Machine Learning Solutions for Monitoring U.S. Transmission Grid with Large-scale Real-world PMU Data

Prof. Nanpeng Yu
Department of Electrical and Computer Engineering
Department of Computer Science
(cooperating faculty)
Director of Energy, Economics, and Environment Research Center
nyu@ece.ucr.edu
951.827.3688

UNIVERSITY OF CALIFORNIA, RIVERSIDE

Outline

- Project Background and Objectives
- Technical Accomplishments
  - Power System Event Detection
  - Power System Event Classification
  - Synthetic Power System Event Data Creation
- Lessons Learned and Next Steps

Project Background and Objectives

- Department of Energy's Transmission Research Program
  - FOA 1861 – Big Data Analysis of Synchrophasor Data (Oct. 2019 – Mar. 2022)
- Project Objectives
  - Derive value from the vast amounts of Phasor Measurement Unit (PMU) Data
  - Provide actionable information on the use of Machine Learning and Artificial Intelligence methods on large PMU datasets
  - Enable faster grid analytics and modeling
- First-of-its-kind PMU dataset
  - Covers each of three U.S. interconnections (~450 PMU, 30 & 60 Hz reporting rate)
  - Covers 2 years including event logs (27 TB)
  - Is real data with inconsistencies, varying quality levels, and flaws (66% - 70% good data)
  - Is anonymized to protect the data providers (lack of location and topology information)

Technical Accomplishments

- PMU Data Quality Improvement
  - Online PMU Missing Value Replacement via Event-Participation Decomposition
- **Power System Event Detection**
  - Graph Signal Processing-based Event Detection
  - **Voltage Event Detection Using Optimization with Structured Sparsity-Inducing Norms**
  - **Power System Event Detection with Bidirectional Anomaly Generative Adversarial Networks**
- **Power System Event Classification**
  - **Deep Neural Network-based Power System Event Classification**
  - Classify System Events with a Small Number of Training Labels with Transfer Learning
  - Adversarial Attacks on Deep Neural Network-based Power System Event Classification Models
- Power System Dynamic Parameter Estimation
  - Dynamic Parameter Estimation with Physics-based Neural Ordinary Differential Equations
- **Synthetic Power System Event Data Creation**
  - **pmuBAGE: The Benchmarking Assortment of Generated PMU Events**
- Power System Event Signature Library
  - A Dynamic Behavior-based Power System Event Signature Library

Voltage Event Detection Using Optimization with Structured Sparsity-Inducing Norms

- Background: *Is PMU Data Matrix Low Rank?*
  - Low-rank property of PMU data matrix holds up during normal operations
    - Largest singular value of Q data matrix accounts for 99.988% of the variance
  - The low-rank property of PMU data matrix is no longer valid during voltage events
    - The largest singular value of Q data matrix accounts for only 59.743% of the variance

TABLE I
SINGULAR VALUE DECOMPOSITION OF P, Q, V, AND F DATA MATRICES OVER 1 SECOND (30 SAMPLES)

Data Type	Electrical Quantity	Singular Value Percentage Variance ($\frac{\sigma_i^2}{\sum_i \sigma_i^2}$)			Singular Value Proportion ($\frac{\sigma_i}{\sum_i \sigma_i}$)		
		1st	2nd	3rd	1st	2nd	3rd
Non-event Data	P (Active Power)	99.999261%	0.000536%	0.000109%	99.242040%	0.229736%	0.103609%
	Q (Reactive Power)	99.988472%	0.008789%	0.001427%	97.134683%	0.910695%	0.366895%
	V (Voltage Magnitude)	99.999995%	0.000005%	0.000000%	99.963393%	0.022302%	0.003274%
	F (Frequency)	99.999999%	0.000000%	0.000000%	99.996304%	0.000824%	0.000654%
Event Data	P (Active Power)	95.003242%	4.933310%	0.045058%	78.391273%	17.863547%	1.707207%
	Q (Reactive Power)	$\boxed{59.743182\%}$	40.185730%	0.068034%	53.278058%	43.695845%	1.797904%
	V (Voltage Magnitude)	99.545736%	0.447828%	0.006371%	92.892350%	6.230519%	0.743139%
	F (Frequency)	99.999994%	0.000006%	0.000000%	99.971389%	0.023498%	0.001269%

Row-Sparse Structure of Residual PMU Data Matrix*

- Key Observations
  - Voltage related events trigged by system faults are often regional events
  - The $X - L$ during voltage event periods have row-sparse structure
  - Rows of residual matrix correspond to PMUs highly impacted by the event
- Main Idea
  - Decompose the streaming PMU data matrix X into
    - A low-rank matrix L, a row-sparse event-pattern matrix S, and a noise matrix G
  - Extract anomaly features from L & S
  - Use clustering algorithm to identify power system voltage events

