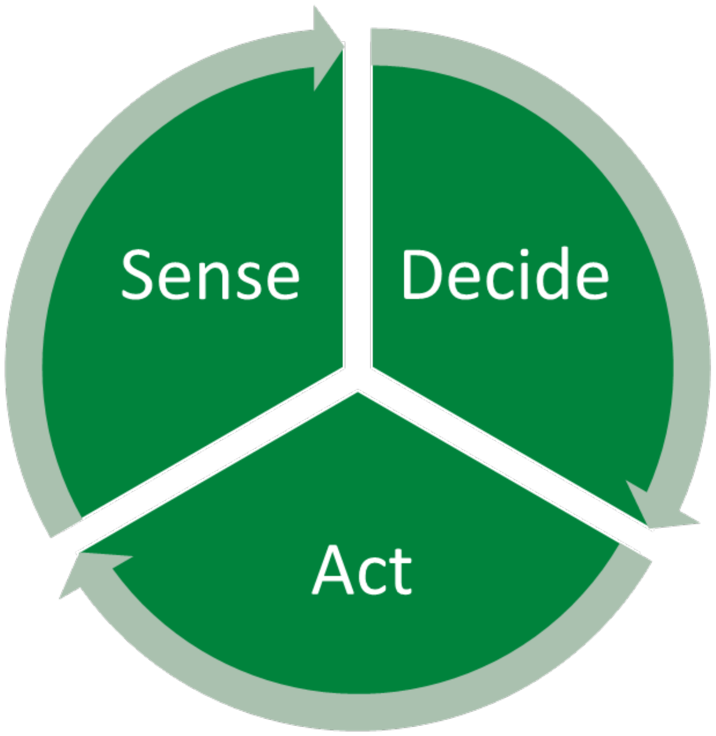
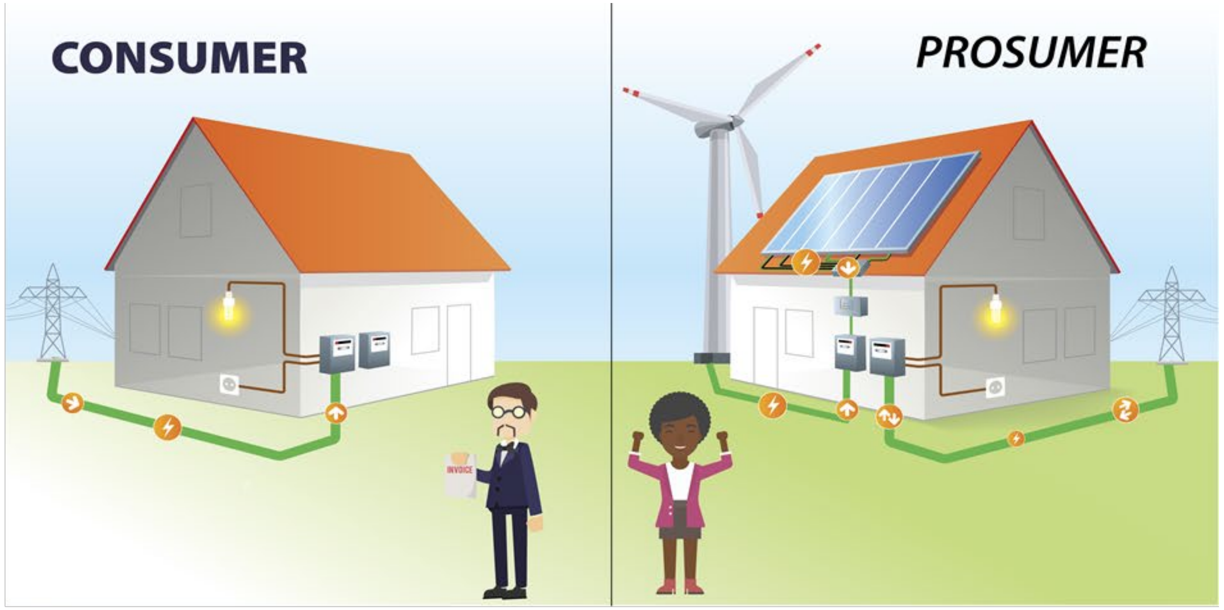
Situational Awareness in Distribution Networks - Doing More with Less!

Dr. Bala Natarajan
Clair N Palmer and Sara M Palmer Professor
Mike Wiegers Department of Electrical and Computer Engineering
Kansas State University
bala@ksu.edu

Foundations of Intelligence



Grid Edge Evolution

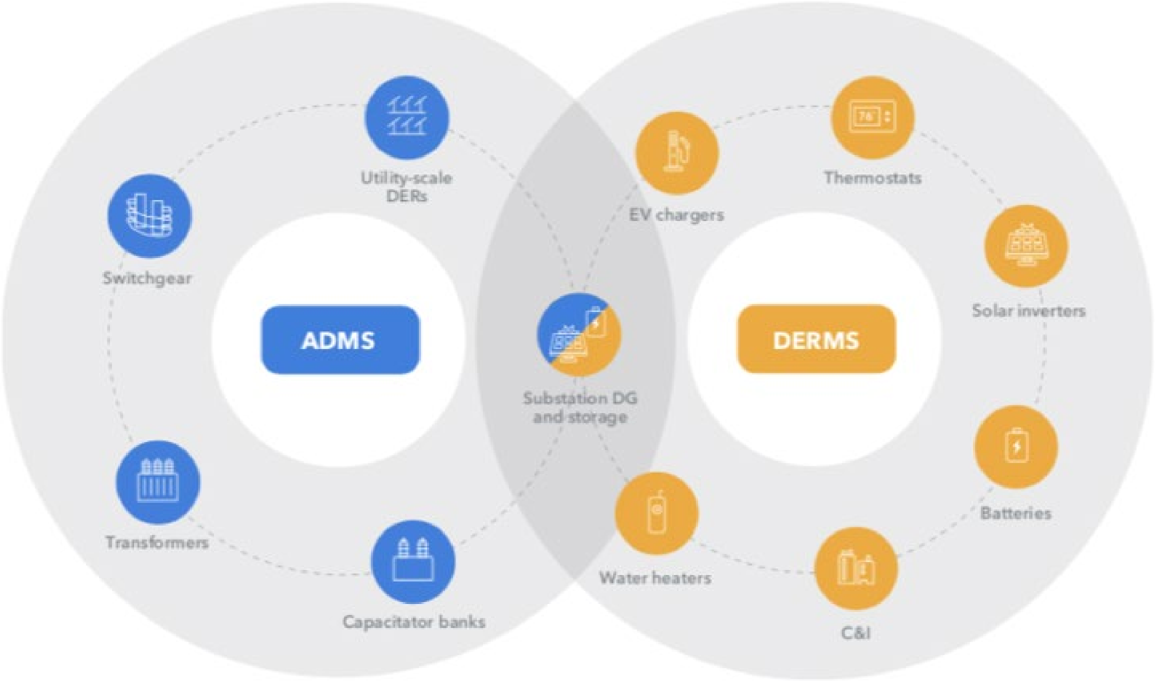


Ideas and Solutions

- Smart inverters
- Blockchain technology
- Network management
- PV curtailment
- Optimal DG/Storage placement
- Advanced protection devices
- Forecasting based management
- Capacitor bank controllers
- Smart Storage
- Distributed control
- Outage detection
- Cyber security solutions
- Charge scheduling
- Demand response

Situational Awareness is Key!

Most grid operators lack either visibility or control of behind-the-meter devices



DERMS transform the DER/EV/BTM assets into solutions

Exploiting Information for system situational awareness

Information is Power!

Utility standpoint

- Have less than what we would like
- Lack visibility or control of BTM devices
- Collected data is in different databases
- Data is collected at different time scales
- Data collected may not be consistent/reliable
- Data is not used for operational purposes due to the delays involved

Grand Challenge

We may never have the quantity or the quality of data we need for complete situational awareness at the grid edge. Can we do more with less?

Doing More with Less!

Distribution system state estimation (DSSE)

Integrating multi-timescale measurements

Integrating Topology and Phase Identification

Distribution grid is unobservable

Currently few measurements are available

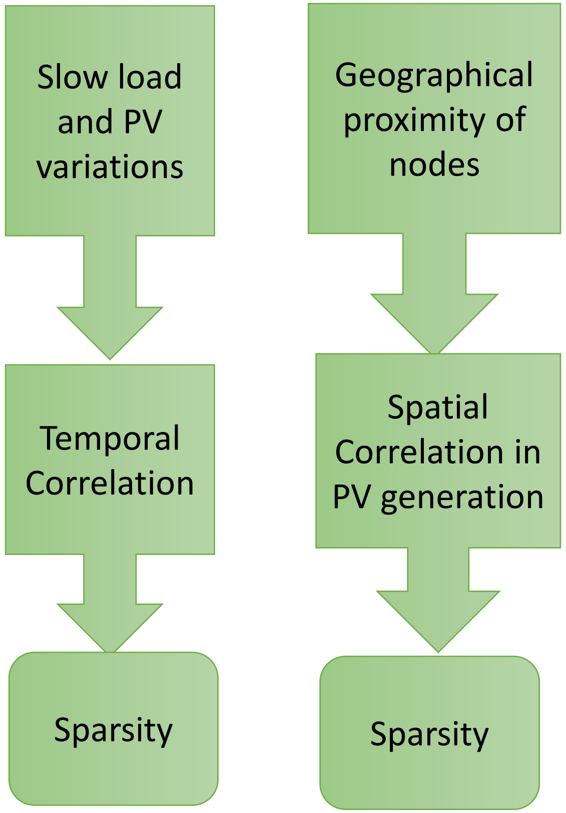
Theorem 1: If there exists a vertex-disjoint set of

Question 1: How to estimate the states using a small number of measurements?

