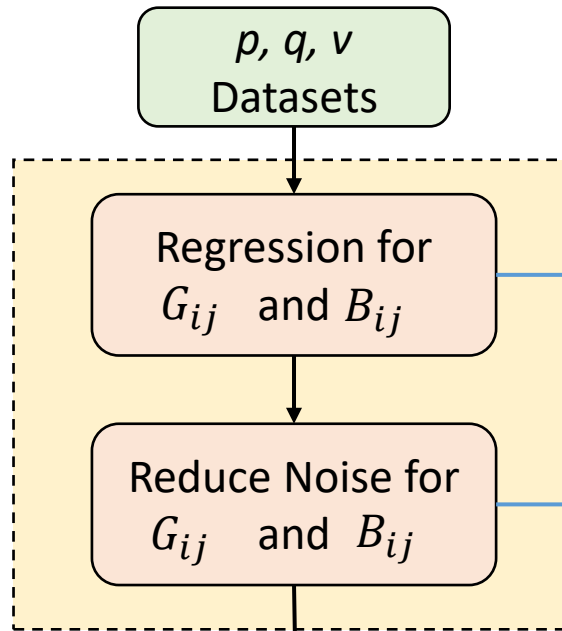
- Structure of three phase balanced or unbalanced power grid Topology and Line Parameter Identification



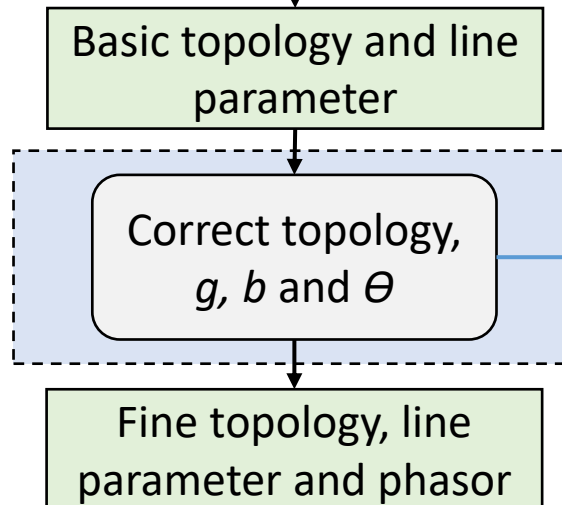
Balanced network: Framework

A two-step framework

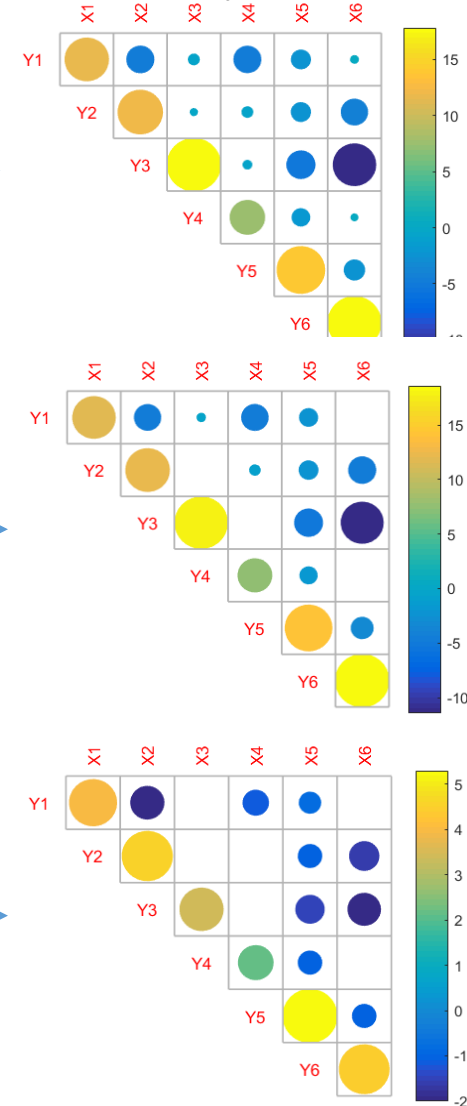
Step 1:
Basic Identification



Step 2:
Fine Identification



G_{ij} after different processes
(6-bus test case)



Balanced network: Step 1 Basic Identification

- **Basic identification** aims to provide approximate topology and line parameter for fine identification.

$$\begin{cases} p_i = \sum_{j=1}^n v_i v_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \\ q_i = \sum_{j=1}^n v_i v_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \end{cases}$$

Approximation

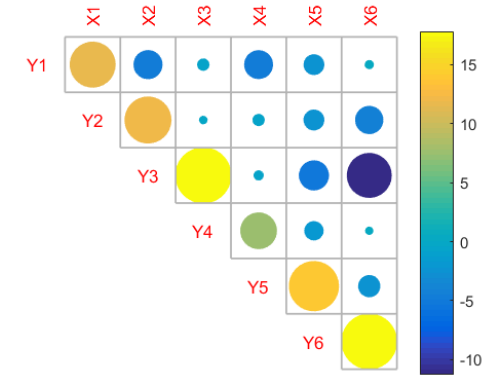
$$\begin{cases} G_{ij}^{\#} = G_{ij} + \theta_{ij} B_{ij} \\ B_{ij}^{\#} = B_{ij} - \theta_{ij} G_{ij} \end{cases}$$

Regression equations

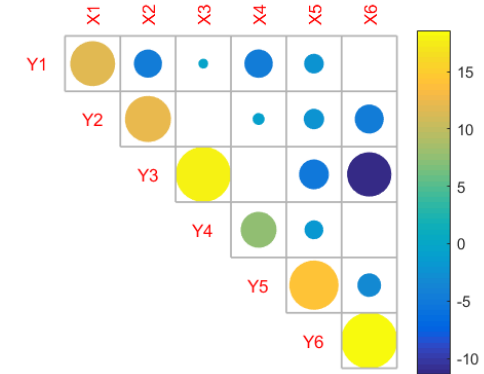
$$\begin{aligned} [P_i/V_i] &= [G_{ij}^{\#}] \cdot [V_j] \\ [Q_i/V_i] &= -[B_{ij}^{\#}] \cdot [V_j] \end{aligned}$$

Symmetrization

$$\begin{aligned} [G_{ij_S}^{\#}] &= ([G_{ij}^{\#}] + [G_{ij}^{\#T}])/2 \\ [B_{ij_S}^{\#}] &= ([B_{ij}^{\#}] + [B_{ij}^{\#T}])/2 \end{aligned}$$



G_{ij} matrix after regression



G_{ij} matrix after reducing noise

Balanced network: Step 2 Fine Identification

- After the basic identification, the **fine identification** further calculates the more accurate topology and parameters.

A. Newton-Raphson Method

0. Using known p, q to solve g, b, θ , we can apply Newton-Raphson Method.

$$\begin{bmatrix} \Delta p \\ \Delta q \end{bmatrix}_{[1 \times 2n]} = \begin{bmatrix} \frac{\partial p}{\partial g} & \frac{\partial p}{\partial b} & \frac{\partial p}{\partial \theta} \\ \frac{\partial q}{\partial g} & \frac{\partial q}{\partial b} & \frac{\partial q}{\partial \theta} \end{bmatrix} \cdot \begin{bmatrix} \Delta g \\ \Delta b \\ \Delta \theta \end{bmatrix}_{[1 \times (2m+n-1)]}$$

m: number of branches
n: number of buses

1. The amount of variables ($2m+n-1$) is larger than constraints ($2n$). Multiple samples would further reduce the error from wrong measurements.

M-Sample Newton-Raphson Method

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix}_{[1 \times 2M \cdot n]} = \begin{bmatrix} \frac{\partial P}{\partial g} & \frac{\partial P}{\partial b} & \frac{\partial P}{\partial \theta} \\ \frac{\partial Q}{\partial g} & \frac{\partial Q}{\partial b} & \frac{\partial Q}{\partial \theta} \end{bmatrix} \cdot \begin{bmatrix} \Delta g \\ \Delta b \\ \Delta \theta \end{bmatrix}_{[1 \times (2m+M \cdot (n-1))]}$$

M: Number of samples

Balanced network: Step 2 Fine Identification

A. Newton-Raphson Method

2. Jacobian matrices built from M samples are usually **not square**. To solve that problem, we can get the unique optimal approximation solution using Penrose-Moore generalized inverse.

generalized inverse

$$\begin{bmatrix} \Delta \mathbf{g} \\ \Delta \mathbf{b} \\ \Delta \boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{P}}{\partial \mathbf{g}} & \frac{\partial \mathbf{P}}{\partial \mathbf{b}} & \frac{\partial \mathbf{P}}{\partial \boldsymbol{\theta}} \\ \frac{\partial \mathbf{Q}}{\partial \mathbf{g}} & \frac{\partial \mathbf{Q}}{\partial \mathbf{b}} & \frac{\partial \mathbf{Q}}{\partial \boldsymbol{\theta}} \end{bmatrix}^{\dagger} \cdot \begin{bmatrix} \Delta \mathbf{P} \\ \Delta \mathbf{Q} \end{bmatrix}$$

3. The line parameter and voltage angel are renewed in each iteration.

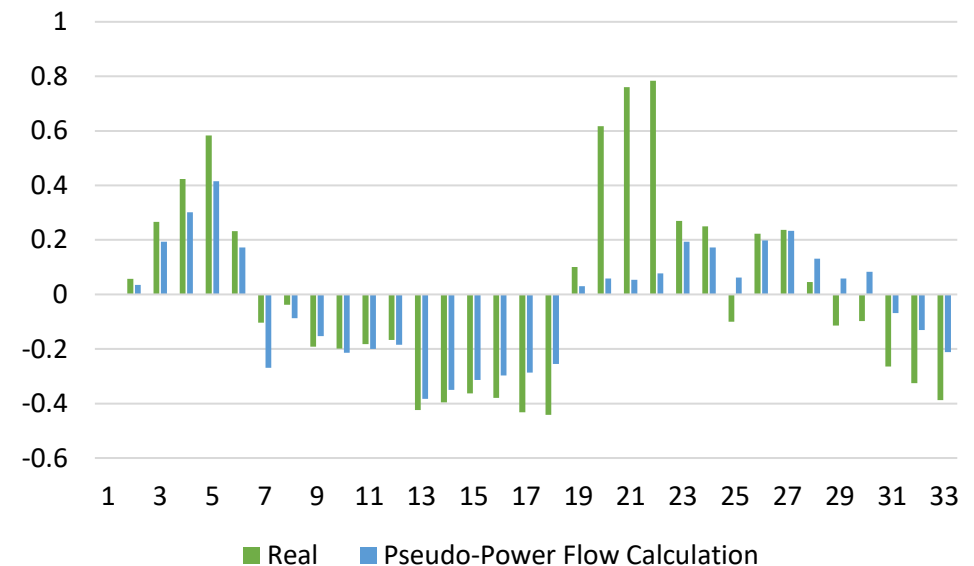
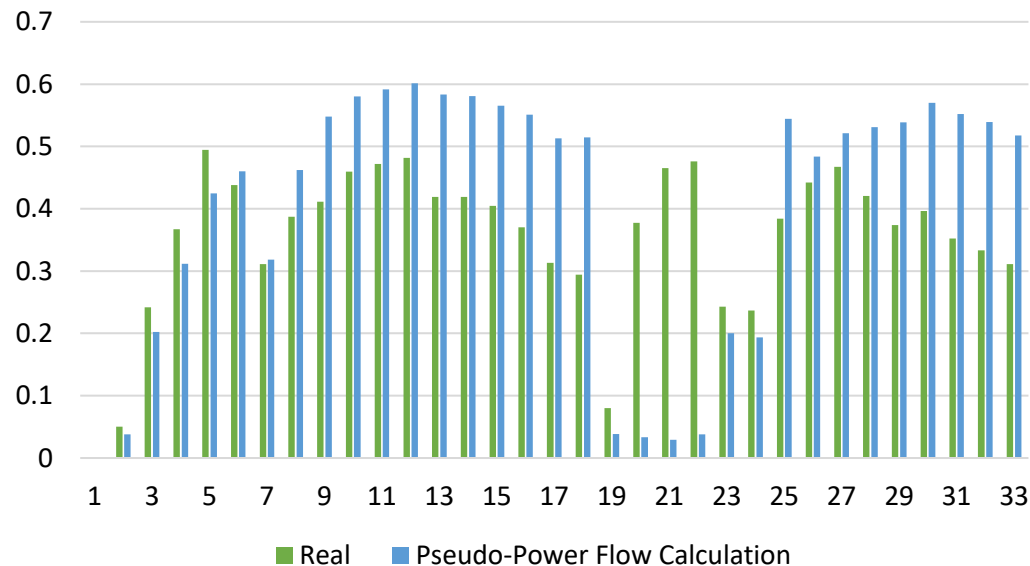
$$\begin{bmatrix} \mathbf{g} \\ \mathbf{b} \\ \boldsymbol{\theta} \end{bmatrix}^{(k+1)} = \begin{bmatrix} \mathbf{g} \\ \mathbf{b} \\ \boldsymbol{\theta} \end{bmatrix}^{(k)} + \begin{bmatrix} \Delta \mathbf{g} \\ \Delta \mathbf{b} \\ \Delta \boldsymbol{\theta} \end{bmatrix}$$

Another problem: To start the iteration , how to select the initial value of \mathbf{g} , \mathbf{b} , $\boldsymbol{\theta}$?

