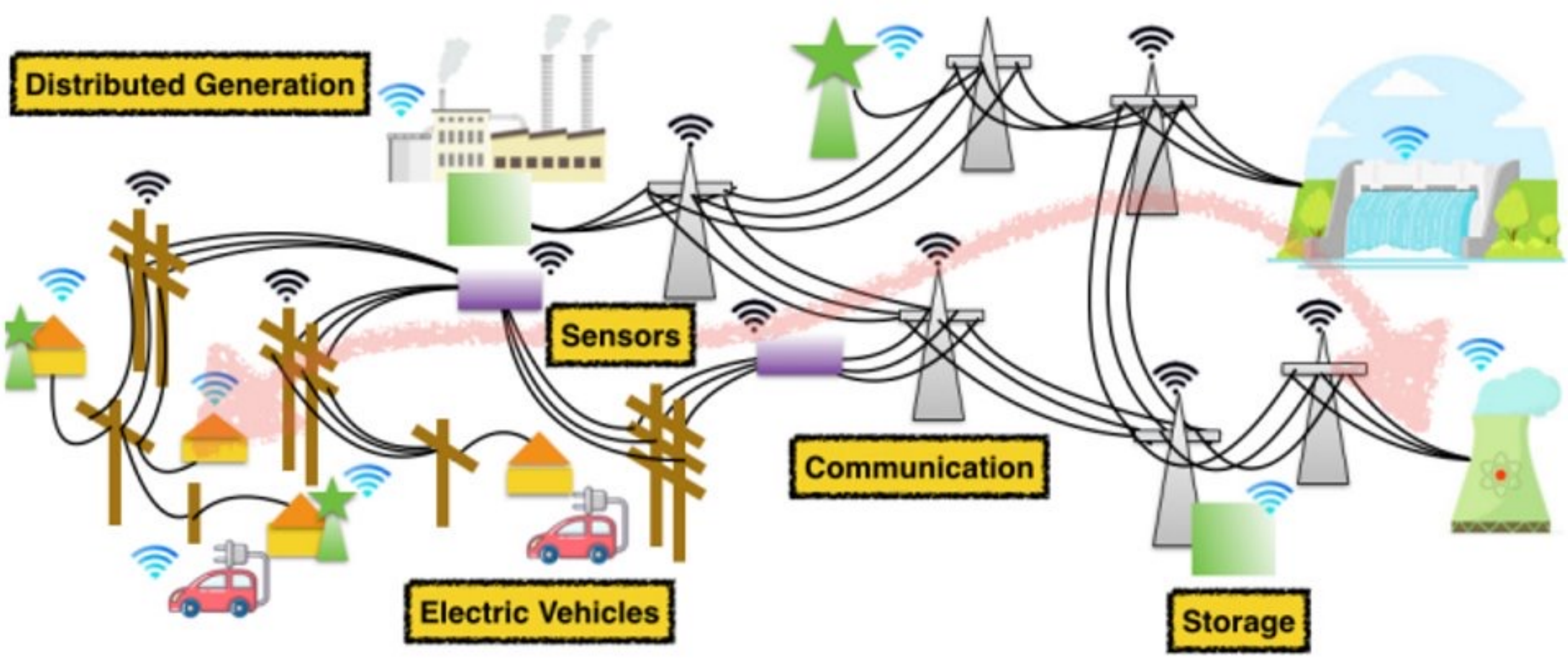
RECONSTRUCTION OF GRID MEASUREMENTS IN THE PRESENCE OF ADVERSARIAL ATTACKS

Amirmohammad Naeini, Samer El Kababji, Pirathayini Srikantha
York University
Nov 7, 2022