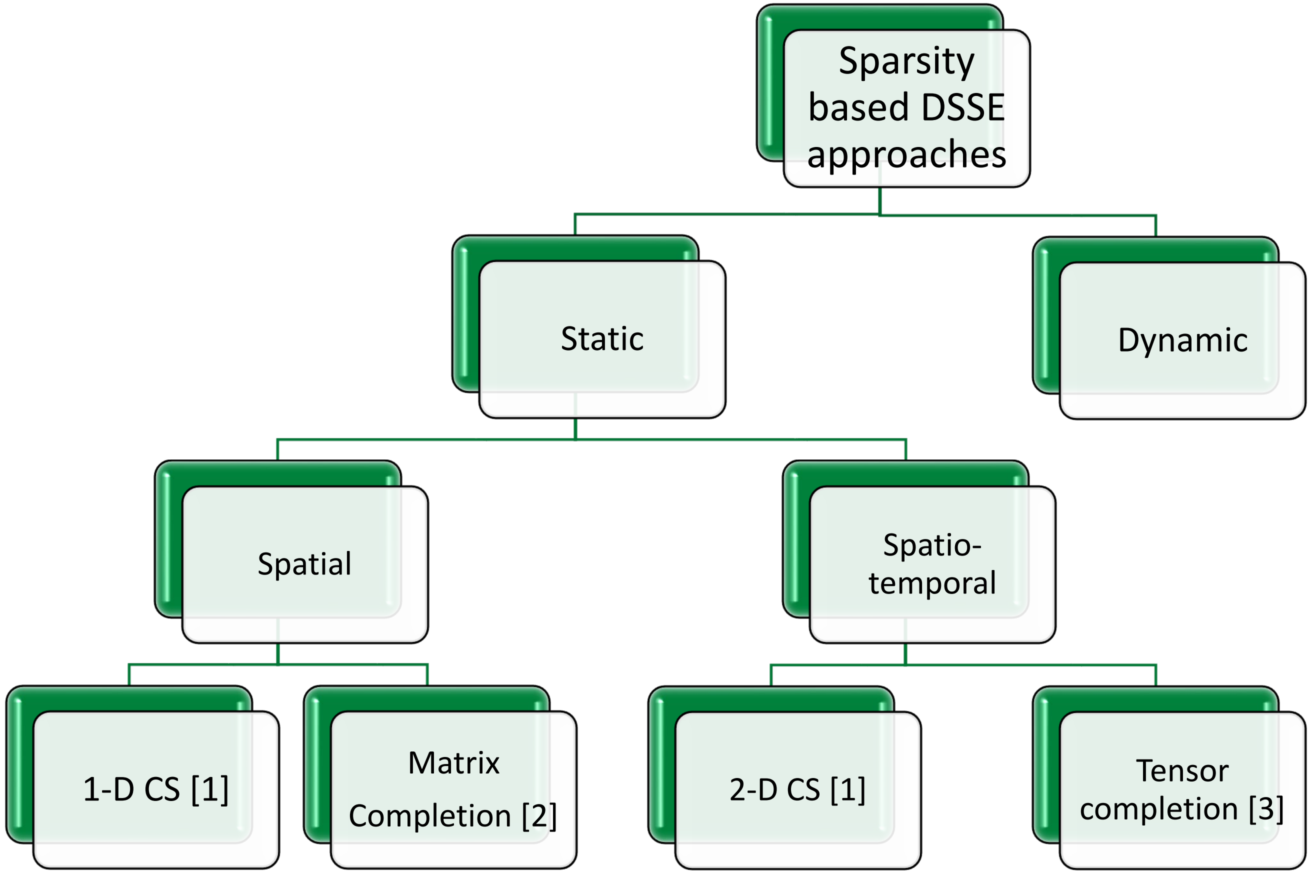
Question 2: How robust are the estimators to bad data, system uncertainties and cyber attacks

network/data processing

Sparsity-based state estimation



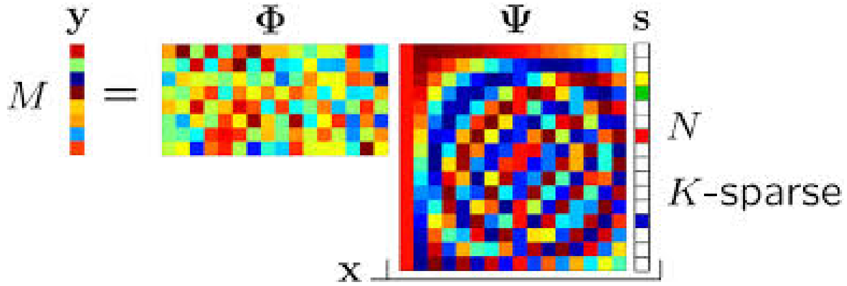
Sparsity in distribution systems



Sparsity-aware DSSE approaches

Compressed Sensing based DSSE

Estimate the signal of interest $\boldsymbol{x} \in \mathbb{R}^{N}$ from compressed measurements $\boldsymbol{y} \in \mathbb{R}^{M}$ $(M < N)$



$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{P} \\ \boldsymbol{V} \end{bmatrix}$$

$$\boldsymbol{w} = -\boldsymbol{Y}_{LL}^{-1}\boldsymbol{Y}_{L0}\boldsymbol{V}_{0}$$

$$\boldsymbol{M} = \left(\boldsymbol{Y}_{LL}^{-1} diag(\overline{\widehat{\boldsymbol{V}}}), -j\boldsymbol{Y}_{LL}^{-1} diag(\overline{\widehat{\boldsymbol{V}}})\right)$$

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_{00} & \boldsymbol{Y}_{0L} \\ \boldsymbol{Y}_{L0} & \boldsymbol{Y}_{LL} \end{bmatrix}$$

Admittance Matrix

$$min\|\boldsymbol{s}\|_1 + \gamma\|\boldsymbol{y} - \boldsymbol{\varphi}\boldsymbol{\psi}\boldsymbol{s}\|_2$$

$$subject\ to:$$

$$\|\boldsymbol{V} - \boldsymbol{A}.\boldsymbol{P} + \mathbf{w}\|_2 < \boldsymbol{\epsilon}$$

$$\widehat{\boldsymbol{x}} = \boldsymbol{\psi}\boldsymbol{s}$$

The recovered signal

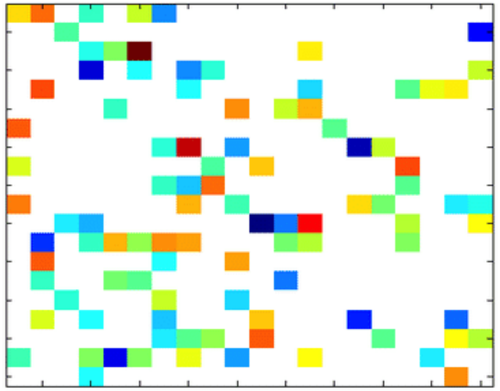
$\boldsymbol{s}$: Sparse signal
$\boldsymbol{y}$: Compressed measurement (Power and Voltage $(\boldsymbol{P}, \boldsymbol{V})$)
$\boldsymbol{\varphi}$: Measurement matrix
$\boldsymbol{\psi}$: Sparsifying matrix (e.g., wavelet)

Matrix completion based DSSE

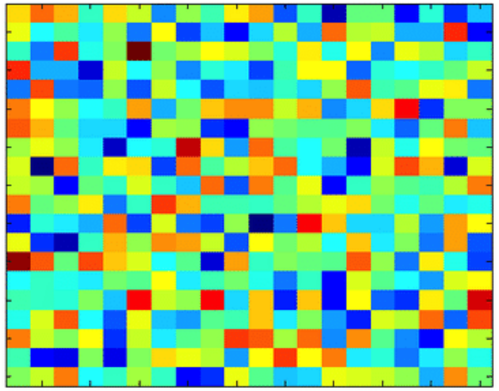
$\boldsymbol{X} = [\boldsymbol{P} \quad \boldsymbol{Q} \quad Re(\boldsymbol{V}) \quad Im(\boldsymbol{V}) \quad |\boldsymbol{V}|]$ ← Matrix of interest

$$\boldsymbol{Y} = (P_\Omega(\boldsymbol{X}))_{ij} = \begin{cases} \boldsymbol{X}_{ij} & (i,j) \in \Omega \\ 0 & otherwise \end{cases}$$

Only a sampled set of entries $\boldsymbol{X}_{ij}$, $(i,j) \in \Omega$ are available.



Measurement (Incomplete) Matrix



Complete Matrix

GOAL: recover the unknown entries in $\boldsymbol{Y}$

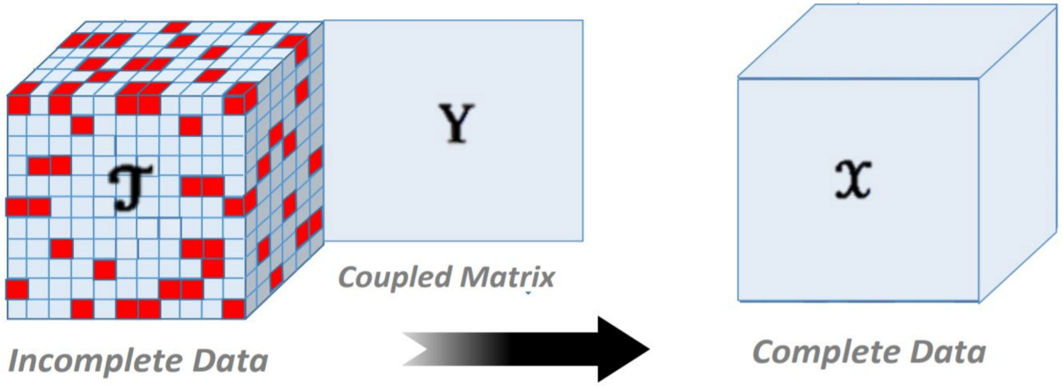
$$\begin{aligned} \min_{\mathbf{X}} \quad & \| \mathbf{X} \|_* + \lambda_1 \| \boldsymbol{Y} - P_\Omega(\boldsymbol{X}) \|_F^2 \\ \text{subject to} \quad & \mathbf{v} = \mathbf{Mz} + \mathbf{w} \end{aligned}$$