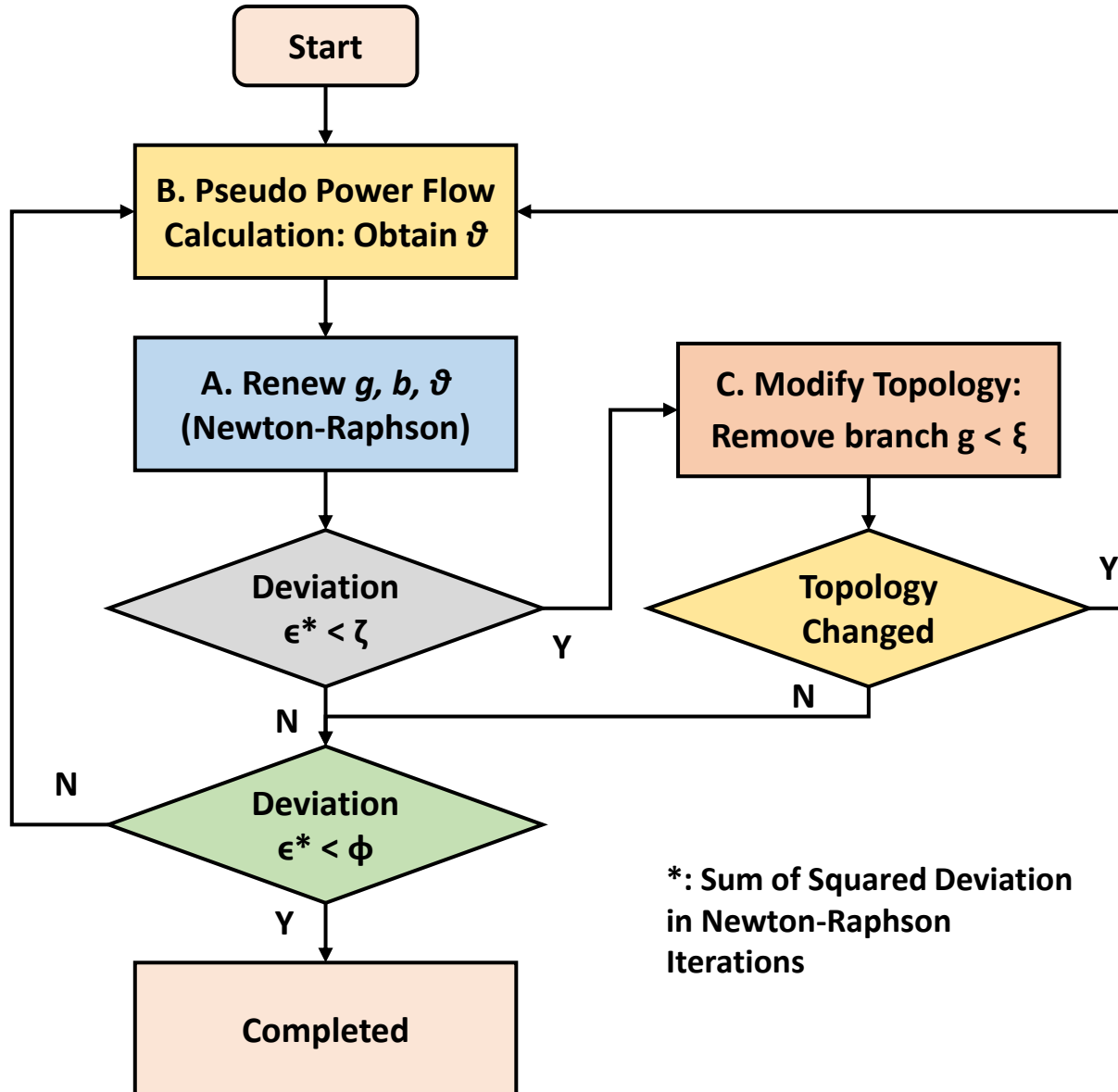
Balanced network: Step 2 Fine Identification

B. Pseudo-Power Flow Calculation

- The key idea of Pseudo-Power Flow Calculation is to regard all the buses (except reference bus) as PQ nodes and **calculate missing voltage angle θ** , using the estimated g , b and measured p , q .
- Pseudo-Power Flow Calculation provides **good initial value** for fine identification.
- Two example of voltage angle: IEEE 33 bus test case (35% Renewables)



Balanced network: Step 2 Fine Identification



*: Sum of Squared Deviation
in Newton-Raphson
Iterations

A. Newton-Raphson Method

B. Pseudo Power Flow Calculation

C. Modify Topology:

Remove small g branches

**ζ (\zetaeta): Avoid wrong topology
modification**

**ϕ : Convergence condition of
Newton-Raphson Method**

Contents

Background

Balanced network

Three phase unbalanced network (Self-study)

Test Cases

Conclusions

Unbalanced network: Framework

- Power flow equation for unbalanced power grid:

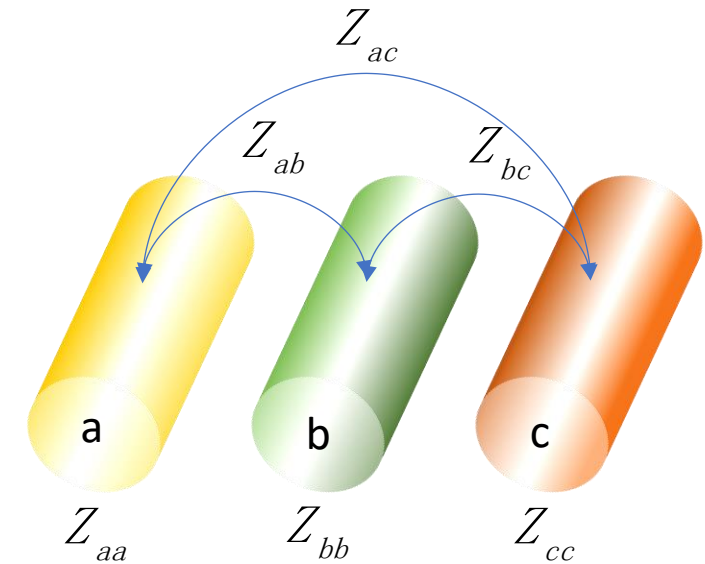
$$\begin{cases} P_i^\alpha = V_i^\alpha \sum_{j=1}^N \sum_{\beta=a,b,c} V_j^\beta (G_{ij}^{\alpha\beta} \cos \theta_{ij}^{\alpha\beta} + B_{ij}^{\alpha\beta} \sin \theta_{ij}^{\alpha\beta}) \\ Q_i^\alpha = V_i^\alpha \sum_{j=1}^N \sum_{\beta=a,b,c} V_j^\beta (G_{ij}^{\alpha\beta} \sin \theta_{ij}^{\alpha\beta} - B_{ij}^{\alpha\beta} \cos \theta_{ij}^{\alpha\beta}) \end{cases}$$

$$\begin{bmatrix} G_{11} & G_{12} & \cdots & G_{1N} \\ G_{21} & G_{22} & \cdots & G_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ G_{N1} & G_{N2} & \cdots & G_{NN} \end{bmatrix}$$

Conductance matrix

$$\begin{bmatrix} G_{11}^{aa} & G_{11}^{ab} & G_{11}^{ac} \\ G_{11}^{ba} & G_{11}^{bb} & G_{11}^{cb} \\ G_{11}^{ca} & G_{11}^{cb} & G_{11}^{cc} \end{bmatrix}$$

$$\begin{bmatrix} G_{21}^{aa} & G_{21}^{ab} & G_{21}^{ac} \\ G_{21}^{ba} & G_{21}^{bb} & G_{21}^{cb} \\ G_{21}^{ca} & G_{21}^{cb} & G_{21}^{cc} \end{bmatrix}$$

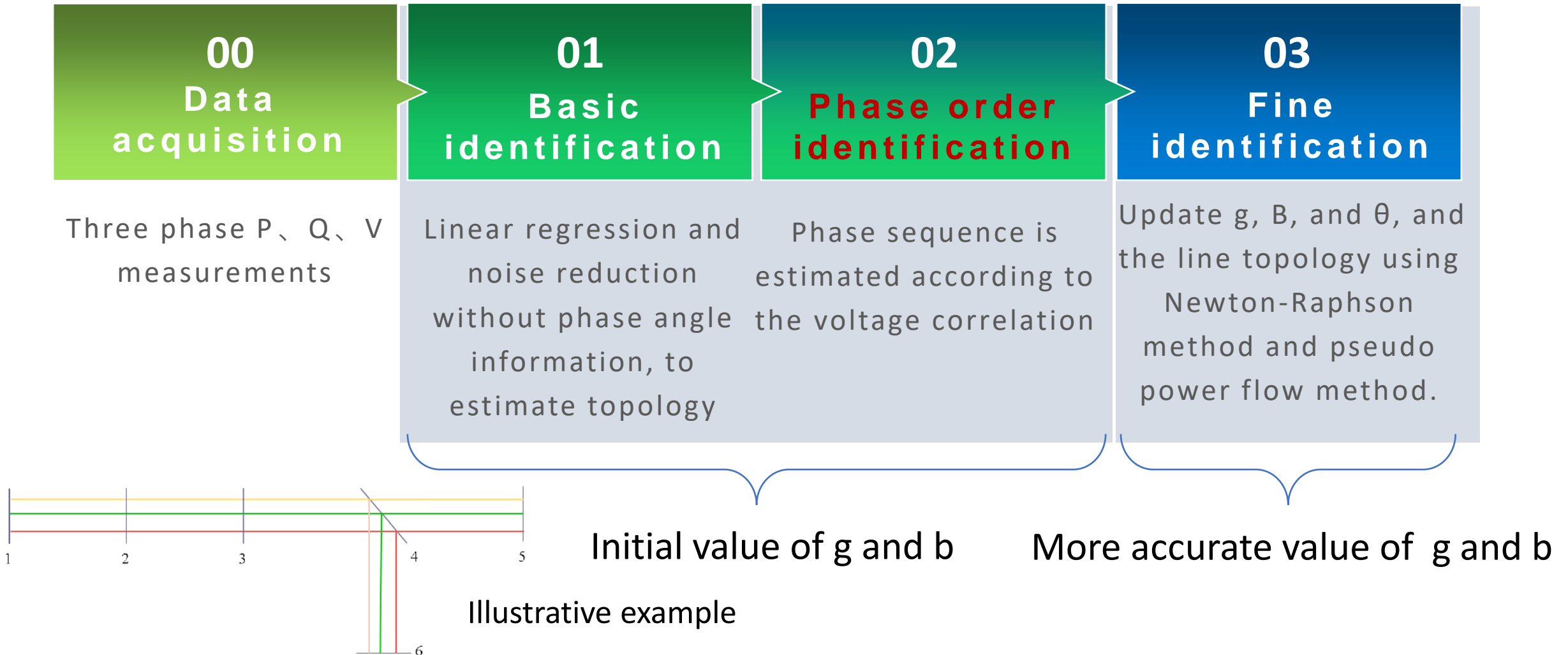


Self impedance and mutual impedance of three phase transmission line

Three phase conductance matrix

Unbalanced network: Framework

- Similar to the balanced network identification framework, but need to **identify phase order** after basic identification