YORK U

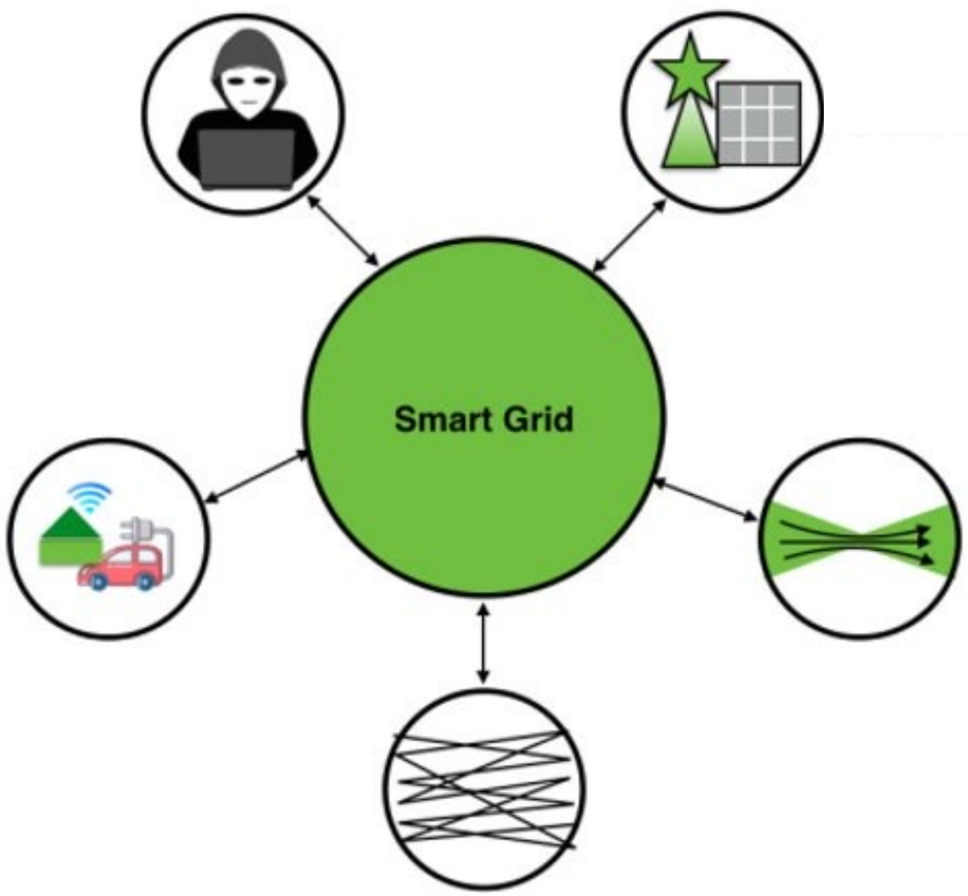
Introduction

- **Climate change:** Proliferation of highly variable renewables
- **Cyber-physical:** Vulnerabilities in the cyber plane
- **Stable operations:** Real-time monitoring resilient to cyber attacks



Motivation

- **Impact of cyber attacks on grid operation**:
  - Renewables: Introduce significant flux in the grid
  - Attacks: Increase modes of instability in the grid
- **Cyber attacks examples**:
  - 2011, Stuxnet worm Iran nuclear plant [1]
  - 2015, Ukraine blackout [2]
  - 2019, Venezuela power grid attack [3]



YORK U

Problem Statement

- **State estimation**:
  - Infer grid states (x) given a set of grid measurements (y)

$$y + \epsilon = H(x); \;\; x = H^{-1}(y + \epsilon)$$

  - Cycle GAN is used for **approximating** H and H^{-1}
- **False data injection**: Common attack in state estimation
  - Cannot detect perturbations to measurements using traditional residual checking
  - Lead to incorrect state inferences

$$x = H^{-1}(y + a + \epsilon)$$

  - Iterative gradients computed using Cycle GAN modules used for **reconstructing** perturbed measurements

