Physics-Informed Machine Learning for Enhancing Robustness and Verification

Wenting Li, Deepjyoti Deka
Theory Division (T-5)
Los Alamos National Laboratory

April 5, 2023

LA-UR-23-23066

Acknowledge: Collaborators

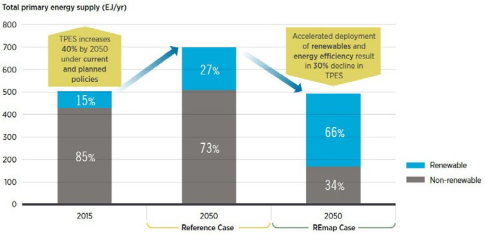
Deepjyoti Deka

Ren Wang

Mario Arrieta Paternina

Krishnamurthy (Dj) Dvijotham

The Increasing Renewables Challenge Power Grids



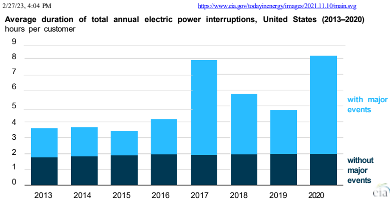


Fig. 1: Global Energy Transformation Prediction: https://www.irena.org (left); Average duration of power interruptions (right)

- When renewable energy **increases constantly**, faults or disturbances also become **more frequent** in these years.

Faults may Trigger Blackouts and Wildfires

Fig. 2: The 2003 blackout causes 50 million people in darkness; Wildfires in California in 2020 cost around $12 billiion

- Faults without efficient monitoring strategies may **trigger blackouts and wildfires**.

Faults may Trigger Blackouts and Wildfires

Fig. 2: The 2003 blackout causes 50 million people in darkness; Wildfires in California in 2020 cost around $12 billiion

- Faults without efficient monitoring strategies may **trigger blackouts and wildfires**.
- Machine learning is promising to be the solution, but **its reliability is not guaranteed** when applied to the stochastic power grids.

Black-box Machine Learning is Powerful but Fragile

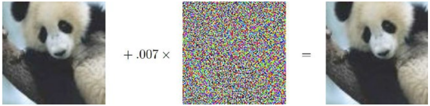


Fig. 3: Panda image is recognized as gibbon by adding trivial noise [1] Goodfellow et al. 2014. .

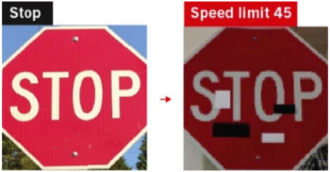


Fig. 4: The stop sign with some marks misleads the deep neural networks [2] K. Eykholt et al., 2018 IEEE/CVF.

- ML is vulnerable and can be **misled or attacked** by noise and perturbations;

Black-box Machine Learning is Powerful but Fragile

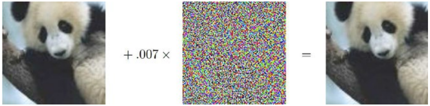


Fig. 3: *Panda image is recognized as gibbon by adding trivial noise [1] Goodfellow et al. 2014. .*

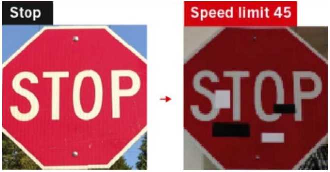


Fig. 4: *The stop sign with some marks misleads the deep neural networks [2] K. Eykholt et al., 2018 IEEE/CVF.*

- ML is vulnerable and can be **misled or attacked** by noise and perturbations;
- Perturbations in stochastic power grids **deteriorate** the performance of ML.

Central Ideas:

Robustify neural networks for fault location through:

1. Designing novel **architectures** by preserving physics
2. Developing physics-constrained optimization for **training**
3. **Certifying** training with physics-informed bound propagation

Central Ideas:

Robustify neural networks for fault location through:

1. Designing novel **architectures** by preserving physics
2. Developing physics-constrained optimization for **training**
3. **Certifying** training with physics-informed bound propagation

Problem Formulation of Fault Location

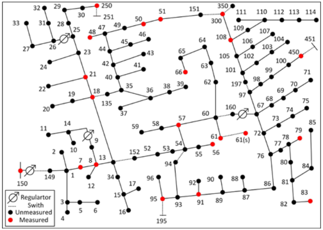


Fig. 5: *The IEEE 123-bus test feeder, where red ones are measured.*

- Given **a few measured node** voltages in Fig. 5, and **partial labels** denoting location.