Unbalanced network: Step 1 Basic Identification

- Identify three phase $\mathbf{G}^\#$ and $\mathbf{B}^\#$ matrices based on **linear regression**.
 - Since there is no phase angle information in hand, the phase angle difference between two nodes is assumed to be either 0 or $\pm 120^\circ$

$$\begin{cases} G_{ij}^{\alpha\beta\#} = G_{ij}^{\alpha\beta} \cos \theta_{ij}^{\alpha\beta} + B_{ij}^{\alpha\beta} \sin \theta_{ij}^{\alpha\beta} \approx -\frac{1}{2} G_{ij}^{\alpha\beta} \pm \frac{\sqrt{3}}{2} B_{ij}^{\alpha\beta} \\ B_{ij}^{\alpha\beta\#} = B_{ij}^{\alpha\beta} \cos \theta_{ij}^{\alpha\beta} - G_{ij}^{\alpha\beta} \sin \theta_{ij}^{\alpha\beta} \approx -\frac{1}{2} B_{ij}^{\alpha\beta} \mp \frac{\sqrt{3}}{2} G_{ij}^{\alpha\beta} \end{cases} \quad (\alpha \neq \beta) \quad \begin{cases} G_{ij}^{\alpha\beta\#} = G_{ij}^{\alpha\beta} \cos \theta_{ij}^{\alpha\beta} + B_{ij}^{\alpha\beta} \sin \theta_{ij}^{\alpha\beta} \approx G_{ij}^{\alpha\beta} \\ B_{ij}^{\alpha\beta\#} = B_{ij}^{\alpha\beta} \cos \theta_{ij}^{\alpha\beta} - G_{ij}^{\alpha\beta} \sin \theta_{ij}^{\alpha\beta} \approx B_{ij}^{\alpha\beta} \end{cases} \quad (\alpha = \beta)$$

- From three phase power flow equation to linear regression equation:

$$\begin{cases} P_i^\alpha = V_i^\alpha \sum_{j=1}^N \sum_{\beta=a,b,c} V_j^\beta (G_{ij}^{\alpha\beta} \cos \theta_{ij}^{\alpha\beta} + B_{ij}^{\alpha\beta} \sin \theta_{ij}^{\alpha\beta}) \\ Q_i^\alpha = V_i^\alpha \sum_{j=1}^N \sum_{\beta=a,b,c} V_j^\beta (G_{ij}^{\alpha\beta} \sin \theta_{ij}^{\alpha\beta} - B_{ij}^{\alpha\beta} \cos \theta_{ij}^{\alpha\beta}) \end{cases} \quad \Rightarrow \quad \begin{cases} \mathbf{P} / \mathbf{V} = \mathbf{V} \bullet \mathbf{G}^\# \\ \mathbf{Q} / \mathbf{V} = -\mathbf{V} \bullet \mathbf{B}^\# \end{cases}$$

Using M measurement

$$\begin{cases} [\mathbf{P} / \mathbf{V}]_M = [\mathbf{V}]_M \bullet \mathbf{G}^\# \\ [\mathbf{Q} / \mathbf{V}]_M = -[\mathbf{V}]_M \bullet \mathbf{B}^\# \end{cases}$$

$$\mathbf{G}^\# = ([\mathbf{V}]_M^T [\mathbf{V}]_M)^{-1} [\mathbf{V}]_M^T [\mathbf{P} / \mathbf{V}]_M$$

$$\mathbf{B}^\# = -([[\mathbf{V}]_M^T [\mathbf{V}]_M)^{-1} [\mathbf{V}]_M^T [\mathbf{Q} / \mathbf{V}]_M$$

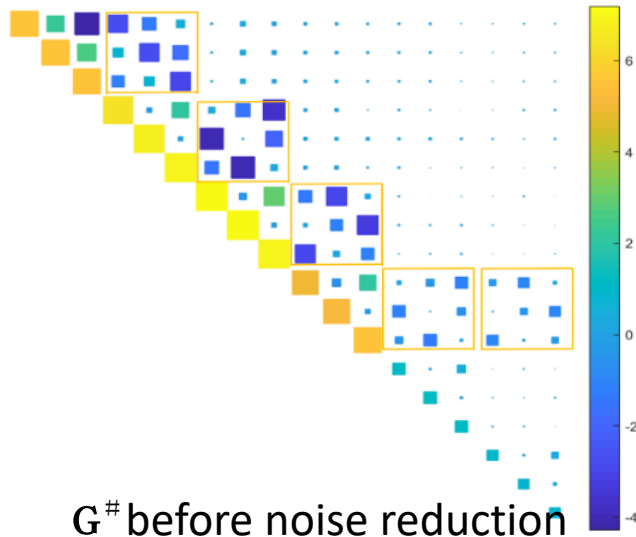
Pseudo conductance and admittance matrix

Unbalanced network: Step 1 Basic Identification

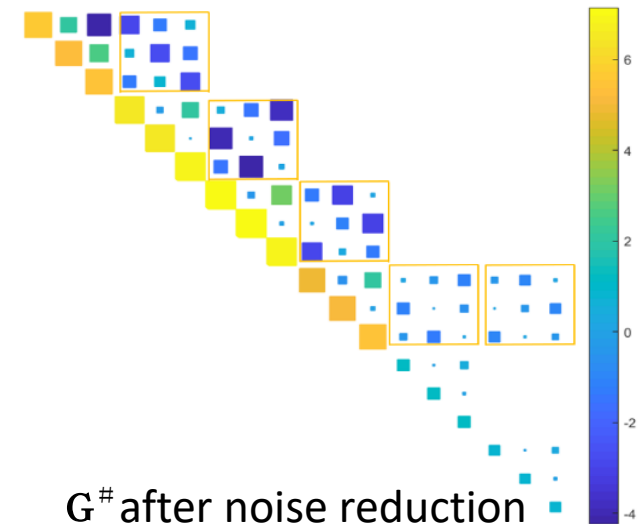
- Identify three phase $\mathbf{G}^\#$ and $\mathbf{B}^\#$ matrices based on **linear regression**.
- Noise reduction and linear regression iteration to obtain **unordered** $\mathbf{G}^\#$ and $\mathbf{B}^\#$.

$$\lambda(i, j)_{i \neq j} = \frac{G_{ij}^{aa\#} + G_{ij}^{bb\#} + G_{ij}^{cc\#}}{G_{ii}^{aa\#} + G_{ii}^{bb\#} + G_{ii}^{cc\#}} < \eta$$

$$[\mathbf{P}_i^\alpha / \mathbf{V}_i^\alpha]_M = [\mathbf{V}_i^\alpha]_M \bullet [\mathbf{G}_i^{\alpha\#}]$$



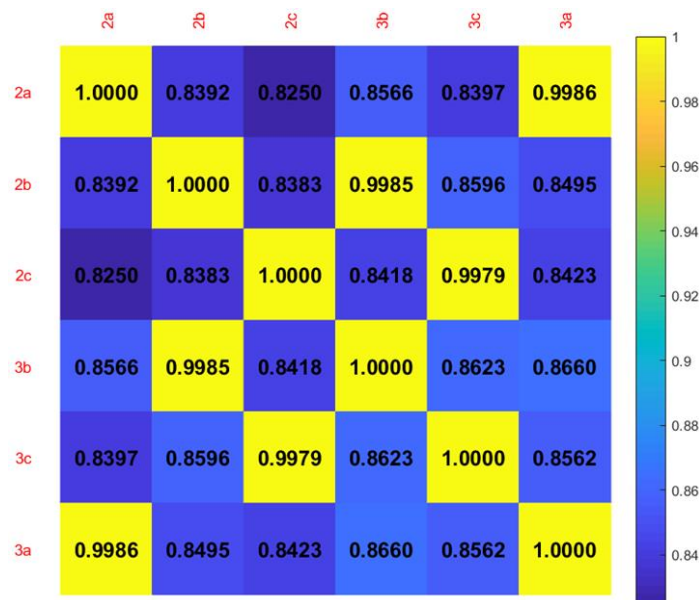
Noise reduction



Unbalanced network: Step 2 Phase order identification

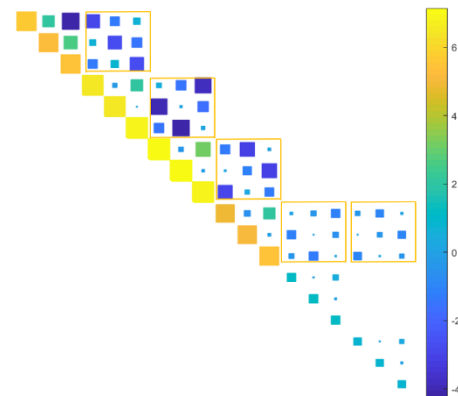
In the three-phase unbalanced distribution network, the voltage correlation between node of the same phase is significantly larger than that of different phase, so we can use the correlation information to identify the phase order.

The phase order of $\mathbf{G}^\#$ and $\mathbf{B}^\#$ can be adjusted by left and right multiply basic matrices.



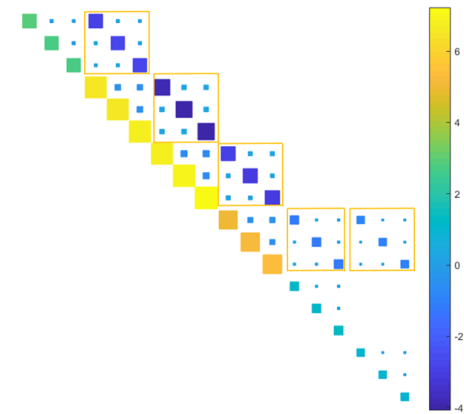
Three phase voltage correlation matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}
 \begin{bmatrix} G_{22}^{aa\#} & G_{22}^{ab\#} & G_{22}^{ac\#} & G_{23}^{ab\#} & G_{23}^{ac\#} & G_{23}^{aa\#} \\ G_{22}^{ba\#} & G_{22}^{bb\#} & G_{22}^{bc\#} & G_{23}^{ba\#} & G_{23}^{bc\#} & G_{23}^{ba\#} \\ G_{22}^{ca\#} & G_{22}^{cb\#} & G_{22}^{cc\#} & G_{23}^{ca\#} & G_{23}^{cb\#} & G_{23}^{ca\#} \\ G_{23}^{ab\#} & G_{23}^{bb\#} & G_{23}^{cb\#} & G_{33}^{bb\#} & G_{33}^{bc\#} & G_{33}^{ba\#} \\ G_{23}^{ac\#} & G_{23}^{bc\#} & G_{23}^{cc\#} & G_{33}^{cb\#} & G_{33}^{cc\#} & G_{33}^{ca\#} \\ G_{23}^{aa\#} & G_{23}^{ba\#} & G_{23}^{ca\#} & G_{33}^{ab\#} & G_{33}^{ac\#} & G_{33}^{aa\#} \end{bmatrix}
 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$



$\mathbf{G}^\#$ before phase order identification

Phase order adjustment



$\mathbf{G}^\#$ after phase order identification

Unbalanced network: Step 2 Phase order identification

If the distribution network has a tree structure, the topology can be further adjusted based on voltage correlation.

From the root node:

Rank the voltage correlation coefficient from the highest to lowest



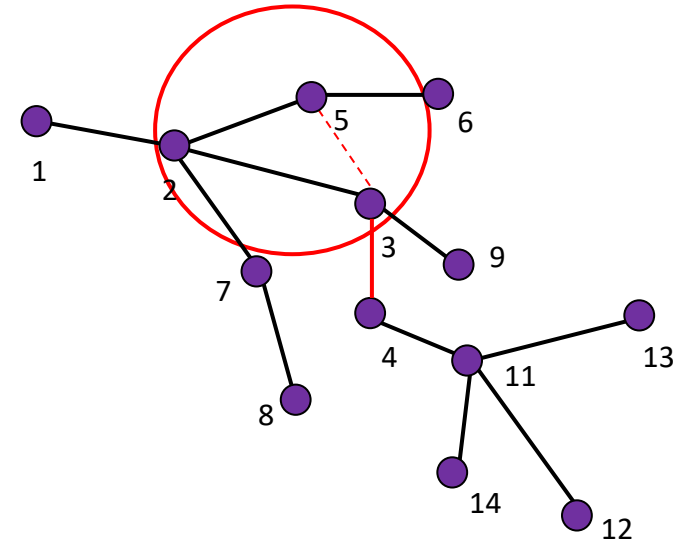
Choose the other node with highest correlation



Whether the branch exists and whether it will form a loop



Confirm the branch and move to the next node.



Correlation coefficient order: $R_{35} > R_{34} > R_{49} > \dots$

For each node, the voltage correlation of the node directly connected to it is usually greater than that of the node not directly connected to it.