Tensor Completion based DSSE

When, we have different matrices of interest across time:

$$\mathcal{M} = [\boldsymbol{M}_1 \quad \boldsymbol{M}_2 \quad \ldots. \quad \boldsymbol{M}_{t-1} \quad \boldsymbol{M}_t]$$

Tensor of interest



$$\min. \left(\sum_{i=1}^{3} \beta_i \, \|\boldsymbol{\mathcal{X}}_i\|_* \right)$$

$$subject\ to:$$

$$P_\Omega(\boldsymbol{\mathcal{X}}_i) = P_\Omega(\boldsymbol{\mathcal{M}}_i)$$

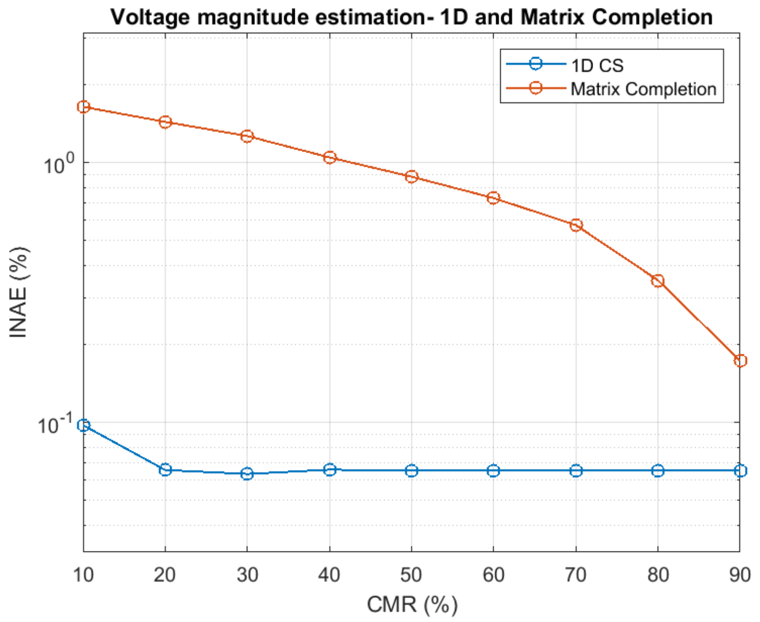
$$\boldsymbol{V_i} = \boldsymbol{A}.\boldsymbol{P_i} + \boldsymbol{W}$$

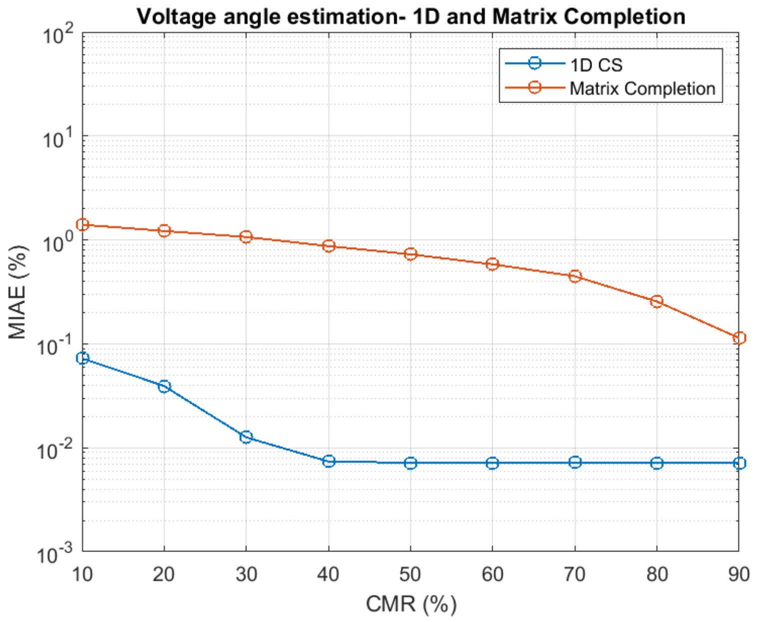
GOAL: recover the unknown entries in $\mathcal{M}$

β_i - positive constant value.
Suffix i denotes the unfolding operation along the mode i

Sample Simulation results

Comparison of Sparsity-based DSSE - 1D- Compressive Sensing and Matrix Completion based DSSE approaches [4]





Key observations-

- *Estimate system states accurately even at 30% of available measurements.*
- *Compressive sensing exploits sparsity of a signal, hence performs better than Matrix completion*

R. Madbhavi, B. Natarajan and B. Srinivasan, "Enhanced Tensor Completion Based Approaches for State Estimation in Distribution Systems," in *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 5938-5947, Sept. 2021

[4] Dahale, Shweta, Hazhar Sufi Karimi, Kexing Lai, and Balasubramaniam Natarajan. "Sparsity Based Approaches for Distribution Grid State Estimation-A Comparative Study." IEEE Access, (2020): 198317-198327.

Alam, SM Shafiul, Balasubramaniam Natarajan, and Anil Pahwa. "Distribution grid state estimation from compressed measurements." *IEEE Transactions on Smart Grid* 5.4 (2014): 1631-1642.

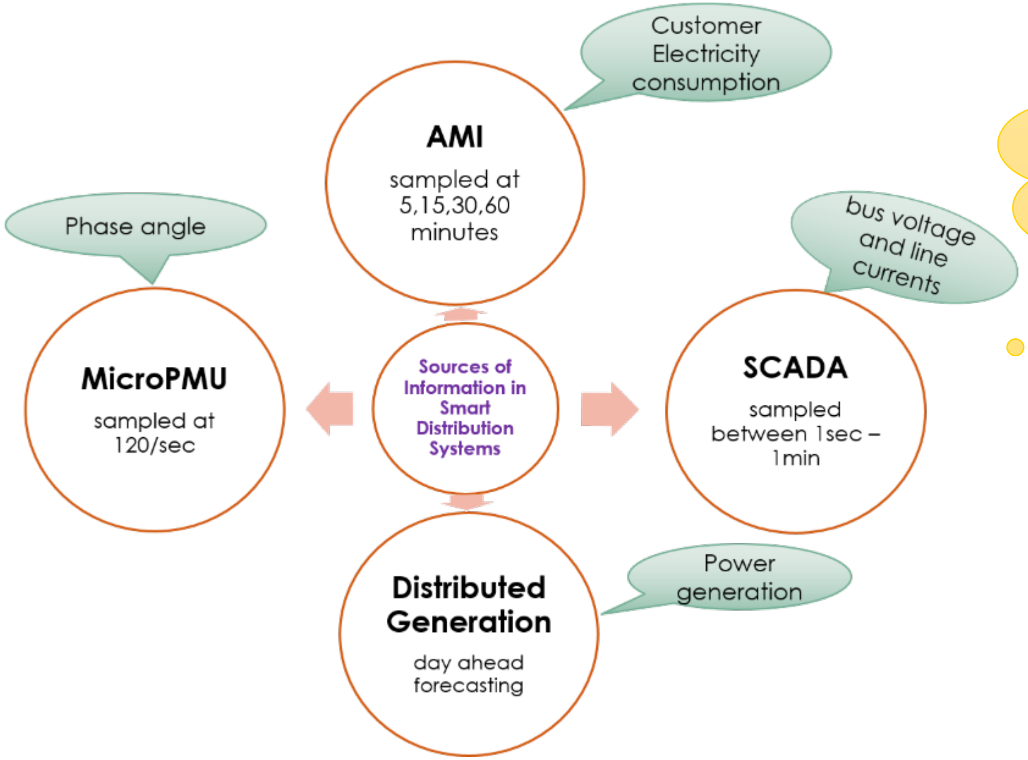
Doing More with Less!

Distribution system state estimation (DSSE)

Integrating multi-timescale measurements

Integrating Topology and Phase Identification

Multi-time scale data issue



Sources of information in smart distribution systems

How to reconcile the heterogenous measurements and estimate states in low-observable distribution systems?

Issues -

- Limited measurements
- Unevenly sampled data
- Missing and corrupt data
- Different accuracy levels