[a] Brahma 2011; Džafić et al. 2016.
[b] Majidi, Etezadi-Amoli, and Fadali 2014; Che[r] Ghasemi, and Daisy 2018.

Problem Formulation of Fault Location

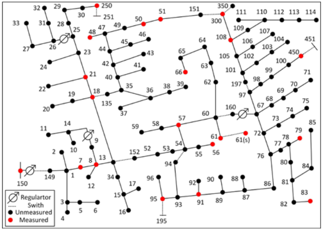


Fig. 5: The IEEE 123-bus test feeder, where red ones are measured.

- Given **a few measured node** voltages in Fig. 5, and **partial labels** denoting location.
- Goal: **Robust to** sparse observation, low label rates, varying loads and topology changes when predicting faults on the node level.

[a]Brahma 2011; Džafić et al. 2016.
[b]Majidi, Etezadi-Amoli, and Fadali 2014; Chen, Ghasemi, and Daisy 2018.

Problem Formulation of Fault Location

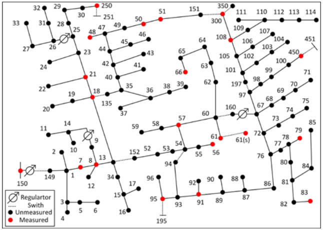


Fig. 5: The IEEE 123-bus test feeder, where red ones are measured.

- Given **a few measured node** voltages in Fig. 5, and **partial labels** denoting location.
- Goal: **Robust to** sparse observation, low label rates, varying loads and topology changes when predicting faults on the node level.
- Traditional methods:
  - Hardware[a];
  - Impedance-based, Traveling-wave-based, Knowledge-based[b]

---

[a] Brahma 2011; Džafić et al. 2016.
[b] Majidi, Etezadi-Amoli, and Fadali 2014; Che Ghasemi, and Daisy 2018.

Our Main Contributions [1]

Our approach: a two-stage graph neural network framework:

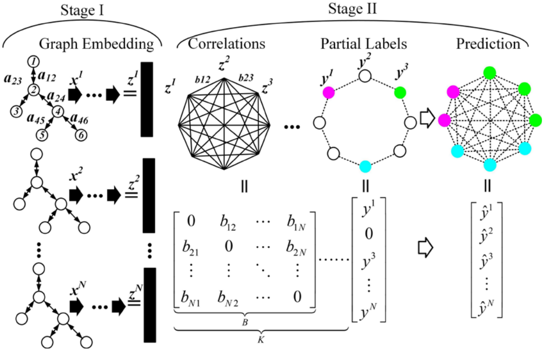


Fig. 6: Our two-stage graph neural network framework

[1] **Wenting Li**, Deepjyoti Deka, "PPGN: Physics-Preserved Graph Networks for Fault Location with Limited Observation and Labels", Hawaii International Conference on System Sciences (HICSS), 2023

Our Main Contributions

Our approach: a two-stage graph neural network framework:

- G_I in stage I learns the graph embedding of power networks for the challenge of **low observability**.

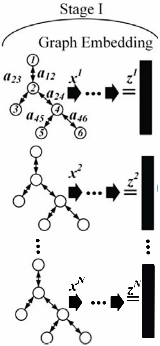


Fig. 7: Our graph learning at stage I

Our Main Contributions

Our approach: a two-stage graph neural network framework:

- G_I in stage I learns the graph embedding of power networks for the challenge of **low observability**.