Contents

Background

Balanced network

Three phase unbalanced network (Self-study)

Test Cases

Conclusions

Case Study

Balanced network

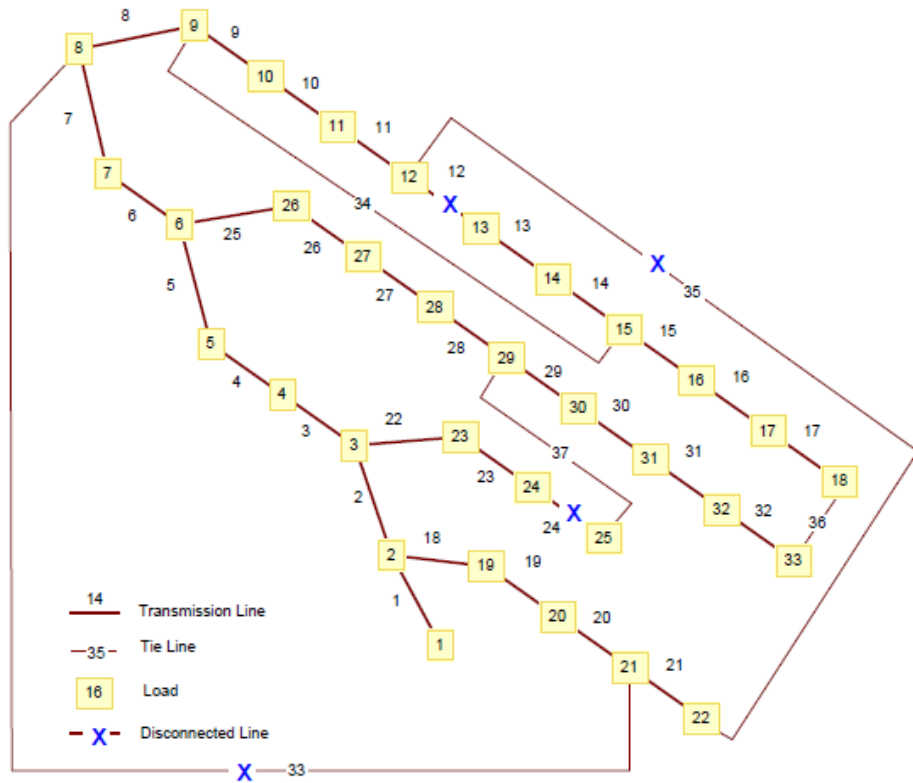
- IEEE 33-bus Test Case (High Renewables Penetration)
- IEEE 123-bus Test Case (Large Distribution Network)

Three phase unbalanced network

- IEEE 34-bus Test Case
- IEEE 123-bus Test Case (Large Distribution Network)

Balanced network: IEEE 33-bus Test Case (High Renewables Penetration)

A. IEEE 33-feeder case



Renewables on bus 2、 5、 7、 11、
15、 16、 22、 26、 30.

35% Renewable Energy Penetration.

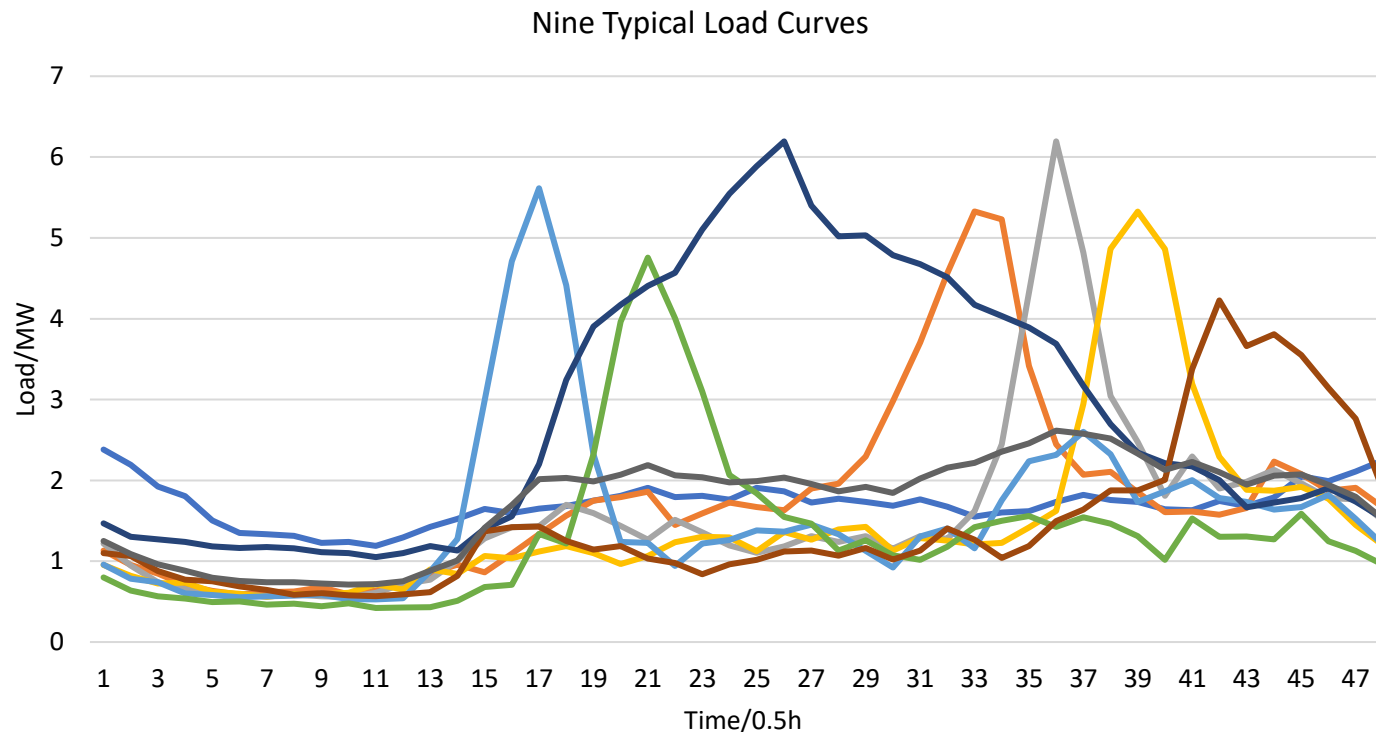
Sampling: 24h 12min/record

0.5% additional error to active and reactive power

Balanced network: IEEE 33-bus Test Case (High Renewables Penetration)

Load:

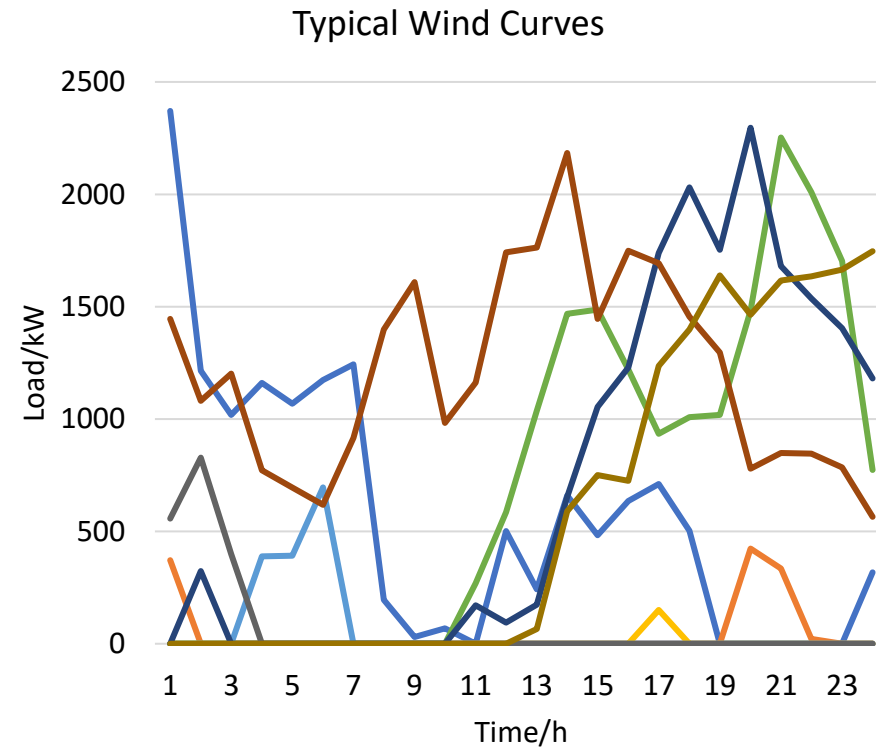
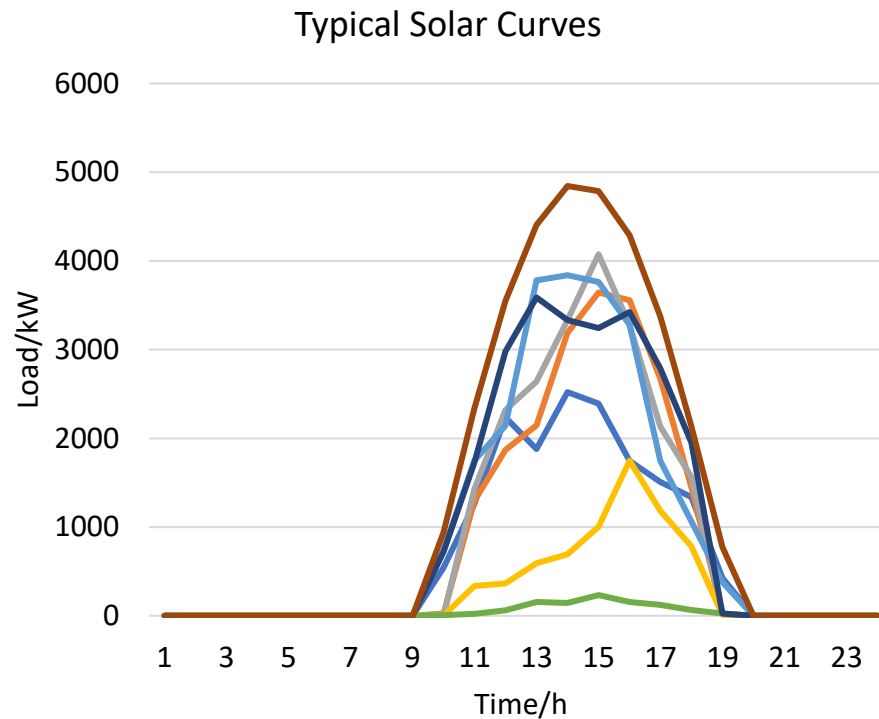
Smart Meter Electricity Trial data from The Research Perspective Ltd. It includes power load curves from 1000 residents and small companies.



Balanced network: IEEE 33-bus Test Case (High Renewables Penetration)

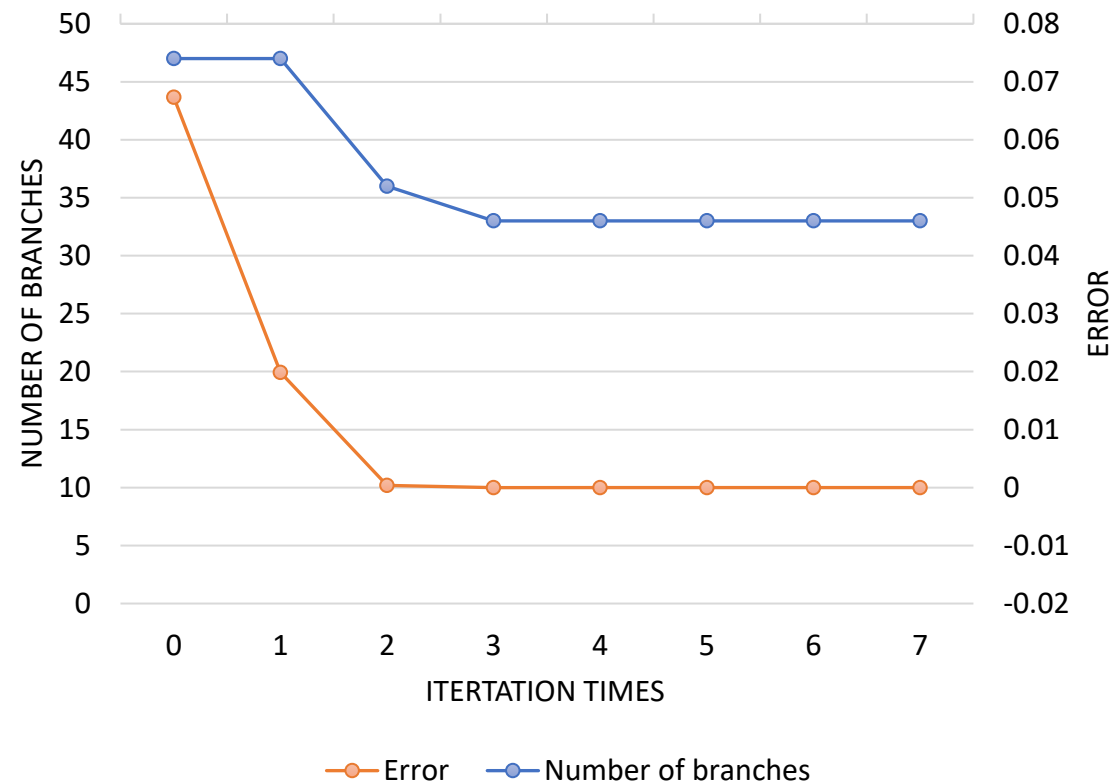
Renewables:

- 51 PV stations and 34 wind farms in China



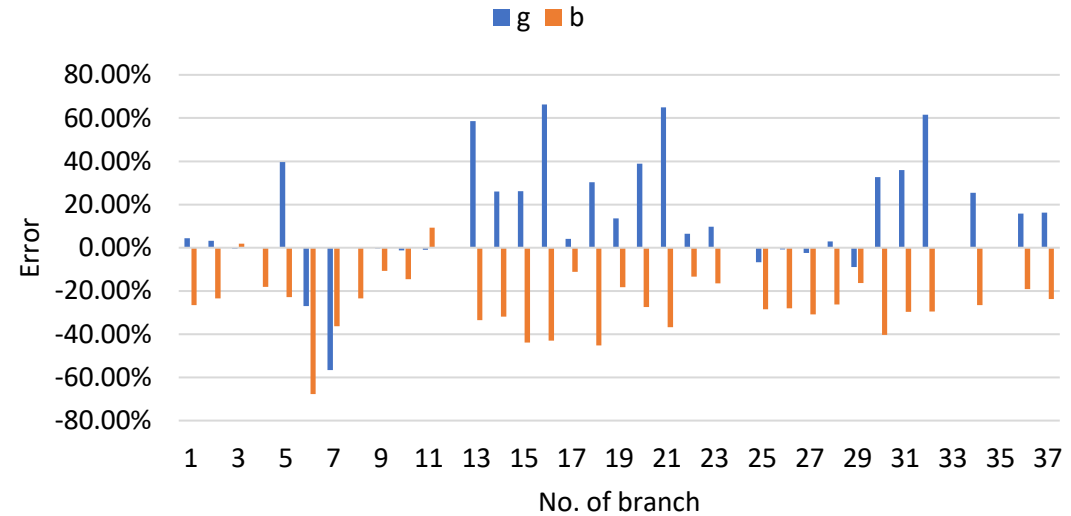
Balanced network: IEEE 33-bus Test Case (High Renewables Penetration)

There are 14 wrong branches after basic identification;
Topology is corrected in 3 iterations;
Convergence after 7 iterations.

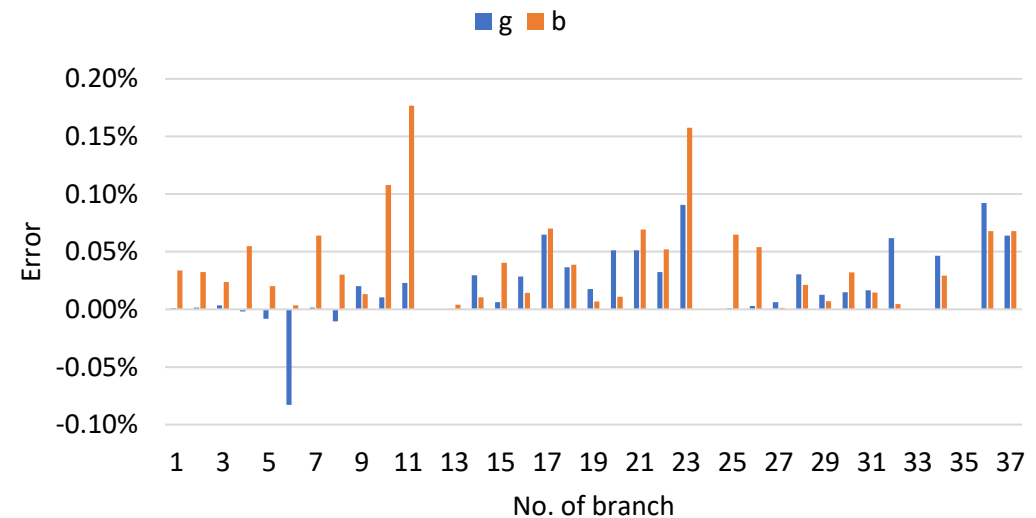


Balanced network: IEEE 33-bus Test Case (High Renewables Penetration)

g, b estimation errors are 20.9% and 26.5% after basic identification.

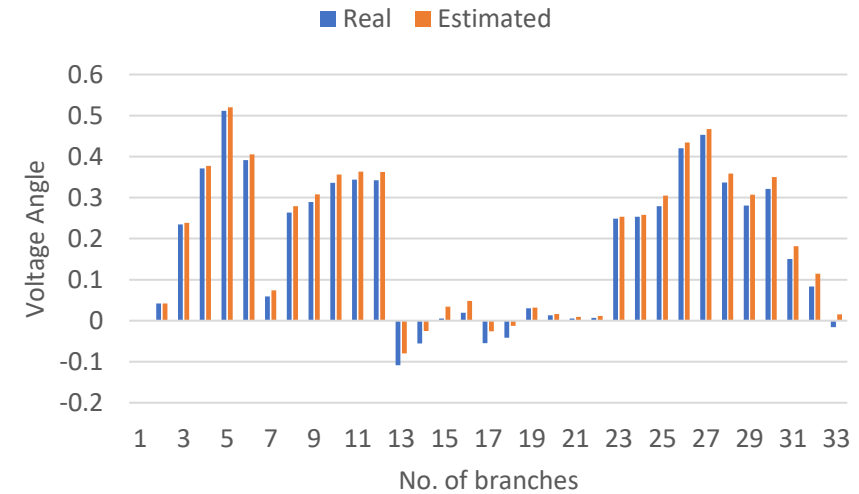
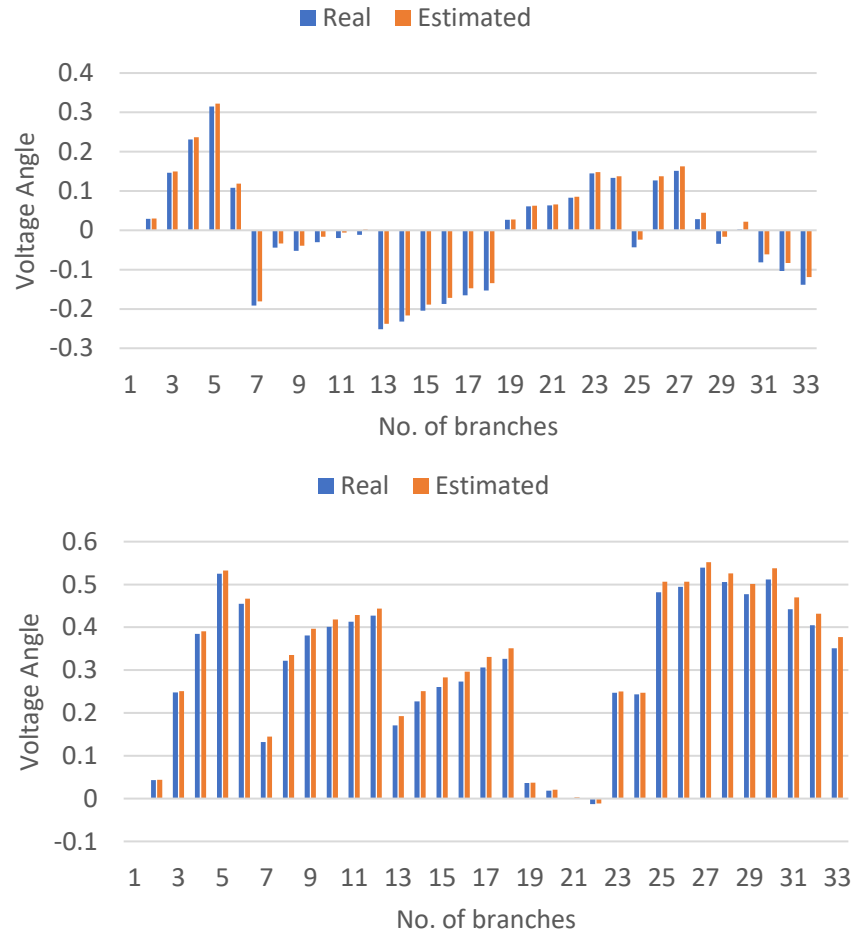


While 0.03%,0.04% after fine identification.



Balanced network: IEEE 33-bus Test Case (High Renewables Penetration)

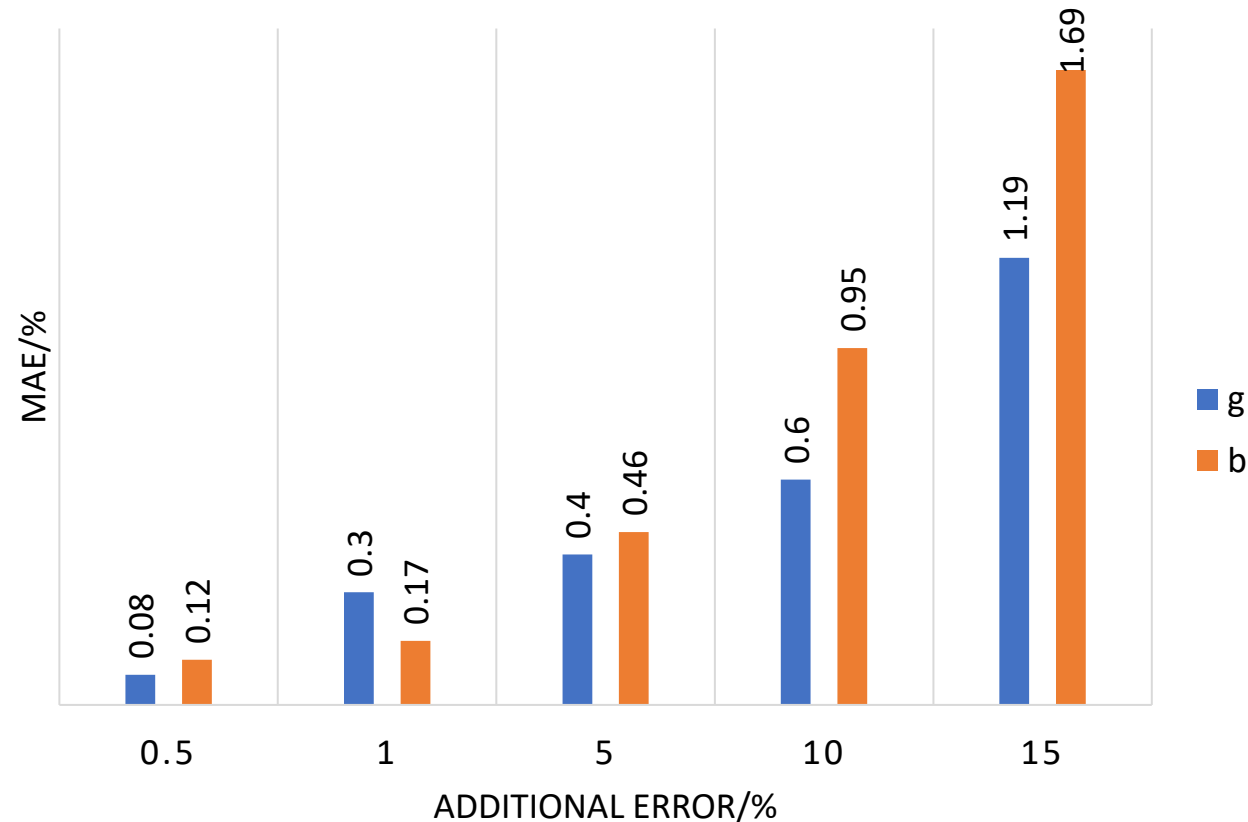
Voltage angle:



Voltage angle after fine identification.
Deviation < 0.02°

Balanced network: IEEE 33-bus Test Case (High Renewables Penetration)

With the increasing of additional error, the estimation error of g and b line parameters also increases. The method proposed in this paper has a high degree of redundancy when the measurement error is significant.