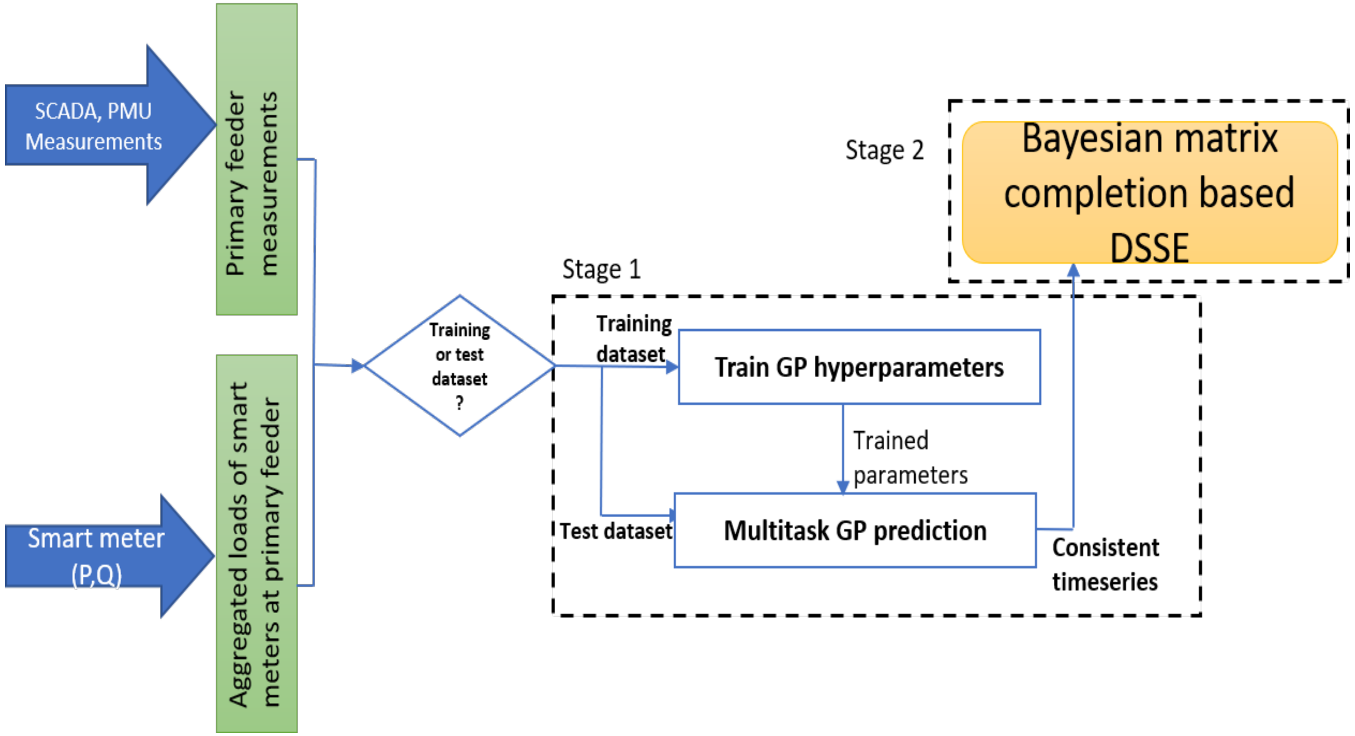
Related work

- Authors in [5] incorporated a simple linear interpolation technique along-with Weighted least squares state estimation technique.
- [6] uses historical timeseries data from smart meters to find kNN cluster centers and then use for imputation approach to impute the time-series data.
- Approach in [7] uses a data collation method to reconcile heterogeneous measurements and a Kalman filter method to perform DSSE. The data collation consists of an exponential moving average method to extrapolate the slow-rate measurements.

Shortcomings

- Existing approaches do not effectively exploit the correlations among the time-series data.
- Requires large measurement redundancy and performs poorly in case of intermittent measurements.
- Do not provide uncertainty bounds on the imputed measurements.

Proposed GP framework



Proposed GP Framework [8], [9]

Stage 1 – Multitask GP approach.
Stage 2 – Bayesian matrix completion based DSSE.

- Stage 1 reconciles measurements at finest time resolution and provides uncertainty bounds
- Stage 2 incorporates the measurements and the associated uncertainty to estimate the system states in low-observable distribution systems.

[8] Dahale, Shweta, and Balasubramaniam Natarajan. "Multi time-scale imputation aided state estimation in distribution system." 2021 IEEE Power & Energy Society General Meeting (PESGM IEEE, 2021

[9] Dahale, Shweta, and Balasubramaniam Natarajan. "Bayesian Framework for Multi-timescale State Estimation in Low-Observable Distribution Systems." IEEE Transactions on Power Systems (2022).

Stage 1: Multitask GP approach

- Gaussian process (GP) is defined as a collection of random variables, any finite collection of them has a joint normal distribution.
- GP prior function is defined over time for each measurement as,

$$f(\boldsymbol{x}) = \mathcal{GP}\left(m_\phi(\boldsymbol{x_i}), k_\theta(\boldsymbol{x_i}, \boldsymbol{x_i}')\right)$$

where, $m_\phi(\cdot)$ is the mean function (e.g., deep neural network)

$k_\theta(\cdot,\cdot)$ is the kernel function (e.g., radial basis kernel function)

GP prior parameters ψ = $\{\phi, \theta\}$ optimized by maximizing log marginal likelihood

Inputs:

- Incomplete $\mathbf{P}$, $\mathbf{Q}$ measurements sampled at 15-min interval
- Incomplete $\mathbf{V}$ measurements sampled at 1-min interval

Output:

- $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{V}$ measurements obtained at finest time resolution along-with their variances.

Training dataset $\mathcal{M}$ - $\{\boldsymbol{x_i}, \boldsymbol{y_i}\}_{i=1}^{P}$

Test dataset $\mathcal{D}$ = $\{\tilde{\boldsymbol{x}}_{\boldsymbol{i}}, \tilde{\boldsymbol{y}}_{\boldsymbol{i}}, \tilde{\boldsymbol{x}}_{\boldsymbol{i}}{}^*\}_{i=1}^{P}$

$\mathcal{M}$

GP Prior - $f(\boldsymbol{x})$ = $\mathcal{GP}\left(m_\phi(\boldsymbol{x_i}), k_\theta(\boldsymbol{x_i}, \boldsymbol{x_i}')\right)$

Likelihood - $y(\boldsymbol{x}) \sim \mathcal{N}(f(\boldsymbol{x}), \sigma^2)$

$\psi = \{\phi, \theta\}$

Goal: Predict $\{\tilde{\boldsymbol{y}}_{\boldsymbol{i}}{}^*\}_{i=1}^{P}$ using $\tilde{\boldsymbol{x}}_{\boldsymbol{i}}$ and $\tilde{\boldsymbol{y}}_{\boldsymbol{i}}$ and GP prior parameters.

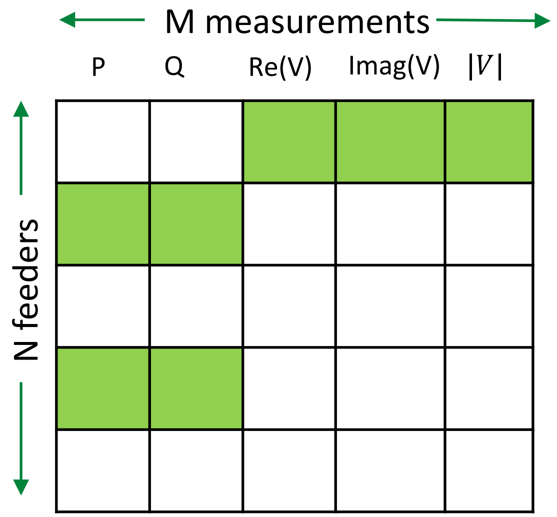
$\mathcal{D}$

$p(\tilde{\mathbf{y}}_i^*|\tilde{\mathbf{x}}_i^*, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \mathcal{N}(\mathbf{m}^*, \mathbf{K}^*)$

where $\mathbf{m}^* = m_\phi(\tilde{\mathbf{x}}^*) + K_\theta^{*x}(K_\theta^{xx} + \sigma^2 I)^{-1}(\tilde{\mathbf{y}} - m_\phi(\tilde{\mathbf{x}}))$

$\mathbf{K}^* = K_\theta^{**} - K_\theta^{*x}(K_\theta^{xx} + \sigma^2 I)^{-1}(\tilde{\mathbf{y}} - K_\theta^{x*})$

Stage 2 -Bayesian Matrix completion based DSSE



$\mathbf{a}_{.i}$ - i^{th} column of A
$\mathbf{a}_{i.}$ - i^{th} row of A

A) Observation and noise models

$$\mathbf{D} = P_{\Omega}(\mathbf{X} + \mathbf{N})$$

$$\mathbf{p}(\mathbf{D}|\mathbf{A},\mathbf{B}) = \prod_{(i,j)\in\Omega} \mathcal{N}(D_{ij}|X_{ij},\beta_{ij}^{-1})$$

GP output $-\mathbf{D}$ and β

B) Low-Rank Modeling

- $\mathbf{X} = \mathbf{A}\mathbf{B}^{\mathbf{T}} = \sum_{i=1}^{k} \mathbf{a}_{.i}\mathbf{b}_{.i}^{T}$

$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}}$. Set $\mathbf{A} = \mathbf{U}\mathbf{S}^{\frac{1}{2}}$ and $\mathbf{B}^{\mathbf{T}} = \mathbf{S}^{1/2}\mathbf{V}^{\mathbf{T}}$

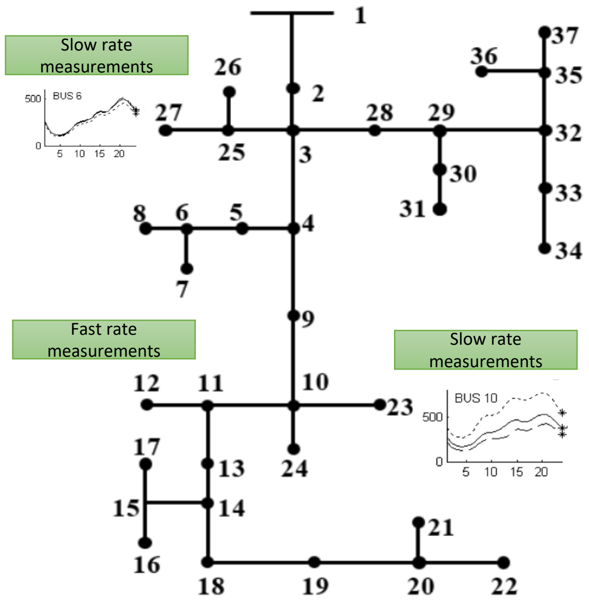
- $\mathbf{p}(\mathbf{A}|\boldsymbol{\gamma}) = \prod_{i=1}^{n} \mathcal{N}(\mathbf{a}_{.i}|\mathbf{0},\gamma_i^{-1}\mathbf{I}_m)$
- $\mathbf{p}(\mathbf{B}|\boldsymbol{\gamma}) = \prod_{i=1}^{k} \mathcal{N}(\mathbf{b}_{.i}|\mathbf{0},\gamma_i^{-1}\mathbf{I}_n)$
- $\mathbf{p}(\gamma_i) = Gamma(a,\frac{1}{b})$

C) Obtain an estimate of $\mathbf{A}, \mathbf{B}$ and $\boldsymbol{\gamma}$ by using the joint probability distribution given as,

$$\mathbf{p}(\mathbf{D},\mathbf{A},\mathbf{B},\boldsymbol{\gamma}) = \mathbf{p}(\mathbf{D}|\mathbf{A},\mathbf{B})\mathbf{p}(\mathbf{A}|\boldsymbol{\gamma})\mathbf{p}(\mathbf{B}|\boldsymbol{\gamma})\mathbf{p}(\boldsymbol{\gamma})$$