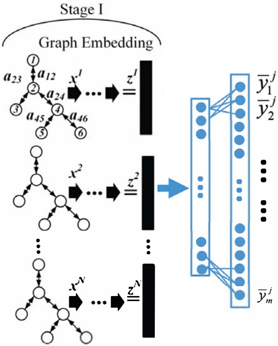


Fig. 8: Our graph learning at stage I

Our Main Contributions

Our approach: a two-stage graph neural network framework:

- G_I in stage I learns the graph embedding of power networks for the challenge of **low observability**.
- The key is the **adjustable adjacency** A of G_I using shortest distance.

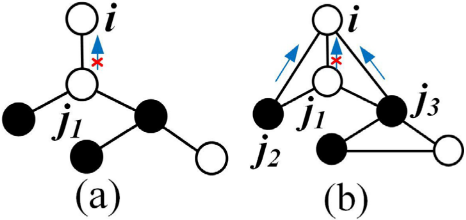


Fig. 9: (a) original graph; (b) reduced graph defined by A

Our Main Contributions

Our approach: a two-stage graph neural network framework:

- G_{II} in stage II further enhances location accuracy to face the **challenge of low label rates**.
  - The key of adjacency B of G_{II} with the output of G_I: **neighborhood** property [1].

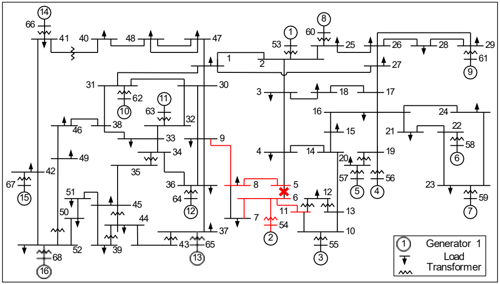


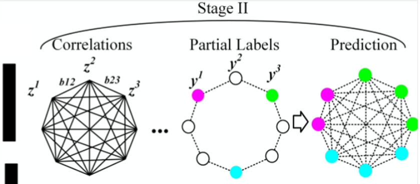
Fig. 10: The neighborhood property.

[1]W. Li, D. Deka, "Real-Time Faulted Line Localization and PMU Placement in Power Systems Through Convolutional Neural Networks", Transaction on Power System, 2019.

Our Main Contributions

Our approach: a two-stage graph neural network framework:

- G_{II} in stage II further enhances location accuracy to face the **challenge of low label rates**.
    - G_{II} represents the **correlations of labeled and unlabeled** data samples.



$$\underbrace{\underbrace{\begin{bmatrix} 0 & b_{12} & \cdots & b_{1N} \\ b_{21} & 0 & \cdots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \cdots & 0 \end{bmatrix}}_{B} \cdots\cdots}_{K} \begin{bmatrix} y^1 \\ 0 \\ y^3 \\ \vdots \\ y^N \end{bmatrix} \Rightarrow \begin{bmatrix} \hat{y}^1 \\ \hat{y}^2 \\ \hat{y}^3 \\ \vdots \\ \hat{y}^N \end{bmatrix}$$

Fig. 11: Our stage II graph learning framework

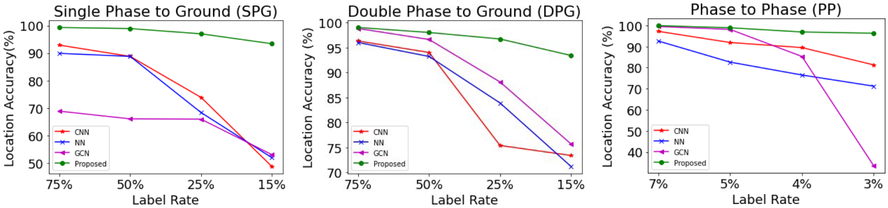


Fig. 12: LAR Comparison at different label rates[2]

- 24480 testing cases by **OpenDSS** in the IEEE 123-node benchmark system.