YORK U

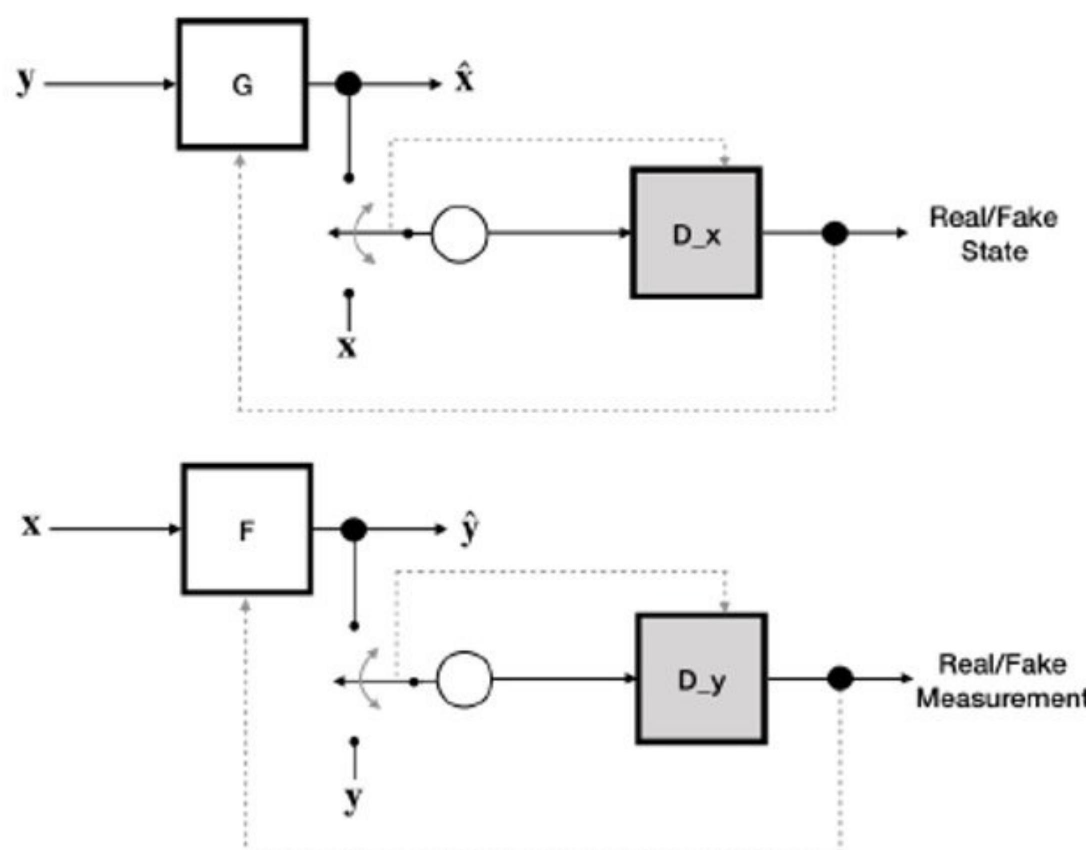
Existing Work

	No knowledge of grid structure	Unsupervised Training Dataset	Recovering from higher rates of perturbation
Proposed Method	✓	✓	✓
GAN Based method [4]	x	x	?
Numerical Method [5]	✓	N/A	?