→ Joint probability distribution

Simulation results

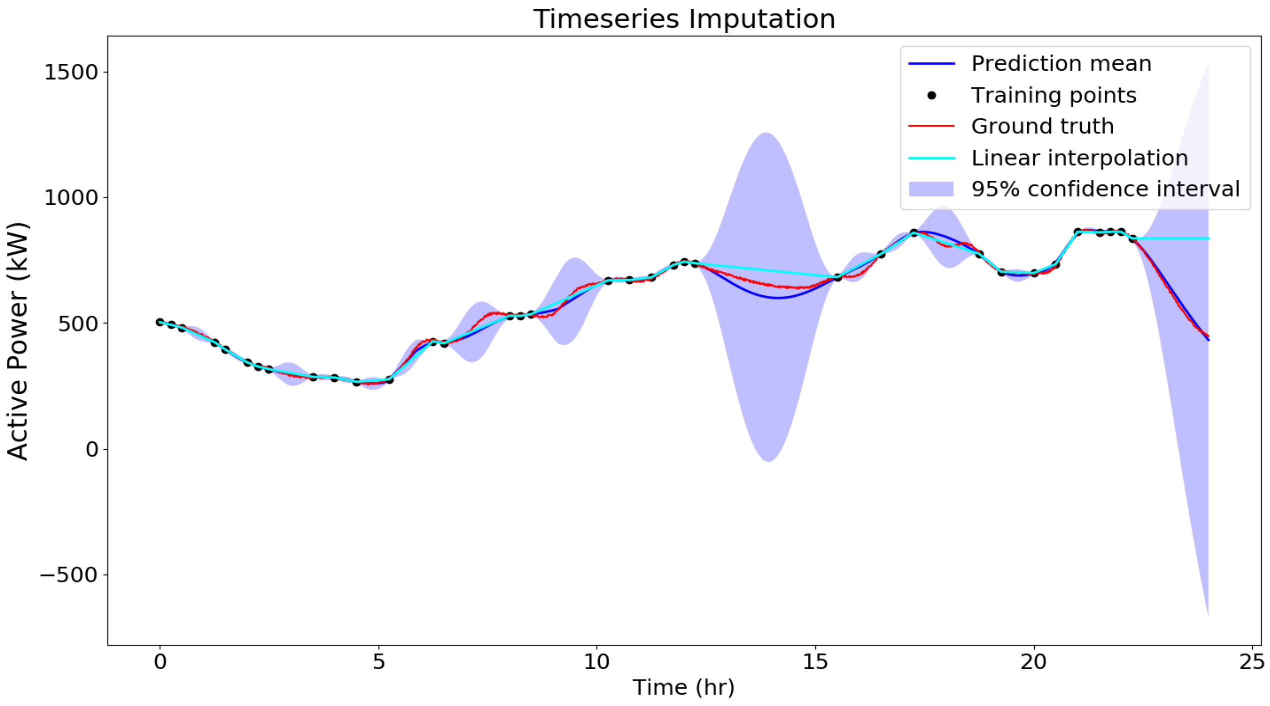


IEEE 37 test system

Slow rate measurements – Averaged at 15 min
Fast rate measurements – Sampled at 1 min

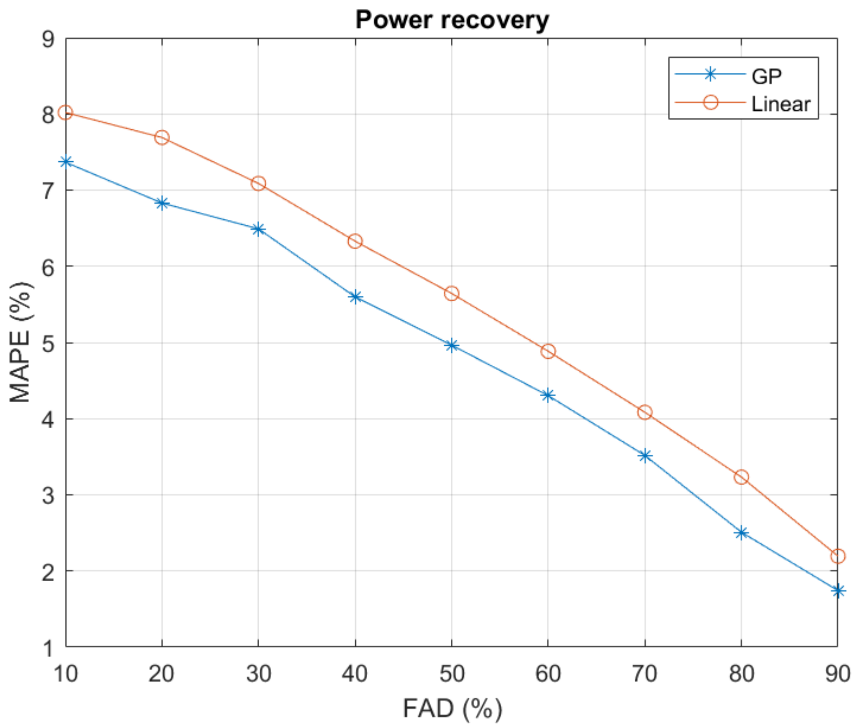
parameter	value
MLP layer	Two hidden layer (size 64 each)
Activation	ReLU
No. of epochs	20
Optimizer	Adam

Bayesian MC initial parameters	value
Initial γ	10e-5
Convergence criteria	10e-17

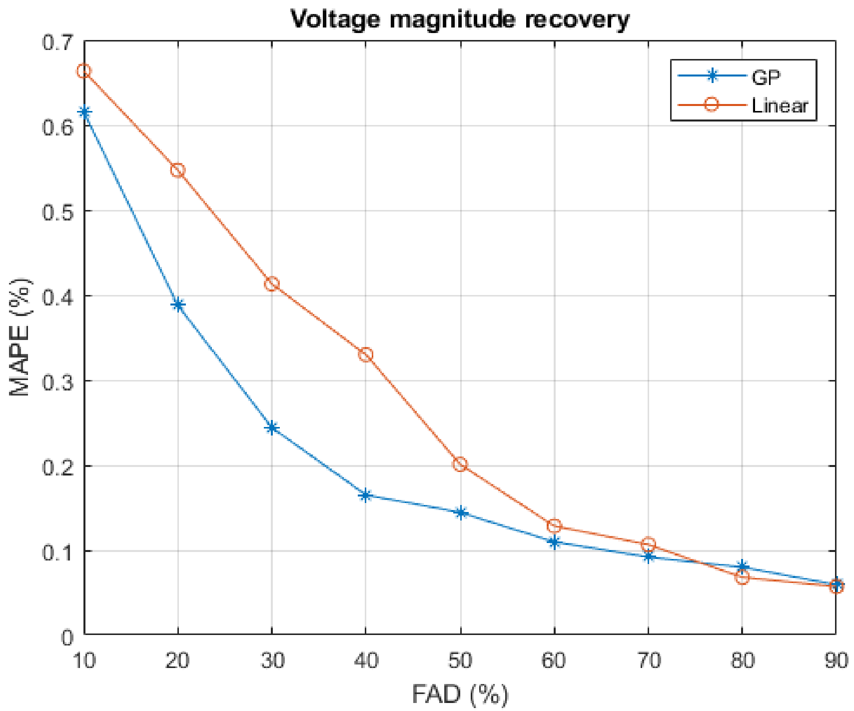


Data imputation of a slow-rate measurement timeseries at node 4

Simulation results



Power recovery performance at different FADs



Voltage magnitude recovery performance at different FADs

Key observations:

- *State estimates using Bayesian matrix completion with measurements from multitask GP approach performs better than linearly interpolated timeseries measurements.*

Multitask Recursive GP approach

- Recursively imputes multiple unevenly sampled measurements [10]
- Incorporate spatial information among measurements by leveraging graphical structure of the grid.
- A computationally efficient approach with the flexibility to perform batch-wise or real-time processing of measurements.

[10] Dahale, Shweta, and Balasubramaniam Natarajan. "Recursive Gaussian Process over graphs for Integrating Multi-timescale Measurements in Low-Observable Distribution Systems " IEEE Transactions on Power Systems (under review).

Multitask Recursive GP approach

Proposed approach: Recursive GP with/ without graphs

- `RGP-G/RGP Interpolation`
- **RGP-G/ RGP Prediction**

Goal: Recursively update the GP function and perform imputation as measurements are received at time t =1,…,T

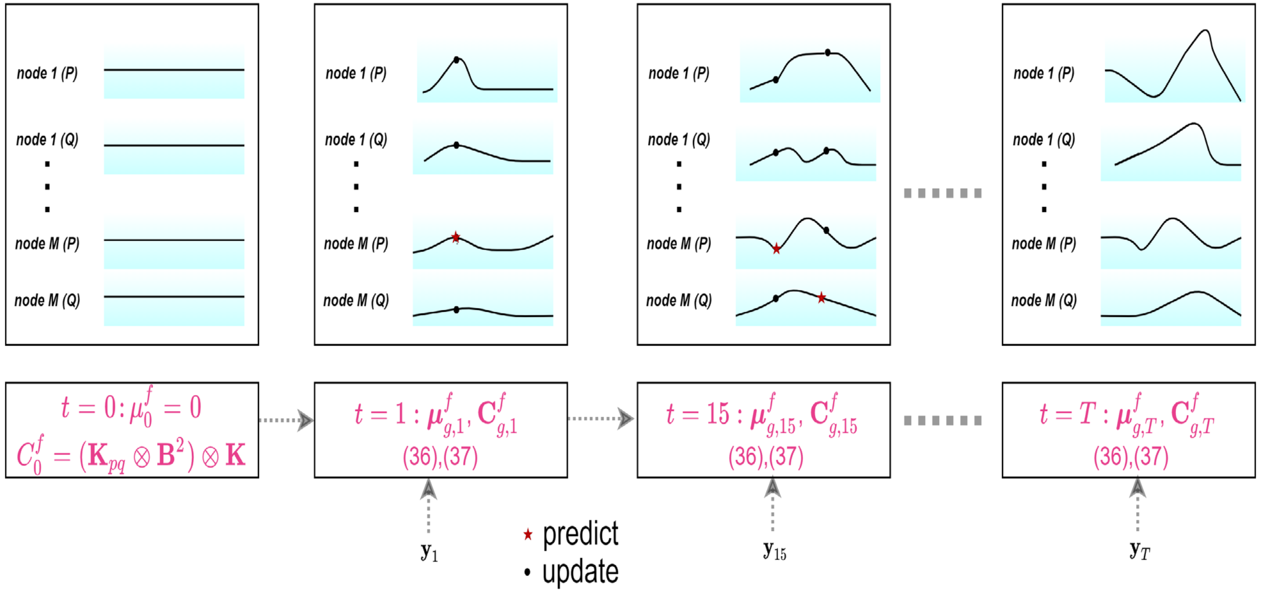
A. RGP-G Interpolation approach

Performs imputation after a batch of measurements in set time-frame are received.