[2] $\text{LAR} = \frac{\text{The number of correctly located faults}}{\text{The total number of faults}}$, $\text{Label rate} = \frac{\text{The number of training data}}{\text{The total number of data}}$

Location Accuracy Rate (LAR) Comparison

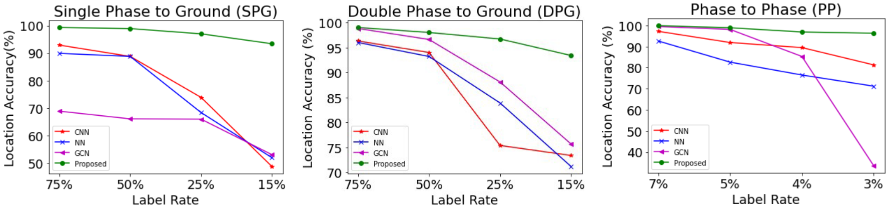


Fig. 12: LAR Comparison at different label rates[2]

- 24480 testing cases by **OpenDSS** in the IEEE 123-node benchmark system.
- System has voltage regulators, overhead/underground lines, switch shunts, and unbalancing loads that vary over time.

[2] $\text{LAR} = \frac{\text{The number of correctly located faults}}{\text{The total number of faults}}$, $\text{Label rate} = \frac{\text{The number of training data}}{\text{The total number of data}}$

Location Accuracy Rate (LAR) Comparison

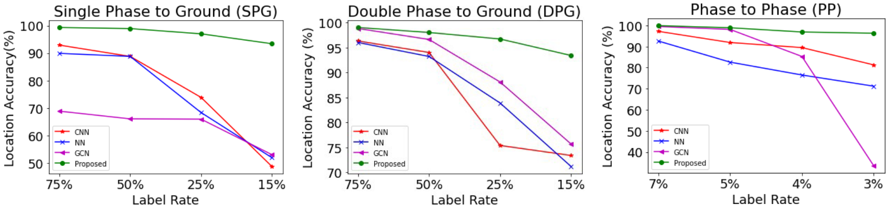


Fig. 12: LAR Comparison at different label rates[2]

- 24480 testing cases by **OpenDSS** in the IEEE 123-node benchmark system.
- System has voltage regulators, overhead/underground lines, switch shunts, and unbalancing loads that vary over time.
- Only **16% of the nodes** in the system are measured (21 measured nodes);

[2] $\text{LAR} = \frac{\text{The number of correctly located faults}}{\text{The total number of faults}}$, $\text{Label rate} = \frac{\text{The number of training data}}{\text{The total number of data}}$

Location Accuracy Rate (LAR) Comparison

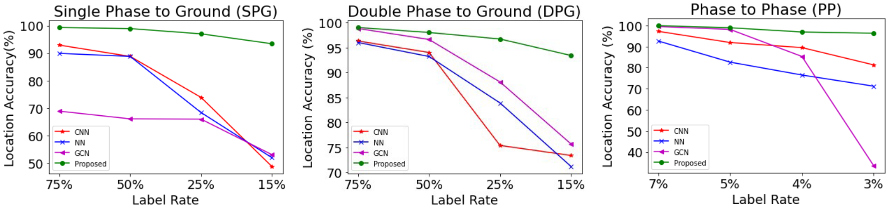


Fig. 12: LAR Comparison at different label rates[2]

- 24480 testing cases by **OpenDSS** in the IEEE 123-node benchmark system.
- System has voltage regulators, overhead/underground lines, switch shunts, and unbalancing loads that vary over time.
- Only **16% of the nodes** in the system are measured (21 measured nodes);
- **Outperforms CNN, NN, and GCN** for various faults.

[2] $\text{LAR} = \frac{\text{The number of correctly located faults}}{\text{The total number of faults}}$, $\text{Label rate} = \frac{\text{The number of training data}}{\text{The total number of data}}$