Machine Learning [6]	x	x	✓

YORK U

Cycle GAN

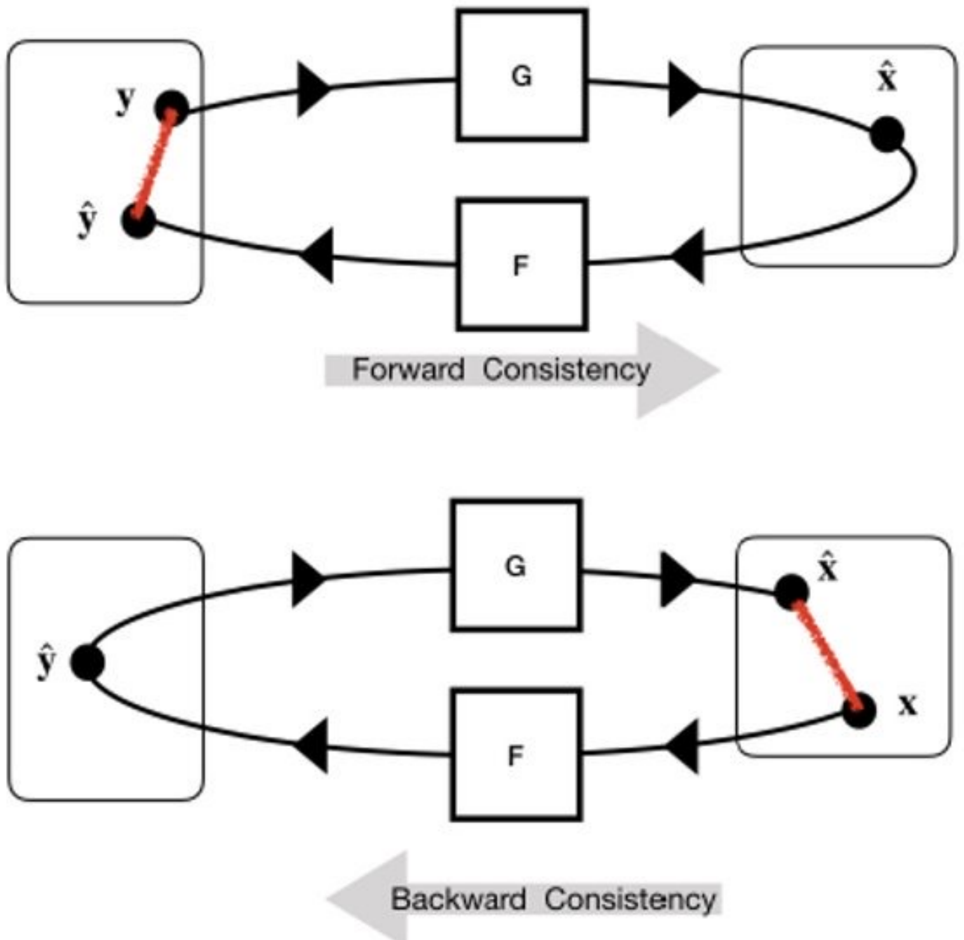
- Composed of two sets of GANs
  - Forward GAN: G is mapping from measurements to states ($G \approx H^{-1}$)
  - Reverse GAN: F is mapping from states to measurements ($F \approx H$)