Initialize: GP prior function with zero mean and covariance
Inputs: Measurements received upto time T.
Consists of two steps:

- **a)** **Predict:** Infer the joint probability GP function using measurements received upto time t -1
- **b)** **Update:** Update the GP function with new measurements received at time t using Kalman filter step

Output: Impute at finest time resolution using updated GP mean and covariance function.



RGP-G interpolation approach

Multitask Recursive GP approach

B. RGP-G Prediction approach

Performs future predictions at desired time resolution using the previously obtained measurements.

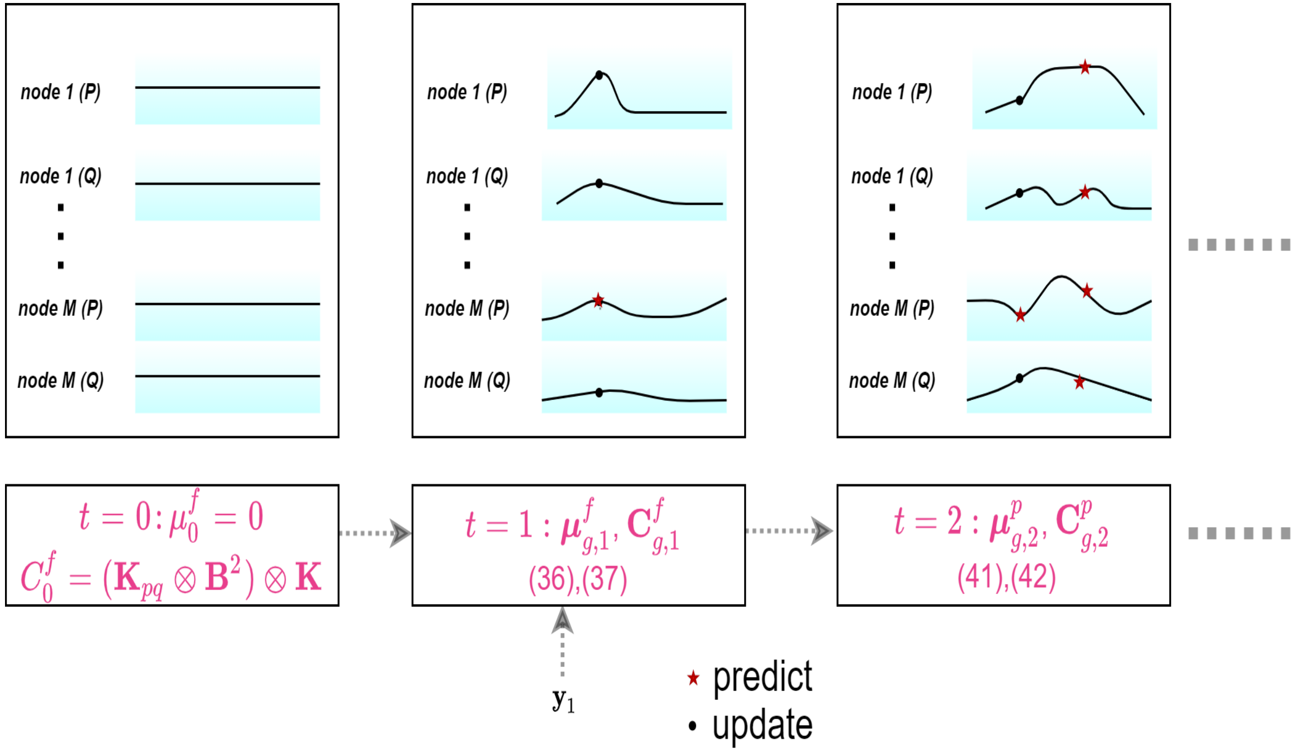
Initialize: GP prior function with mean and covariance
Inputs: Measurements received at time $t = 1$.
Consists of two steps:

a) **Predict:** Infer the joint GP prior using measurements received upto time t -1

b) **Update:** Update the GP function with new measurements received at time t using Kalman filter step

c) Forecast measurements at time t+1 using GP function updated at previous times. Continue until next measurements are received.

Output: Predicted measurements at finest time resolution



RGP-G prediction approach

Simulation results

- IEEE 37 bus system
- AMI and SCADA measurements recursively arrives at time $t = 1, \ldots, T$
- Predictions of AMI measurements: 1-min ahead
- Use the previously updated GP function for predictions.

Scenario	RGP-G prediction	Exponential moving average [11]
0% missing	2.26%	6.44%
10% missing	3.5%	7.75%
20% missing	4.8%	8.13%

MAPE of imputed time-series data of AMI measurements

Key observations:

- *Recursive GP prediction accurately predicts the AMI measurements at finest time resolution.*
- *Offers nearly 40% improvement compared to the exponential moving average method [11]*

Doing More with Less!

Distribution system state estimation (DSSE)

Integrating multi-timescale measurements

Integrating Topology and Phase Identification

System-related Uncertainties

Basic Matrix completion

$$\min_{\mathbf{X}} \quad \| \mathbf{X} \|_* \; + \lambda_2 \| \boldsymbol{Y} - P_\Omega(\boldsymbol{X}) \|_F^2$$
$$\text{subject to} \quad \mathbf{v} = \mathbf{Mz} + \mathbf{w}$$

Linearized Power Flow

$$\boldsymbol{W} = -\boldsymbol{Y}_{LL}^{-1}\boldsymbol{Y}_{L0}\boldsymbol{V}_0$$

$$\boldsymbol{M} = \left(\boldsymbol{Y}_{LL}^{-1} diag(\bar{\widehat{\boldsymbol{V}}}), -j\boldsymbol{Y}_{LL}^{-1} diag(\bar{\widehat{\boldsymbol{V}}})\right)$$

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_{00} & \boldsymbol{Y}_{0L} \\ \boldsymbol{Y}_{L0} & \boldsymbol{Y}_{LL} \end{bmatrix}$$

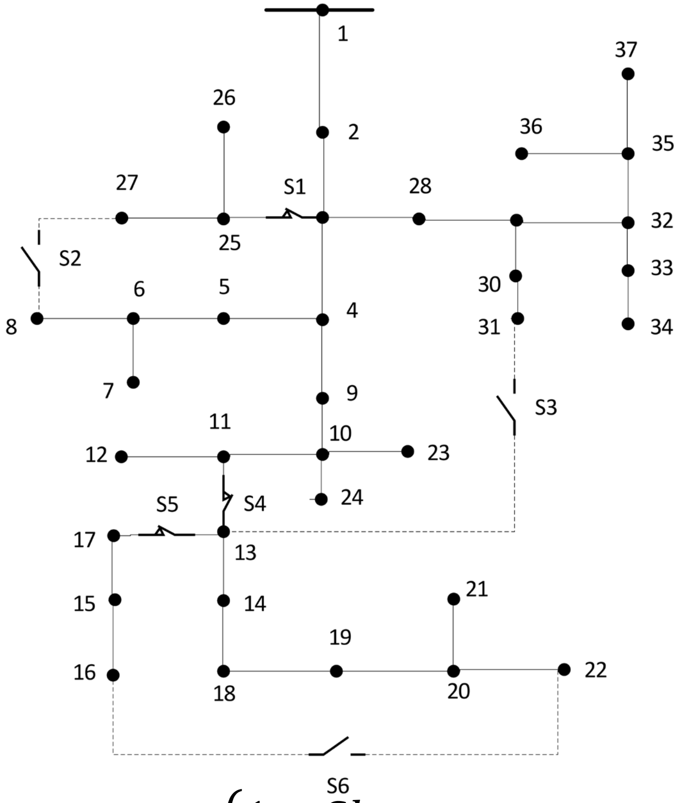
Admittance Matrix

Uncertainties

1- Topology Error (Lacking information about connection among the nodes)

2 – Phase Error (incorrect information about the phase labels)

Joint Topology Identification and State estimation



$$s_i = \begin{cases} 1 & Close \\ 0 & Open \end{cases}$$

Topology Identification:
Knowledge of switch status (open/close) using available measurements

$$min\|\boldsymbol{S}\|_1 + \lambda\|\boldsymbol{y} - \boldsymbol{\varphi\psi S}\|_2$$
$$subject\ to:$$
$$\boldsymbol{Y}_{LL}\boldsymbol{V} - \boldsymbol{M}_y.\boldsymbol{P} + \boldsymbol{Y}_{L0}\boldsymbol{V}_0 = \boldsymbol{0}$$

Linearized Power Flow

$$\boldsymbol{Y}_{LL} = \overline{\boldsymbol{Y}} + s_1\boldsymbol{Y}_1 + \cdots + s_n\boldsymbol{Y}_n$$

$$\boldsymbol{M}_y = \left(diag(\widehat{\overline{\boldsymbol{V}}}), -jdiag(\widehat{\overline{\boldsymbol{V}}})\right)$$

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_{00} & \boldsymbol{Y}_{0L} \\ \boldsymbol{Y}_{L0} & \boldsymbol{Y}_{LL} \end{bmatrix}$$ Admittance Matrix

MINLP

Goals:
1- Estimate Voltage and Power
2- Estimate Switch Status

[12] H. S. Karimi and B. Natarajan, "Joint Topology Identification and State Estimation in Unobservable Distribution Grids," in IEEE Transactions on Smart Grid, vol. 12, no. 6, pp. 5299-5309, Nov. 2021, doi: 10.1109/TSG.2021.3102179.