Balanced network: IEEE 123-bus Test Case (Large Distribution Network)

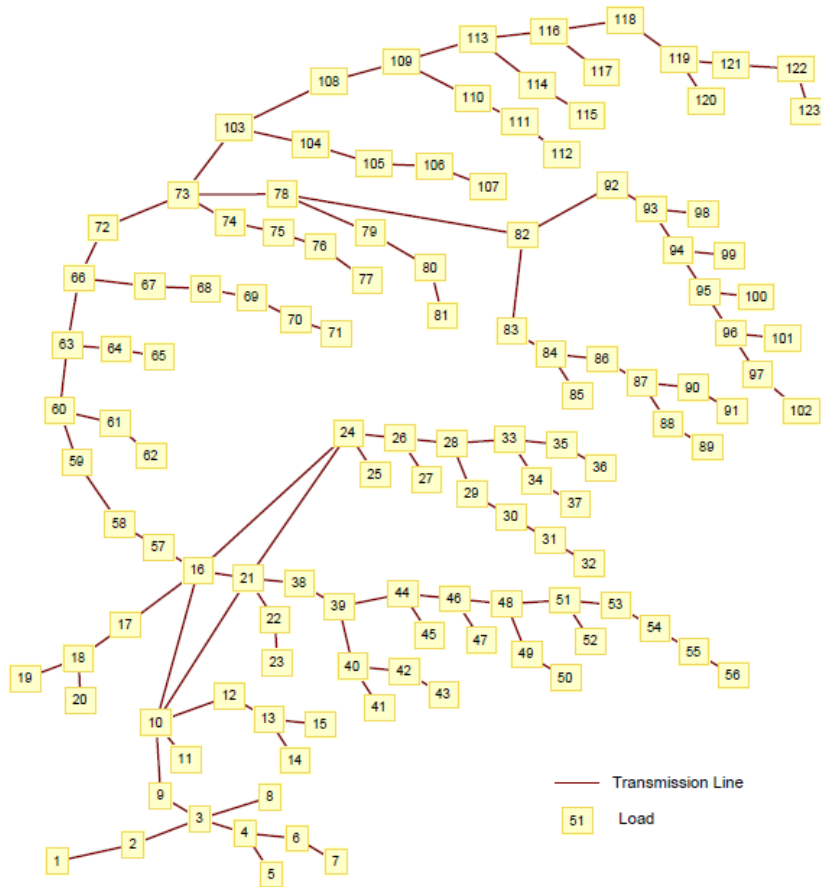


Fig. 7. IEEE 123 Bus Feeder

In this test case, we investigate the impact of the size of measurement samples in basic identification on the whole algorithm.

i.e. 24-hour dataset with 10, 15, 20 and 25 samples per hour

In fine identification, we only select the last 20 samples in basic identification's dataset .

Basic identification

Fine
identification

Data

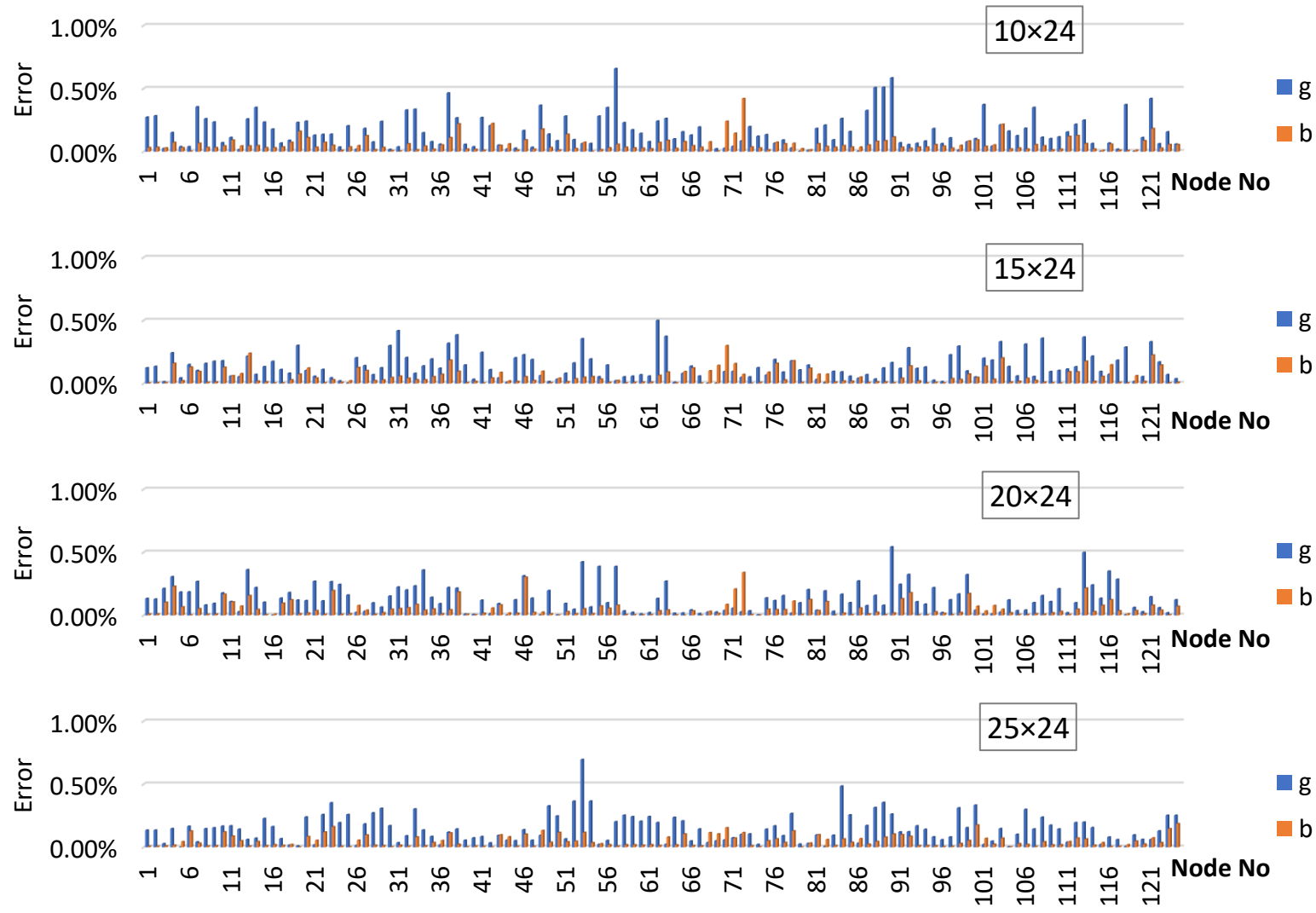
Balanced network: IEEE 123-bus Test Case (Large Distribution Network)

- Though larger data size would increase time consumed in basic identification, with larger data size, the wrong branches after basic identification reduce, and it may save time for the proposed method in fine identification.

| Basic identification Data Size | MAE of g | MAE of b | Number of Iterations | Unidentified Branches after Basic Identification | Time Consumption/s |
|--------------------------------|------------|------------|----------------------|--|--------------------|
| 10×24 | 0.158% | 0.057% | 6 | 32 | 141.3 |
| 15×24 | 0.133% | 0.055% | 9 | 24 | 150.4 |
| 20×24 | 0.133% | 0.051% | 7 | 5 | 108.6 |
| 25×24 | 0.144% | 0.047% | 6 | 1 | 107.3 |

Balanced network: IEEE 123-bus Test Case (Large Distribution Network)

g, b absolute estimation error under different basic identification data size.

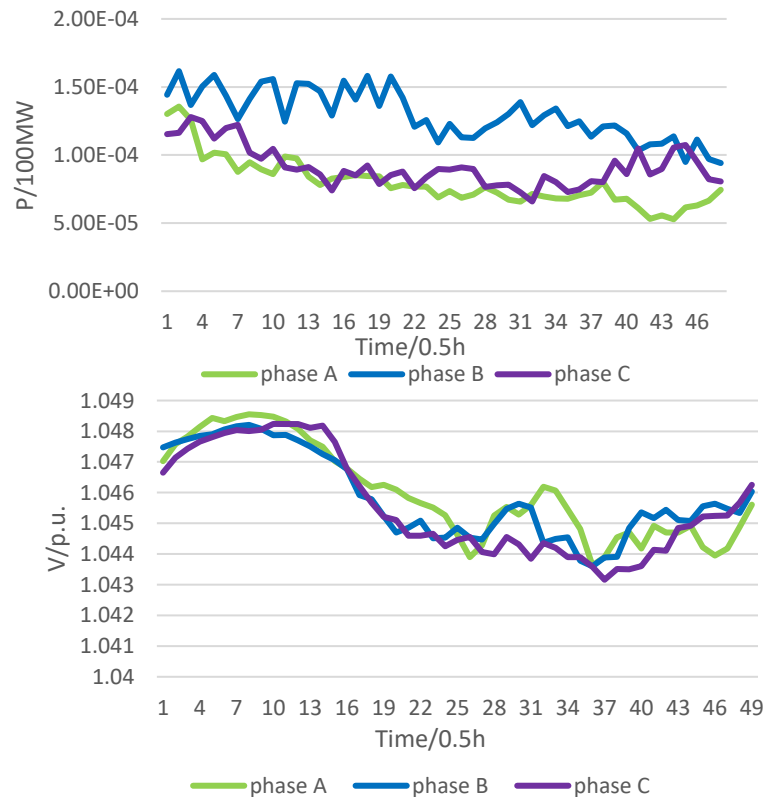


Unbalanced network: IEEE 34-bus Test Case

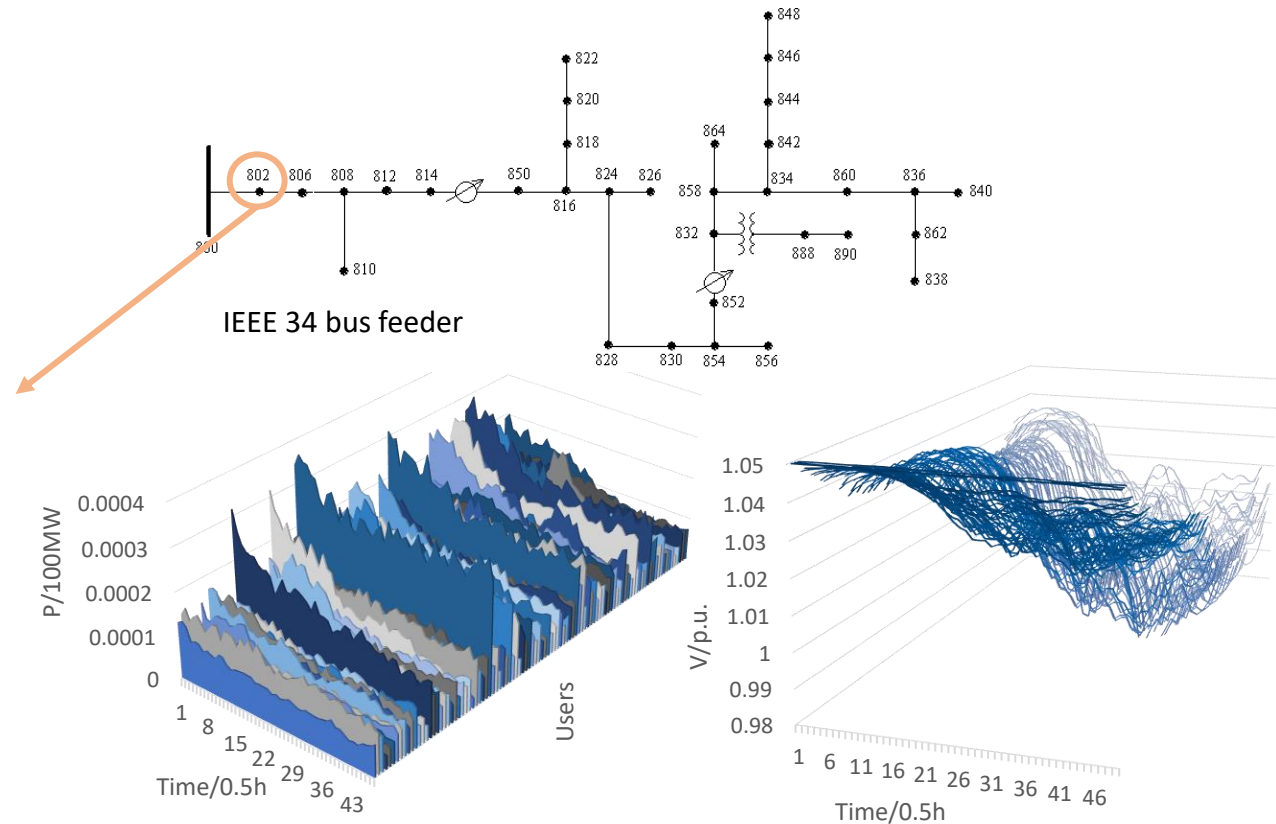
Generate load curve and required active and reactive power data, and distribute them to each bus and each phase, making the network unbalanced.

Dataset with 300 samples is used in regression; dataset with 10 samples is used in fine identification.