Robust to Load Variations

Table 1: *LARs (%)when Loads Vary in Different Ranges*

	Δp (p.u.)	0.53	0.58	0.64	0.69	0.74
SPG	CNN	93.9	85.3	84	83.9	82
	NN	92.5	80	77.4	76.7	74
	GCN	64.3	57.7	56.4	55.6	55.1
	Proposed	**98.9**	**96.6**	**96.3**	**95.8**	**95.1**
DPG	CNN	96.5	88.3	87.8	85.3	82.5
	NN	98	89.3	88.2	86.7	85.1
	GCN	98.3	84.0	83.7	82.2	78.8
	Proposed	**98.4**	**94.1**	**93.7**	**92.7**	**92.2**
PP	CNN	97.5	96.2	96.1	95.1	94.6
	NN	95.6	92.2	90.3	87.9	85.9
	GCN	99.5	96.5	96.5	96.6	96.7
	Proposed	**99.9**	**99.6**	**99.4**	**99.2**	**98.4**

- Generate **another 110160 faults** when load and topology change
- **No retraining** is needed.
- Achieves up to **15% improvement** than the baseline classifiers.

Central Ideas:

Robustify neural networks for fault location through:

1. Designing novel **architectures** by preserving physics
2. Developing physics-constrained optimization for **training**
3. **Certifying** training with physics-informed bound propagation

Background

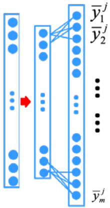


Fig. 13: Neural Networks

- **Motivations**: Perturbations in power grids **degrade** the performance of NNs, shown in the Table 2.

Table 2: The LAR when data perturbs due to different load variations with magnitudes δ_1 per unit (p.u.)

Load Variations (p.u.)	δ_1= 1	δ_1 =1.5	δ_1 =2	δ_1 =3
LAR (%)	96.25	81.61	71.96	57.5

Background

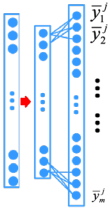


Fig. 13: Neural Networks

- **Motivations**: Perturbations in power grids **degrade** the performance of NNs, shown in the Table 2.
- **Goal:** Train neural networks to be robust to *natural perturbations* in power grids.

Table 2: The LAR when data perturbs due to different load variations with magnitudes δ_1 per unit (p.u.)

Load Variations (p.u.)	$\delta_1 = 1$	$\delta_1 = 1.5$	$\delta_1 = 2$	$\delta_1 = 3$
LAR (%)	96.25	81.61	71.96	57.5

Background

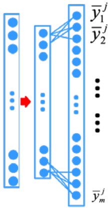


Fig. 13: Neural Networks

Table 2: The LAR when data perturbs due to different load variations with magnitudes δ_1 per unit (p.u.)

Load Variations (p.u.)	$\delta_1 = 1$	$\delta_1 = 1.5$	$\delta_1 = 2$	$\delta_1 = 3$
LAR (%)	96.25	81.61	71.96	57.5

- **Motivations**: Perturbations in power grids **degrade** the performance of NNs, shown in the Table 2.
- **Goal:** Train neural networks to be robust to *natural perturbations* in power grids.
- **The state of the art:**
  - Adversarial training methods can augment the robustness of NNs (Madry et al. 2017, Shafahi et al. 2020, Goodfellow and Begnio 2016).
  - Those perturbations are well-design due to some malicious attacks with **weak capability** to *natural perturbations*.

Our Approach [1]

- Main idea: Obtain the **worst-case perturbation** σ **constrained by physical laws** to train the parameters θ of neural networks without extra training datasets.
- The loss function is $L(\theta, \sigma)$:

[1] **Wenting Li**, Deepjyoti Deka, Ren Wang, Mario Arrieta Paternina,“Physics-Constrained Adversarial Training for Neural Networks in Stochastic Power Grids’, Artificial Intelligence on Transaction, 2022

Our Approach [1]

- Main idea: Obtain the **worst-case perturbation** σ **constrained by physical laws** to train the parameters θ of neural networks without extra training datasets.
- The loss function is $L(\theta, \sigma)$:

$$\min_{\theta} \mathcal{L}(\theta) = \underbrace{\mathcal{L}_{\text{original}}(\theta)}_{\text{Data Fitting}}$$

[1] **Wenting Li**, Deepjyoti Deka, Ren Wang, Mario Arrieta Paternina,"Physics-Constrained Adversarial Training for Neural Networks in Stochastic Power Grids", Artificial Intelligence on Transaction, 2022