Approach

1- MILP

$$\min\|\boldsymbol{S}\|_1 + \lambda\|\boldsymbol{y} - \boldsymbol{\varphi\psi S}\|_2$$
$$subject\ to:$$
$$\boldsymbol{Y}_{LL}\,\boldsymbol{V} - \boldsymbol{M}_{\boldsymbol{y}}.\boldsymbol{P} + \boldsymbol{Y}_{L0}\boldsymbol{V}_0 = \boldsymbol{0}$$
$$-(1-s_i)F \le \boldsymbol{U}_i - \boldsymbol{V} \le (1-s_i)F$$
$$-Fs_i \le \boldsymbol{U}_i \le Fs_i$$

Using an auxiliary variable, we remove the nonlinear term:
$$\boldsymbol{U}_i = s_i\boldsymbol{V}$$

$$\boldsymbol{Y}_{LL} = \overline{\boldsymbol{Y}} + s_1\boldsymbol{Y}_1 + \cdots + s_n\boldsymbol{Y}_n$$

2- Convex Relaxation

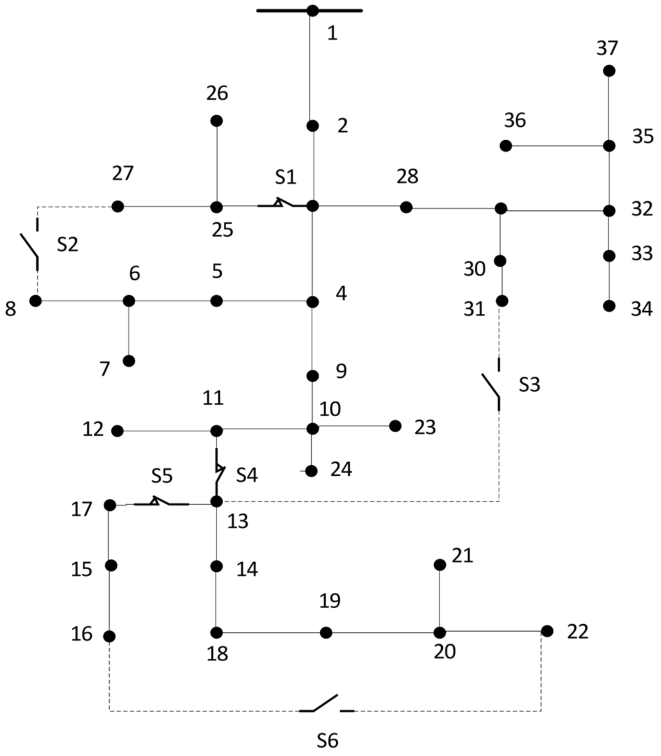
We consider x_i as switch status where
$$0 \le x_i \le 1$$

$$\boldsymbol{Y}_{LL} = \overline{\boldsymbol{Y}} + x_1\boldsymbol{Y}_1 + \cdots + x_n\boldsymbol{Y}_n$$

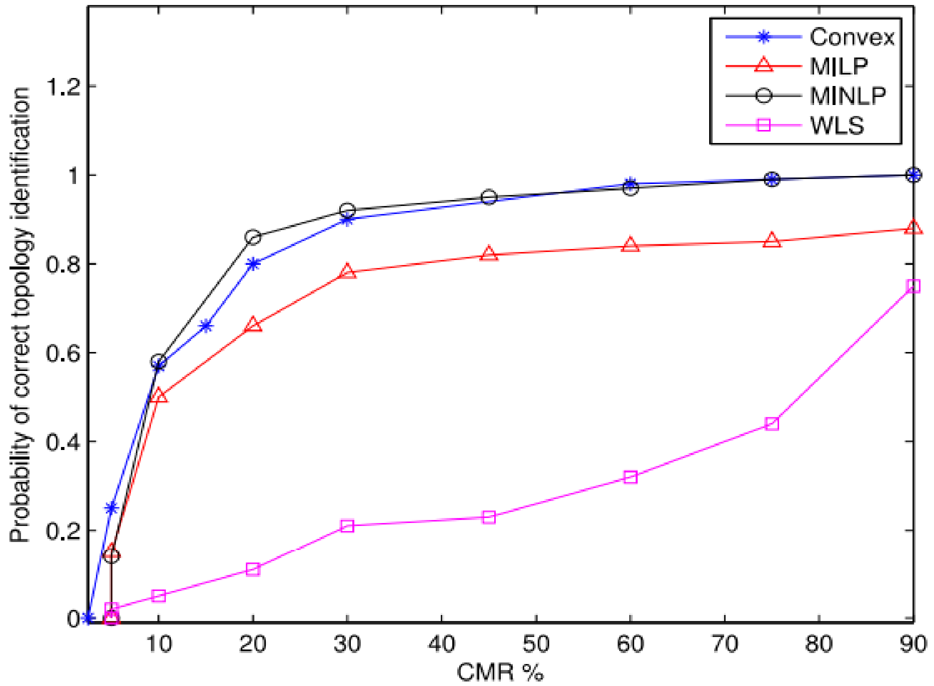
Solution: Alternating Minimization

$$\min\|\boldsymbol{S}\|_1 + \lambda\|\boldsymbol{y} - \boldsymbol{\varphi\psi S}\|_2$$
$$subject\ to:$$
$$\boldsymbol{Y}_{LL}\,\boldsymbol{V} - \boldsymbol{M}_{\boldsymbol{y}}.\boldsymbol{P} + \boldsymbol{Y}_{L0}\boldsymbol{V}_0 = \boldsymbol{0}$$

Results



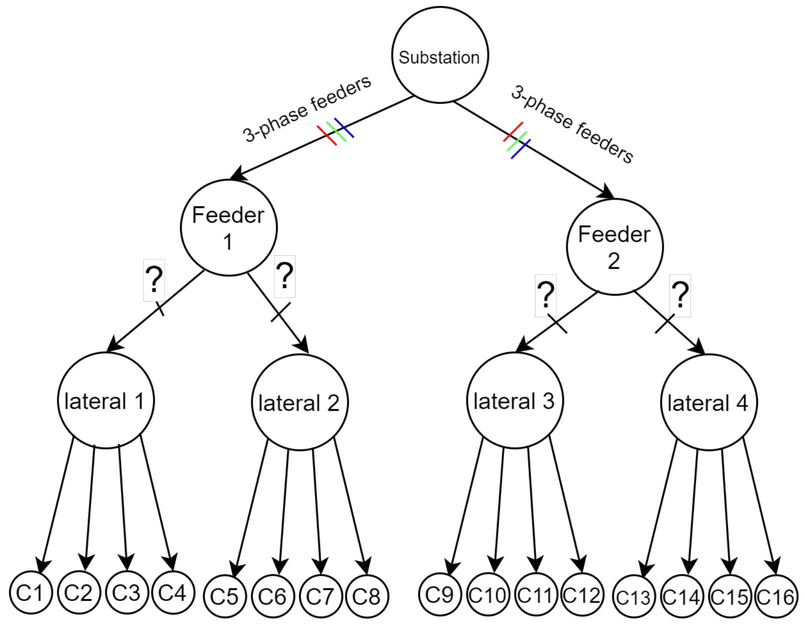
IEEE 37-node test feeder with 6 switches



Probability of accurate topology identification – MINLP, MILP and convex approaches

Phase Identification

- Phase changes occur frequently in the distribution networks which are not always tracked continuously.
- Utilities have limited or unreliable information to identify the phase labels (A, B or C).
- Distribution grid is generally unobservable and measurement data is limited.



Tree representation of network topology with ? representing the unknown phase labels

Question 1: How to provide accurate phase information using limited measurements in the distribution grid?

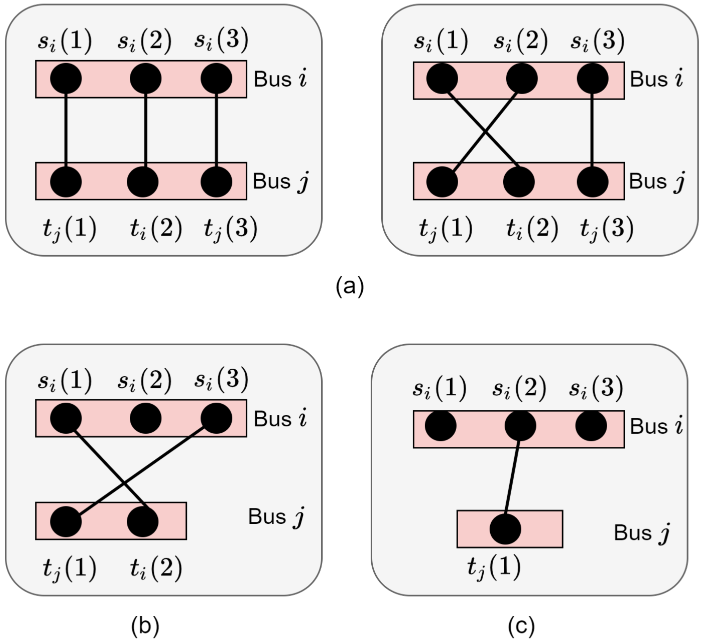
Proposed approach [13]

Theorem 1 - *In a multi-phase distribution grid, if two terminal buses of a branch are connected on the same phase, their phase voltage correlation is the largest [14].*

Inputs:
- Voltage magnitude time-series measurements.
- Set of known bus label indices.