YORK

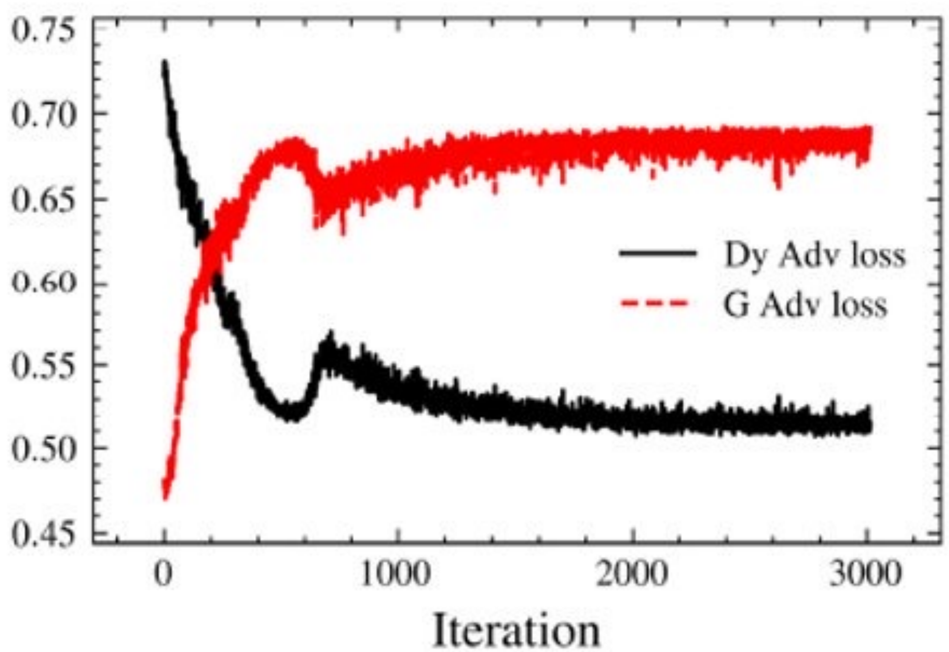
Cycle GAN

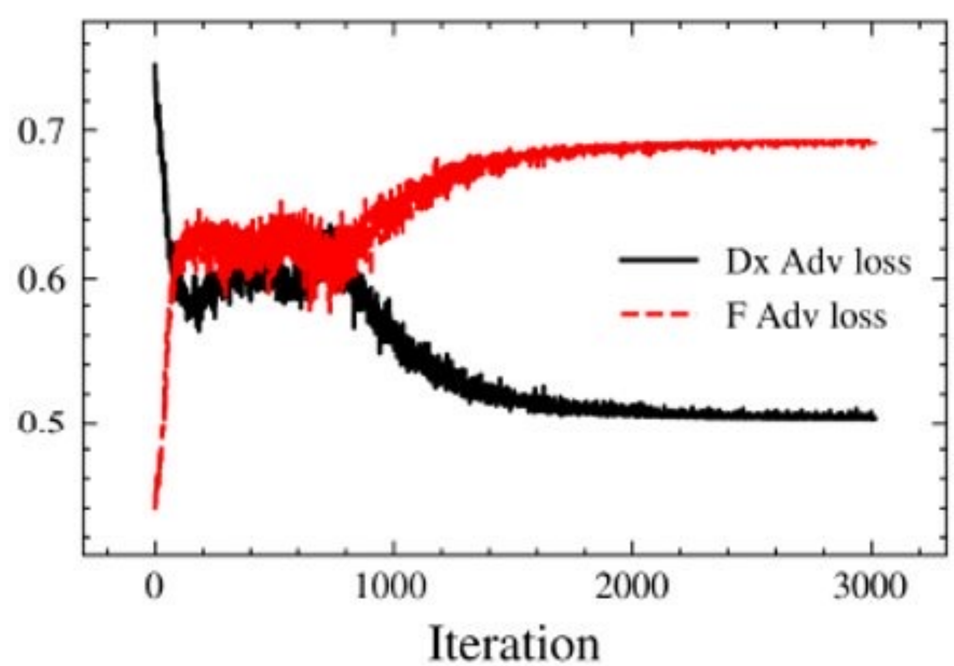
- Mapping between domains: Cycle consistency loss



YORK U

Cycle GAN: Framework





Grid State Generator Neural Network - G	
Nodes	Input: 759 L_1:512, L_2:1024, L_3:2048, L_4:1024, L_5:512 Output: 235
Activation	relu, relu, relu, relu, relu, tanh
Grid State Discriminator Neural Network- D_x	
Nodes	Input: 235 L_1:512, L_2:1024, L_3:256, L_4:64 Output: 1
Activation	relu, relu, relu, relu, sigmoid
Grid Measurement Generator Neural Network - F	
Nodes	Input: 235 L_1:512, L_2:1024, L_3:2048, L_4:1024, L_5:512 Output: 759
Activation	relu, relu, relu, relu, relu, tanh
Grid Measurement Discriminator Neural Network- D_y	
Nodes	Input: 759 L_1:512, L_2:1024, L_3:256, L_4:64 Output: 1
Activation	relu, relu, relu, relu, sigmoid

YORK U

Proposed Algorithm

- **Detection**:
  - Residual-based: If $|y_i - F(G(y))_i| \geq \alpha$,i^{th} component is labelled as attacked
- **Reconstruction**:

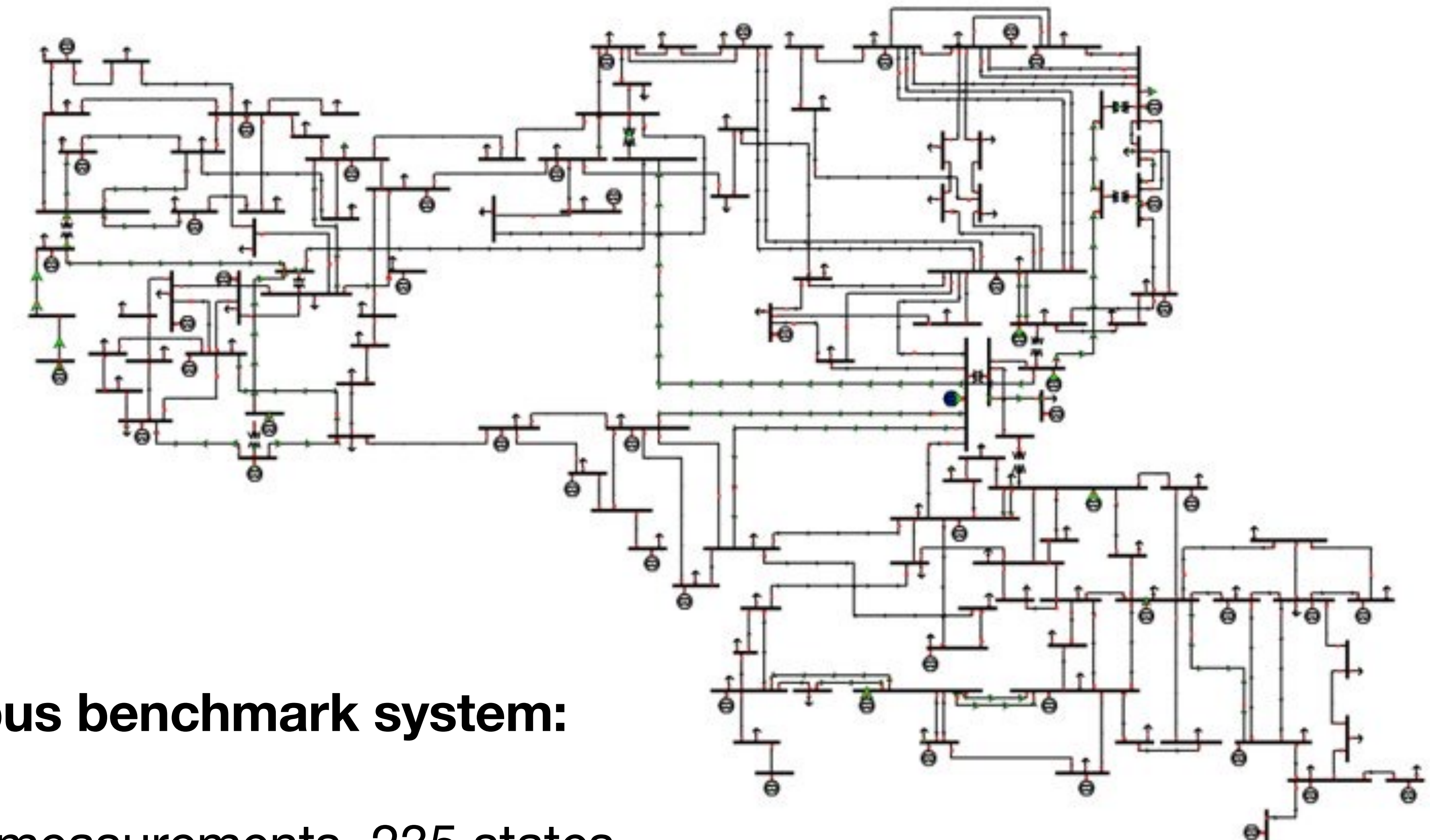
  Problem formulation: $\mathcal{P}_{err}: \min_{y} ||y - F(G(y))||_2^2$

  Gradient computation: $\frac{\partial f}{\partial y} = -2 \cdot (y - F(G(y))(1 - \frac{\partial F(G(y))}{\partial G(y)} \frac{\partial G(y)}{\partial y})$

  Iterative update rule: $y_{t+1} = y_t - 2\beta\, sgn(y_t - m_t)(y_t - m_t) \frac{\partial f(y_t)}{\partial y}$