- Main idea: Obtain the **worst-case perturbation** σ **constrained by physical laws** to train the parameters θ of neural networks without extra training datasets.
- The loss function is $L(\theta, \sigma)$:

$$\min_{\theta} \max_{\sigma} \mathcal{L}(\theta) = \underbrace{\mathcal{L}_{\text{original}}(\theta)}_{\text{Data Fitting}} + \underbrace{\mathcal{L}_{\text{worst \& orignial}}(\theta, \sigma)}_{\text{Robust Errors}}$$

[1] **Wenting Li**, Deepjyoti Deka, Ren Wang, Mario Arrieta Paternina,“Physics-Constrained Adversarial Training for Neural Networks in Stochastic Power Grids”, Artificial Intelligence on Transaction, 2022

Our Approach [1]

- Main idea: Obtain the **worst-case perturbation σ constrained by physical laws** to train the parameters θ of neural networks without extra training datasets.
- The loss function is $L(\theta, \sigma)$:

$$\min_{\theta} \underbrace{\max_{\sigma \in \Omega}}_{\text{Physical range}} \mathcal{L}_{\theta} = \underbrace{\mathcal{L}_{\text{original}}(\theta)}_{\text{Data Fitting}} + \underbrace{\mathcal{L}_{\text{worst \& orignial}}(\theta, \sigma)}_{\text{Robust Errors}}$$

$$+ \underbrace{\mathcal{L}_{\text{physical correclations of } \sigma}(\sigma)}_{\text{Physics Constraints}}$$

[1] **Wenting Li**, Deepjyoti Deka, Ren Wang, Mario Arrieta Paternina,“Physics-Constrained Adversarial Training for Neural Networks in Stochastic Power Grids”, Artificial Intelligence on Transaction, 2022

Datasets

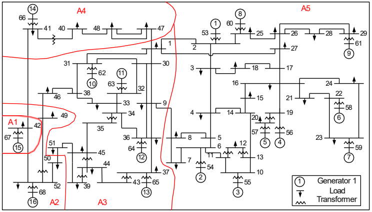


Fig. 14: The IEEE 68-bus benchmark Power Grid with five areas

- IEEE 68-bus system[a] through power system toolbox (PST)[b].
- 560 training data with 36 measured buses.
- Natural perturbations are generated when loads vary with δ_1 and control input changes with δ_2.

[a] Rogers 2012.

[b] Chow and Cheung 1992.

Robustness to Natural Perturbations

Table 3: Location Accuracy Rate (LAR %) when the load vary within δ_1

Type(Method)	δ_1= 1	δ_1=1.5	δ_1=2	δ_1=3
Base	96.25	81.61	71.96	57.5
TRADES	95.89	88.21	76.96	66.61
FGSM	90.89	84.46	75.36	62.68
PGD	90.36	80.18	71.25	58.04
$\text{TRADES}_{\text{regu}}$ (Proposed)	97.86	89.64	81.96	73.39
$\text{TRADES}_{\text{physics}}$ (Proposed)	**98.21**	**91.96**	**82.86**	**74.64**

- $\text{TRADES}_{\text{regu}}$ and $\text{TRADES}_{\text{physics}}$ **Improve** the LAR[3] than existing methods[4] by 1% - 17%.

[3] LAR = $\frac{\text{The number of correctly located events}}{\text{Total number of events}}$

[4] where "Base" denotes the stochastic gradient descent method; "$\text{TRADES}_{\text{regu}}$": When the physical regularization is applied. "$\text{TRADES}_{\text{physics}}$": when regularization and physical ranges are included; TRadeoff-inspired Adversarial DEfense(TRADES); Fast Sign Gradient Method (FGSM); Projected Gradient Descent (PGD).