Recover the Graph phase connectivity by solving an optimization problem using -
Objective function: Minimize the variation of voltage signals in the graph and use the known phase label information

Constraints: Must-link constraints, cannot-link constraints, Valid set of Adjacency matrices.



Different possible configurations for a two bus connection. (a) three phase end (b) two phase end (c) one phase end

[13] Dahale, Shweta and B.Natarajan, “Phase Identification in Unobservable Distribution Systems”, (manuscript in preparation).

Proposed approach

Consider a matrix of voltage magnitude measurements $\mathbf{V} \in \mathbb{R}^{M\times N}$, where N - total time over which the samples are collected.

$\mathbf{V} = [\mathbf{V}_1^{p_1}, \dots, \mathbf{V}_K^{p_K}]^\intercal$ where $\mathbf{V}_i^{p_i}$ is the nodal voltages on bus i with phases p_i,

(e.g.,$\mathbf{V}_i^{p_i} = [\mathbf{V}_i^a, \mathbf{V}_i^b, \mathbf{V}_i^c]$).

The aim is to determine the ϕ labels with respect to the reference of substation phases.

- **Objective function**: Minimize the variation of voltage signals in the graph and use the known phase label information.
- **Constraints**: Must-link constraints, cannot-link constraints, Valid Adjacency matrix.

Known entries in the adjacency matrix ($P_\Omega(\mathbf{A_{known}})$) is defined as,

$$[P_\Omega(\mathbf{A_{known}})]_{mn} = \begin{cases} [\mathbf{A_{known}}]_{mn}, \text{if } (m,n) \in \Omega \\ 0, \text{otherwise} \end{cases}$$

$$\min_{\mathbf{A},\mathbf{S}} \quad \mathrm{tr}(\mathbf{V}^\intercal\mathbf{L}\mathbf{V}) + \eta \parallel \mathbf{vec}(\mathbf{A}) \parallel_1 + \beta \mathrm{tr}(\mathbf{S}^\intercal\mathbf{L}\mathbf{S}) + \lambda_1 \parallel P_\Omega(\mathbf{A} - \mathbf{A_{known}}) \parallel_F^2$$

$$0.5 \le \sum_{m=1}^{p_j} \mathbf{A}(\mathbf{s}_i(l), \mathbf{t}_j(m)) \le 1, \forall i, \forall j, i \sim j$$

$$0.5 \le \sum_{l=1}^{p_i} \mathbf{A}(\mathbf{s}_i(l), \mathbf{t}_j(m)) \le 1, \forall i, \forall j, i \sim j$$

Must-link constraints

$$\mathbf{A}(\mathbf{s}_i(l), \mathbf{s}_i(m)) \le 0, \mathbf{A}(\mathbf{t}_i(l), \mathbf{t}_i(m)) \le 0, \quad m, l = 1, \dots, p_i, i = 1, \dots, K.$$

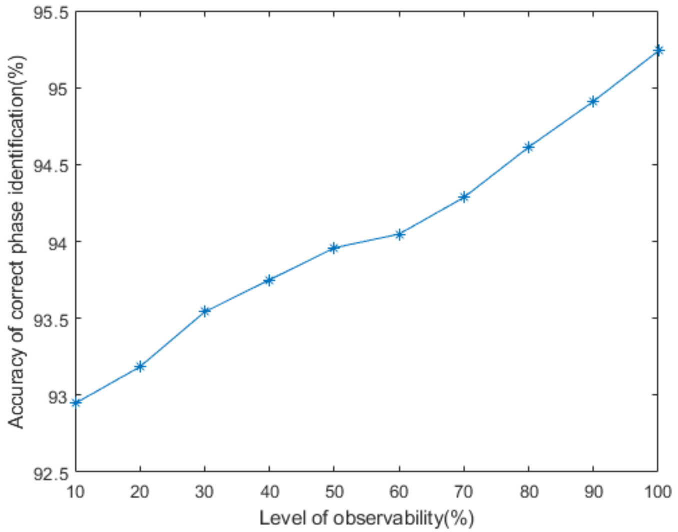
Cannot-link constraints

$$\mathbf{S}^\intercal\mathbf{S} = \mathbf{I}$$

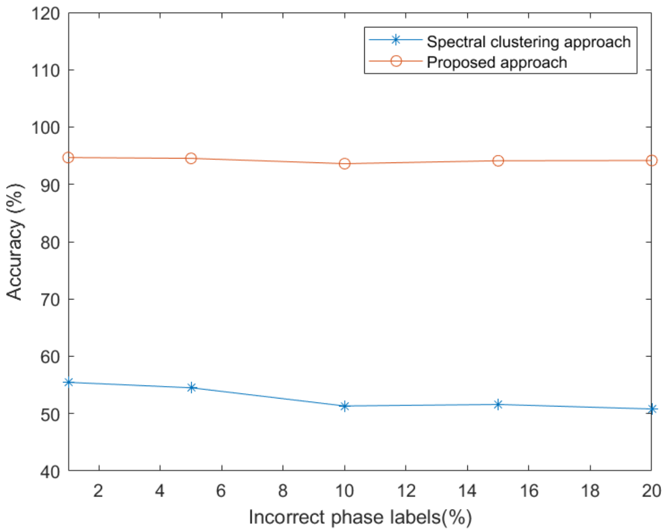
$$\mathbf{A} \in \mathcal{A}$$

Results

- We compared the proposed approach with Spectral Clustering approach [15] on IEEE 37 bus test system (40% known and 60% unknown).
- The power at each bus is an aggregation of customers which are randomly considered in the range 75-100.
- Power flow analysis is run for one day to generate voltage timeseries measurements.



Performance of phase identification approaches in the presence of limited spatial measurements



Performance of phase identification approaches in the presence of incorrect phase labels

Key observations:

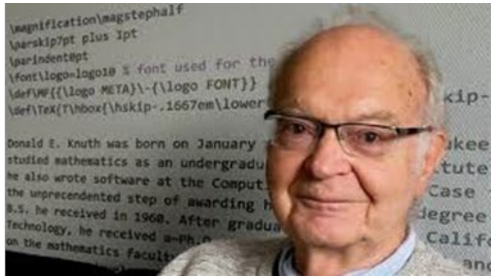
- Across all fraction of available data, the proposed approach is accurate in identifying the phase labels.
- The proposed approach is insensitive to incorrect bus phase labels and achieves high fidelity with an accuracy higher than 90%.

Take-home Message

Focus on Distribution system situational awareness NOW!

- Data and model driven approaches can help
- Less data ⟶ more opportunities to innovate!

"It is much more rewarding to do more with less."
-- Donald Knuth



References

[1] SM Shafiul Alam, Balasubramaniam Natarajan, and Anil Pahwa. Distribution grid state estimation from compressed measurements.   IEEE Transactions on Smart Grid, 5(4):1631–1642, 2014.

[2] Priya L Donti, Yajing Liu, Andreas J Schmitt, Andrey Bernstein, Rui Yang, and Yingchen Zhang. Matrix completion for low-observability voltage estimation.   IEEE Transactions on Smart Grid, 2019.

[3] Rahul Madbhavi, Hazhar Sufi Karimi, Balasubramaniam Natarajan, and Babji Srinivasan. Tensor completion based state estimation in distribution systems.   11th Conference on Innovative Smart Grid Technologies, Washington DC, Feb, 2020.

[4] Dahale, Shweta, Hazhar Sufi Karimi, Kexing Lai, and Balasubramaniam Natarajan. "Sparsity Based Approaches for Distribution Grid State Estimation-A Comparative Study." IEEE Access, (2020): 198317-198327.

[5] Antonio Gómez-Expósito, Catalina Gómez-Quiles, and Izudin Džafić. State estimation in two time scales for smart distribution systems.   *IEEE Transactions on Smart Grid*, 6(1):421–430, 2014.

[6] A. Al-Wakeel, J. Wu, and N. Jenkins, "State estimation of medium voltage distribution networks using smart meter measurements, "Applied energy, vol. 184, pp. 207–218, 2016.

[7] Karimipour, Hadis, and Venkata Dinavahi. "Extended Kalman filter-based parallel dynamic state estimation." *IEEE transactions on smart grid* 6.3 (2015): 1539-1549.

[8] Dahale, Shweta, and Balasubramaniam Natarajan. "Multi time-scale imputation aided state estimation in distribution system." 2021 IEEE Power & Energy Society General Meeting (PESGM). IEEE, 2021

[9] Dahale, Shweta, and Balasubramaniam Natarajan. "Bayesian Framework for Multi-timescale State Estimation in Low-Observable Distribution Systems." IEEE Transactions on Power Systems (2022).

[10] Dahale, Shweta, and Balasubramaniam Natarajan. "Recursive Gaussian Process over graphs for Integrating Multi-timescale Measurements in Low-Observable Distribution Systems " IEEE Transactions on Power Systems (under review).

[11] H. Karimipour and V. Dinavahi, "Extended kalman filter-based parallel dynamic state estimation," IEEE transactions on smart grid, vol. 6, no. 3, pp. 1539–1549, 2015

[12] H. S. Karimi and B. Natarajan, "Joint Topology Identification and State Estimation in Unobservable Distribution Grids," in IEEE Transactions on Smart Grid, vol. 12, no. 6, pp. 5299-5309, Nov. 2021, doi: 10.1109/TSG.2021.3102179.

[13] Dahale, Shweta and B.Natarajan, "Phase Identification in Unobservable Distribution Systems", (manuscript in preparation).

[14] Y. Liao, Y. Weng, G. Liu, Z. Zhao, C.-W. Tan, and R. Rajagopal, "Unbalanced multi-phase distribution grid topology estimation and bus phase identification," IET Smart Grid, vol. 2, no. 4, pp. 557–570, 2019

[15]] L. Blakely and M. J. Reno, "Phase identification using co-association matrix ensemble clustering," IET Smart Grid, vol. 3, no. 4, pp. 490–499, 2020.