Obtain V dataset according to the three-phase power flow equation.



three-phase active power and voltage curves of bus 802 in 24 hours



three-phase active power curve of all buses in 24 hours

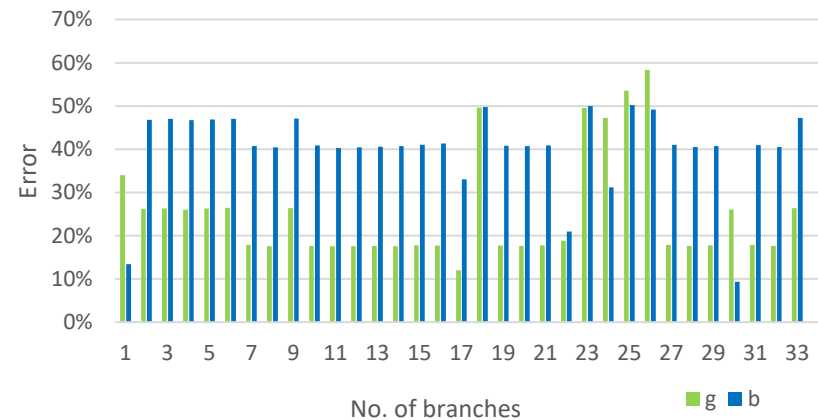
three-phase voltage magnitude of all buses in 24 hours

Unbalanced network: IEEE 34-bus Test Case

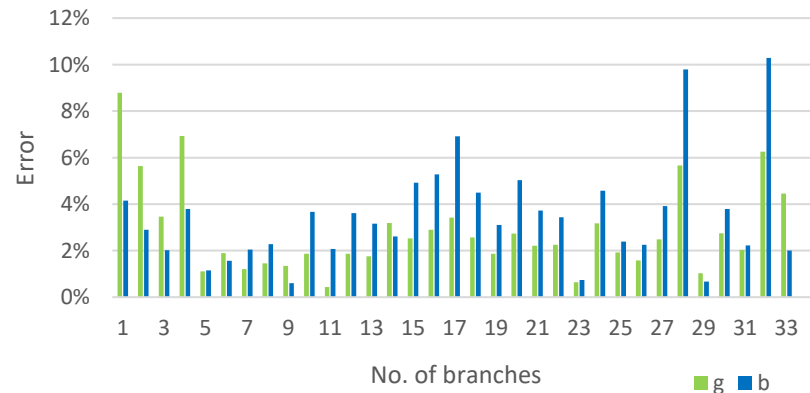
Adding 0.2% additional error to active and reactive power, after the regression and phase order identification, the topology and phase are correct; g, b estimation errors are **25.24%** and **40.27%** respectively.

After **13** iterations, g, b estimation errors are **2.82%** and **3.44%** respectively.

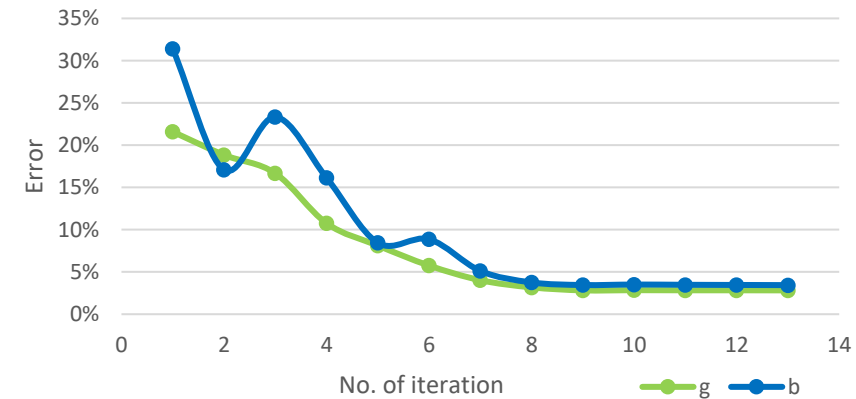
The time consumption of identification is **7.68s**.



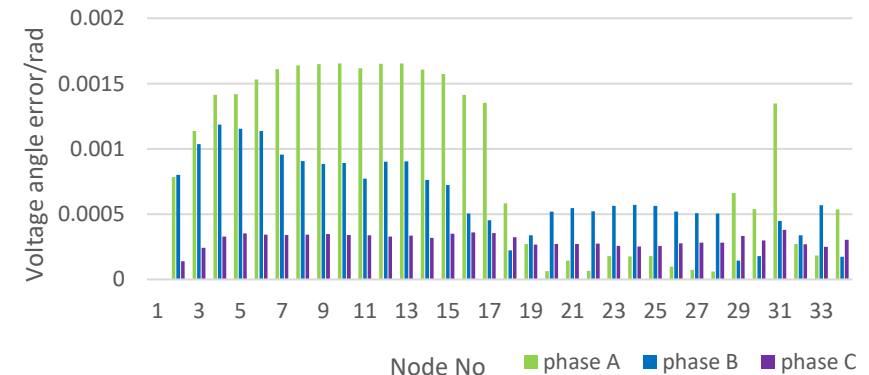
Error before fine identification



Error after fine identification



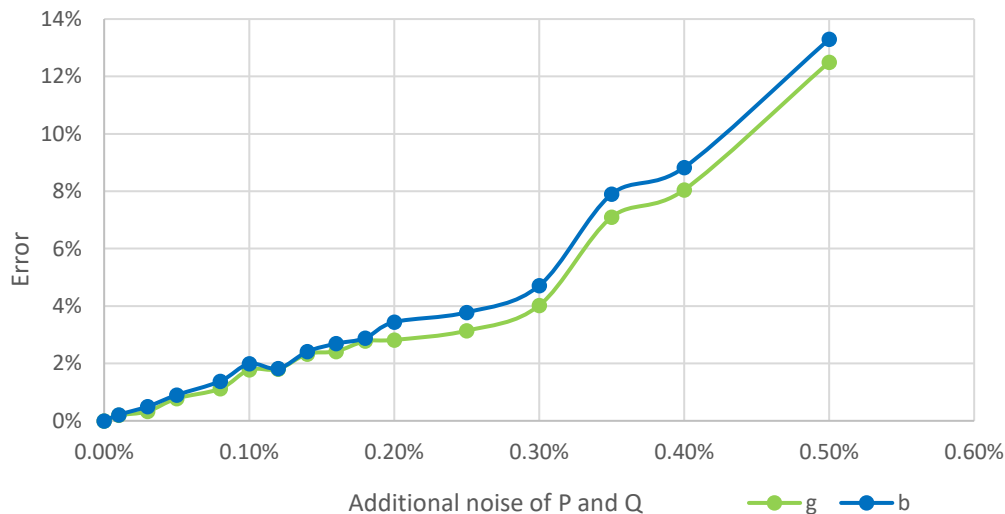
Error in each iteration



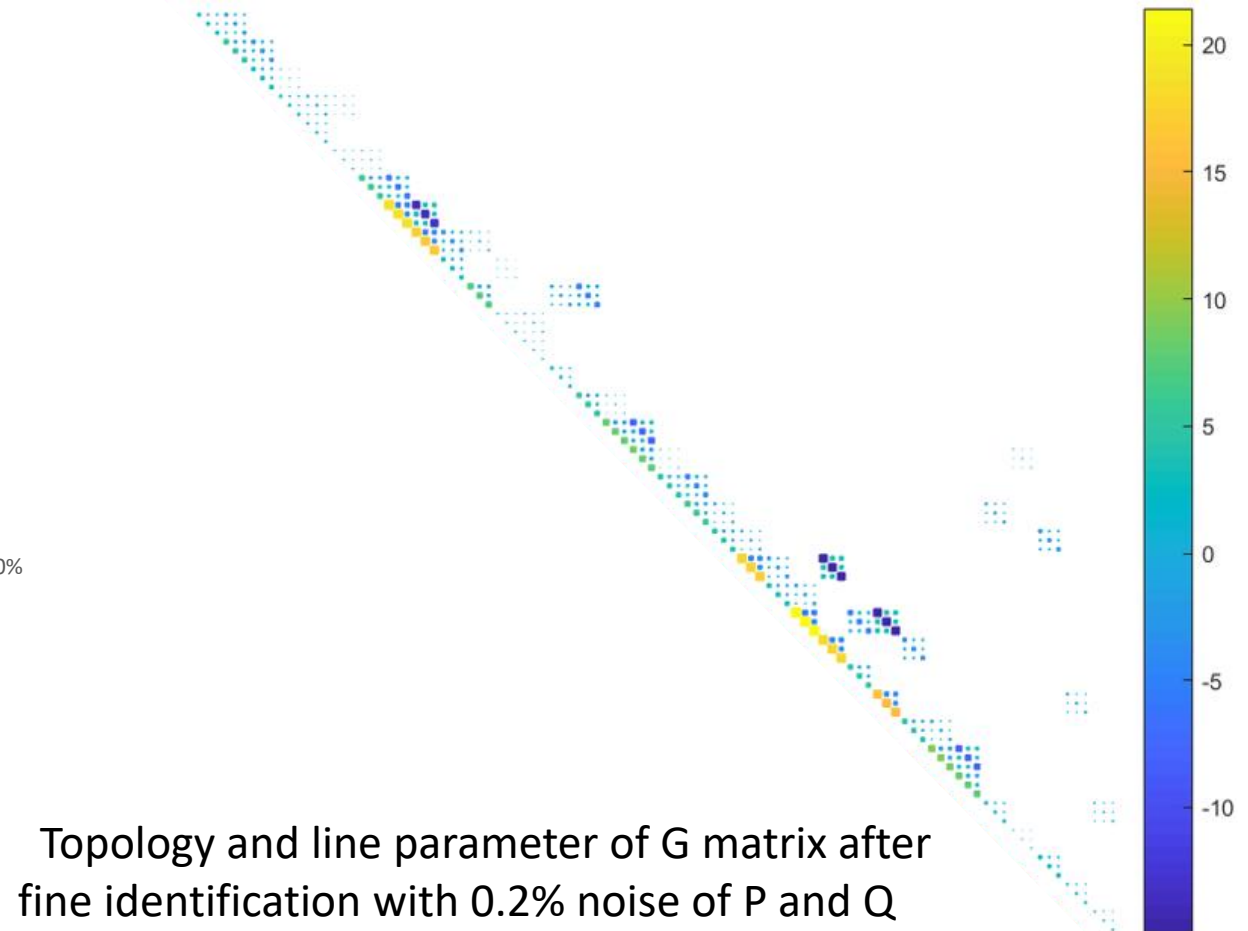
Three-phase voltage angle error after fine identification

Unbalanced network: IEEE 34-bus Test Case

With the incensement of the additional noise of P and Q, the error of g and b also increases, when the noise reaches to **0.5%**, the estimation error is over **10%**, but the topology is still correct.

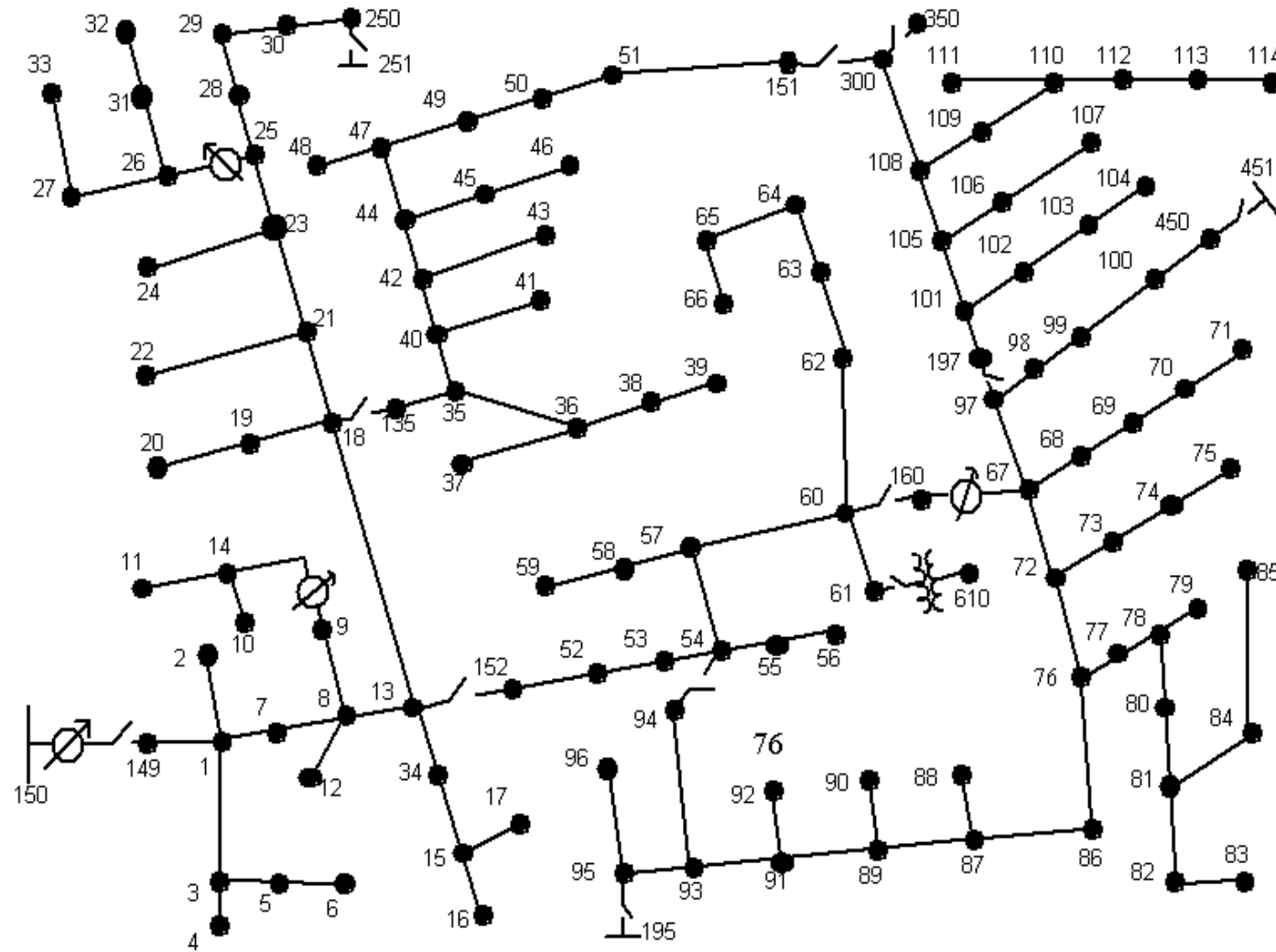


Identification error of different additional noise



Topology and line parameter of G matrix after fine identification with 0.2% noise of P and Q