YORK U

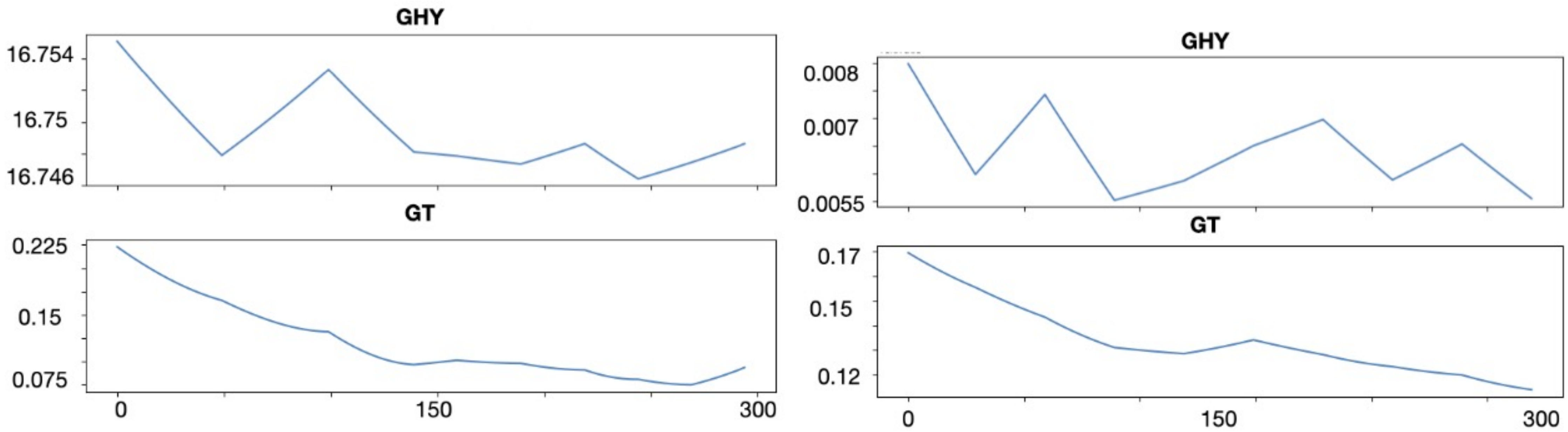
Results



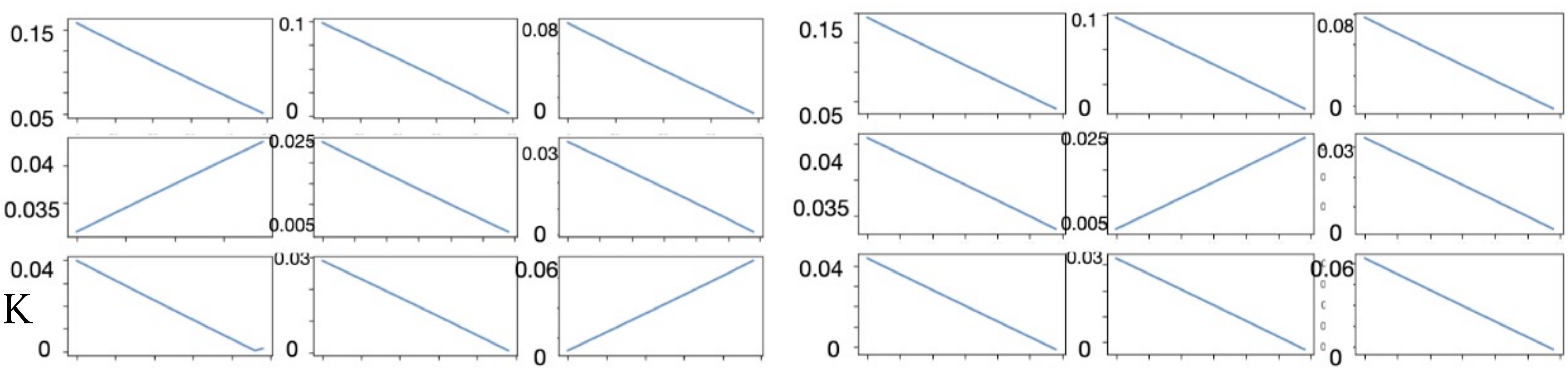
- **IEEE 118-bus benchmark system:**
  - 759 grid measurements, 235 states
- **Attack simulation:**
  - Randomly selected columns , ±10% perturbation

YORK U

Results



G(H(y)) and GT error for two different cases.



Component-wise ground truth error.

Conclusion

- **Novel reconstruction method:**
  - No need for underlying knowledge of grid topology and parameters
  - Inferencing is computationally inexpensive after training Cycle GAN
  - Effective for high rates of perturbation
- **Future work:**
  - Iterative revisions reach 0 most of the time but is not stopped at these points
  - Need to identify an effective stopping criteria

YORK U

References

1. Anwar, A., Mahmood, A. N., & Pickering, M. (2017). Modeling and performance evaluation of stealthy false data injection attacks on smart grid in the presence of corrupted measurements. *Journal of Computer and System Sciences*, *83*(1), 58-72.

2. G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, “The 2015 ukraine blackout: Implica157 tions for false data injection attacks,” IEEE Trans. Power Syst., vol. 32, no. 4, pp. 3317–3318, 158 Jul. 2017.

3. Vaz, Ricardo. “Venezuela’s power grid disabled by cyber attack,” Green Left Weekly, no. 1213, 160 pp. 15, March 2019.

4. Karimipour, H., Dehghantanha, A., Parizi, R. M., Choo, K. K. R., & Leung, H. (2019). A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids. IEEE Access, 7, 80778-80788.

5. Ruan, J., Liang, G., Zhao, J., Qiu, J., & Dong, Z. Y. (2022). An Inertia-based Data Recovery Scheme for False Data Injection Attack. IEEE Transactions on Industrial Informatics.

6. Li, Y., Wang, Y., & Hu, S. (2019). Online generative adversary network based measurement recovery in false data injection attacks: A cyber-physical approach. IEEE Transactions on Industrial Informatics, 16(3), 2031-2043.