Central Ideas:

Robustify neural networks for fault location through:

1. Designing novel **architectures** by preserving physics
2. Developing physics-constrained optimization for **training**
3. **Certifying** training with physics-informed bound propagation

The State of the Art

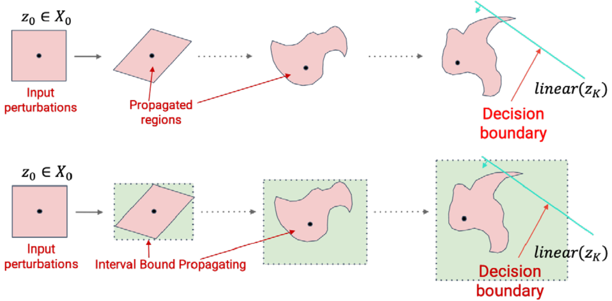


Fig. 15: ***Verification*** *is to check whether the region of NNs' outputs (red areas) satisfies some decision boundary and is* ***intractable****; Interval Bound Propagation (IBP) method is tractable (Gowal et al. 2018)*

- IBP can efficiently verify it by propagating the **interval bounds (green areas)** of the regions.

Our Approach[5]

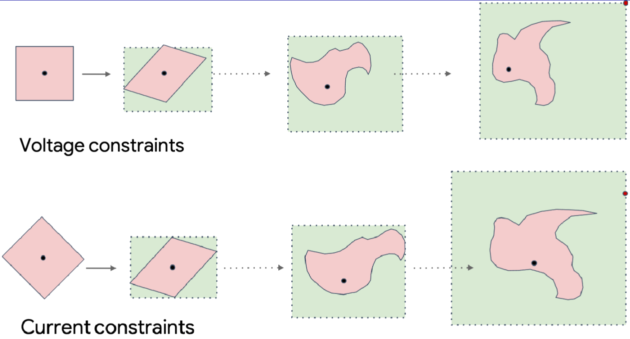


Fig. 16: The voltage limits propagate; the current limits propagate

- **Tighter bounds of the region with physical ranges**, including the voltage limits, current limits and their correlation.

[5] W. Li, K. Dvijotham, D. Deka, "Physics-Constrained Interval Bound Propagation for Robustness Verifiable Neural Networks in Power Grids", AI for Energy Innovation, 2023

Numerical Results

Table 4: Performance comparison when loads randomly vary within δ_1

Type (Method)	$\delta_1 = 1$	δ_1=1.5	δ_1=2	δ_1=3
Base (LAR %)	96.25	81.61	71.96	57.5
IBP (LAR %)	98.93	94.64	84.29	61.43
IBP_{Phy} (LAR %)	**99.64**	**96.61**	**89.29**	**64.29**
IBP (VR %)	93.93	82.14	69.64	47.5
IBP_{Phy} (VR %)	**97.32**	**90.36**	**79.46**	**54.29**

- “Base” denotes that no verification training; “IBP” denotes the baseline and “ IBP_{Phy}” denotes the proposed method.
- We improve the LAR and VR [6] up to 5% and 10% respectively.

[6]

$$\text{LAR} = \frac{\text{The number of correctly located cases}}{\text{Total number of cases}}$$

$$\text{VR} = \frac{\text{The number cases satisfying the specification}}{\text{Total number of cases}}$$

Conclusions and Future Works

- The stochastic and dynamical environments in power grids require machine learning algorithms to be **robust and interpretable**;
- Design a **physics-preserved graph network framework** for fault location, showing superior performance than the state of the art when data is imperfect (low label rates and noisy);
- Propose **a training algorithm with physical constraints** to enhance robustness of neural networks in the perturbed environment;
- Create **a physics-Informed verification** for neural networks in power grids to guarantee the reliability.
- Future works will **generalize the art of guiding machine learning with physics** for extensive applications, such as stability prediction, state estimation.

Thank you! Questions?

Job Search for modelers of microgrids and distribution systems

- Postdoc: https://lanl.jobs/search/jobdetails/microgrid-postdoctoral-research-associate/f47f3ac6-7caf-4213-a325-f1ee3b4489e2
- Scientist: https://lanl.jobs/search/jobdetails/microgrid-scientist-scientist-2/b2696055-3997-4a53-a246-147bc4e9f2fe

Email: wenting@lanl.gov

Los Alamos NATIONAL LABORATORY EST.1943

Center for Nonlinear Studies