Unbalanced network: IEEE 123-bus Test Case

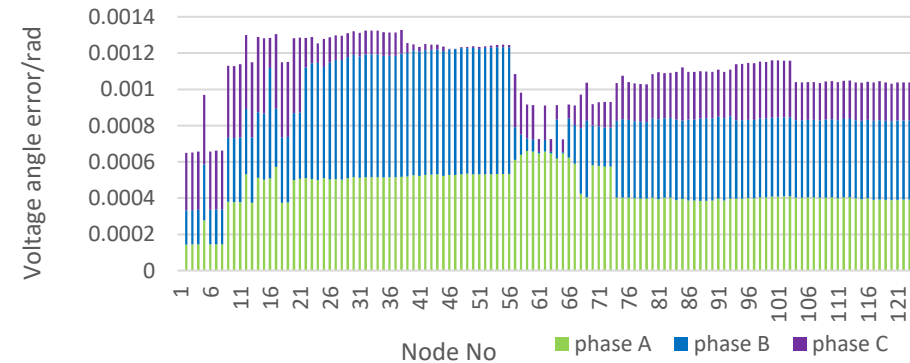
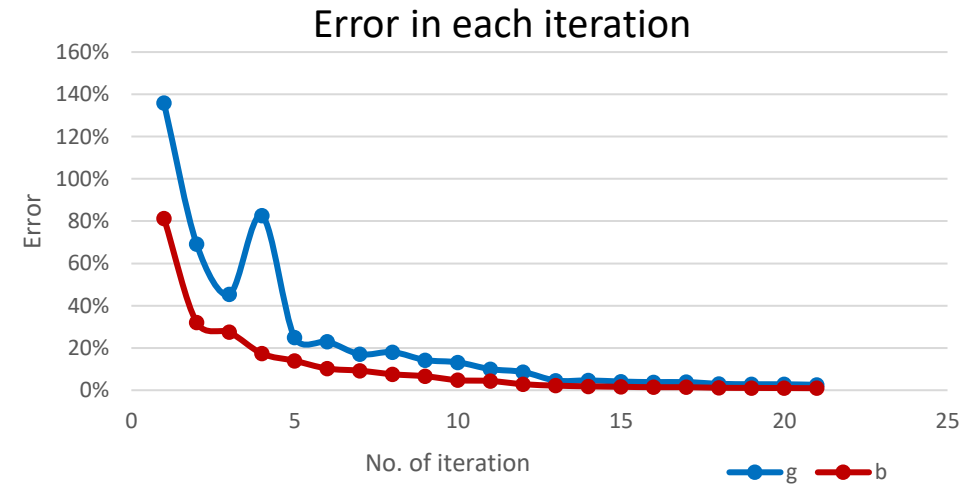
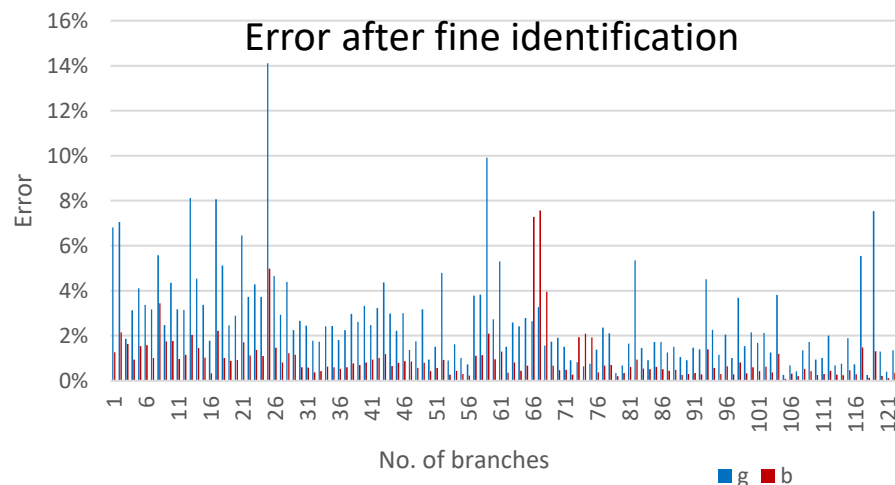
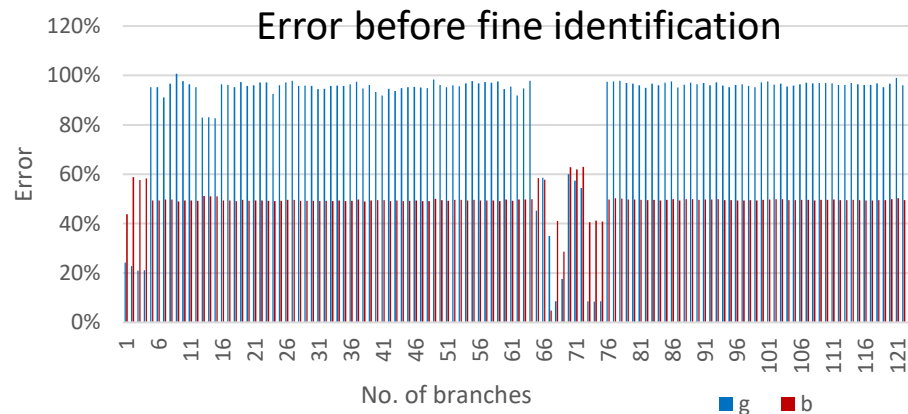


IEEE 123 bus test feeder

Unbalanced network: IEEE 123-bus Test Case

Adding 0.1% additional error to active and reactive power, after the regression and phase order identification, the topology and phase are correct; g, b estimation errors are **89.3%** and **49.1%** respectively.

After **21** iterations, g, b estimation errors are **2.56%** and **0.96%** respectively. The time consumption of identification is **116s**.



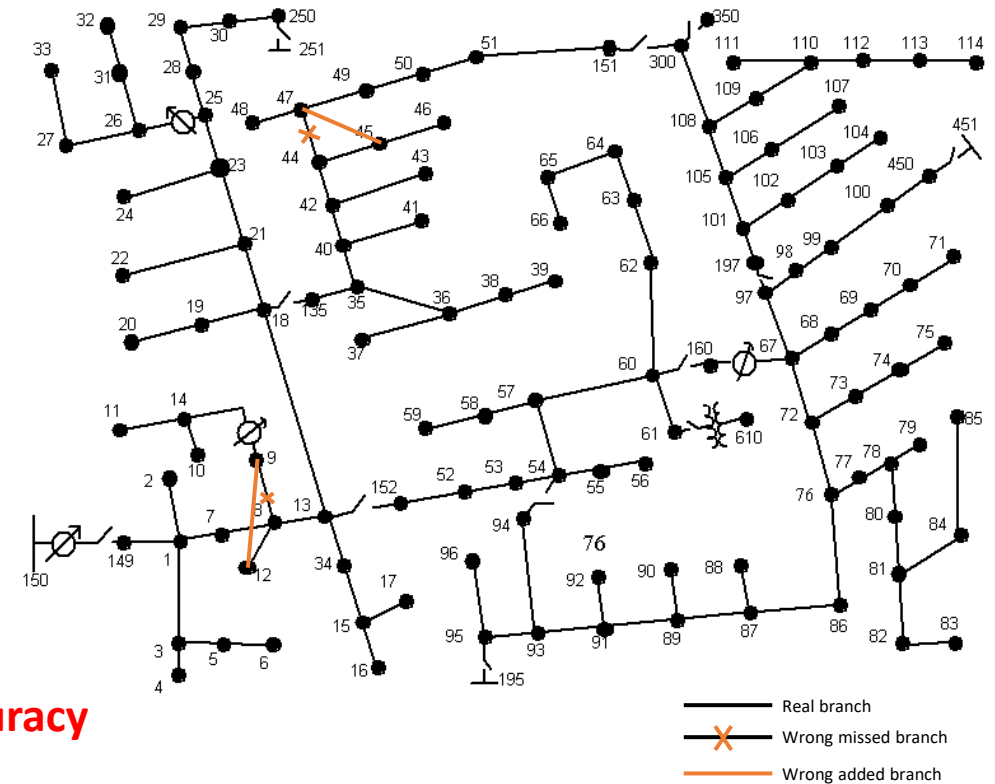
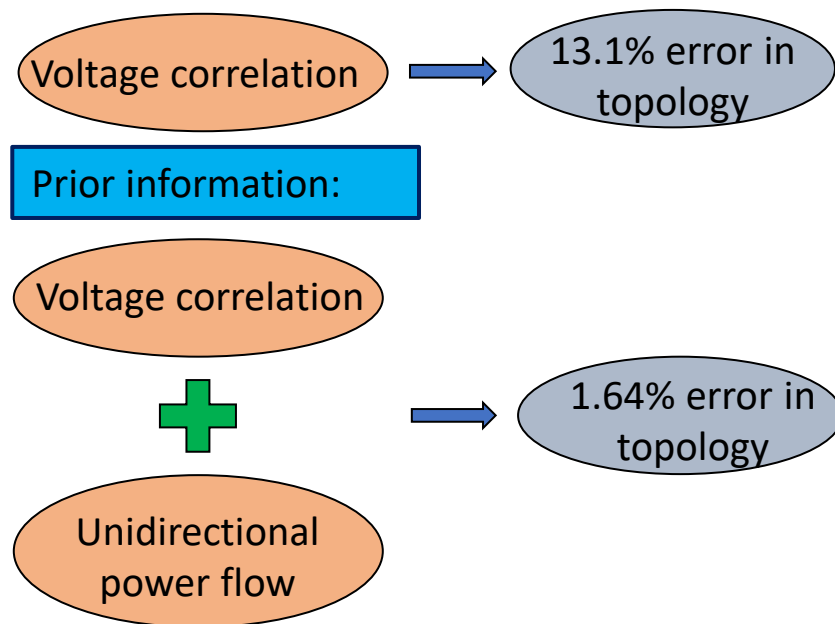
Three-phase voltage angle error after fine identification

Unbalanced network: IEEE 123-bus Test Case

□ Topology identification with voltage noise

The per-unit value of voltage of each bus in distribution network is about 1, and the fluctuation is small, the noise of voltage will possibly change the power flow and has a great impact on the identification of the topology and parameter.

Add **0.01%** additional noise to voltage of each bus, it's found that the regression and fine identification are difficult. So try to use the voltage correlation to directly identify the topology.



If there is some prior node connection information, the accuracy will be further improved.

References

Jiawei Zhang, Yi Wang, Yang Weng, and Ning Zhang*, Topology Identification and Line Parameter Estimation for non-PMU Distribution Network: A Numerical Method. IEEE Transactions on Smart Grid, doi: 10.1109/TSG.2020.2979368

Code available: <https://github.com/AmateurZhang/MatIdentification>

Jiawei Zhang, Peng Wang, Ning Zhang*, Distribution Network Admittance Matrix Estimation with Linear Regression, IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2021.3090250.

宁家鑫, 刘羽霄, 章家维, 李海, 朱桂萍, 钱敏慧, 张宁, 数据驱动的三相配电网络拓扑与线路参数辨识, 中国电机工程学报, 2021,41(08):2615-2628.

Q&A