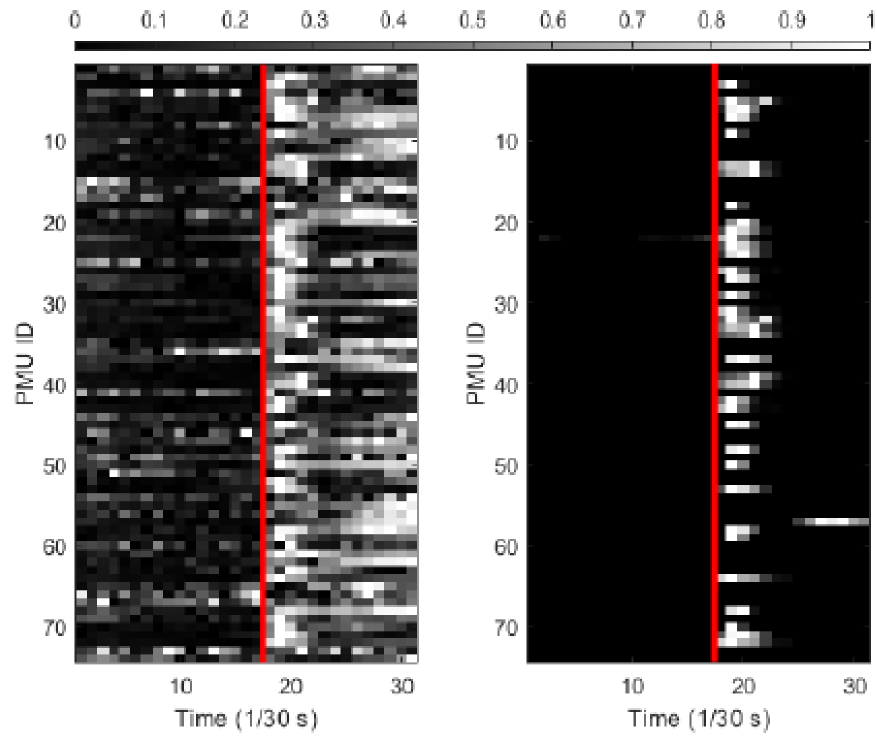
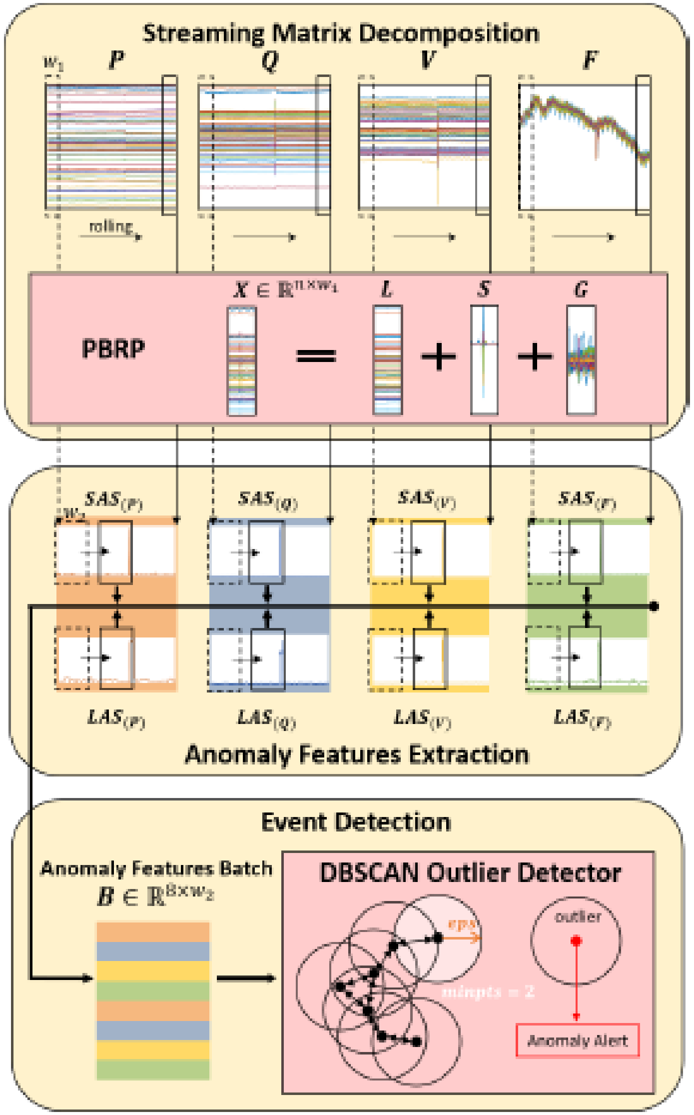


Fig. 2. The heatmap of "$X - L$" (left) and "$X - L - G$" (right) for normalized active power data (scaled from 0 to 1). The event happens approximately at the red line.

* X. Kong, B, Foggo, and N. Yu, "Online Voltage Event Detection Using Optimization with Structured Sparsity-Inducing Norms," *IEEE Transactions on Power Systems*, 2022. DOI: 10.1109/TPWRS.2021.3134945.

Overview of Voltage Event Detection Framework



Step 1: Decompose streaming PMU data matrix $X = L + S + G$

✓ Proposed Algorithm: *Proximal Bilateral Random Projections (PBRP)*

Step 2: Extract anomaly features

✓ l_{21} norm of the row sparse matrix S

✓ Max temporal difference of low-rank matrix L

Step 3: Distinguish normal system operation data from that of the system voltage events

✓ Adopt density-based cluster analysis DBSCAN

✓ Outliers correspond to voltage events

Decompose Matrix with Row-Sparse Structure with Proximal Bilateral Random Projection (PBRP)

$$\min_{L,S} \quad \frac{1}{2}\|X-L-S\|_F^2 \quad s.t. \quad \begin{cases} rank(L) = r, \\ S \text{ is row-sparse}. \end{cases} \Rightarrow \min_{L,S} \quad \frac{1}{2}\|X-L-S\|_F^2 + \lambda\|S\|_{21} \quad s.t. \quad rank(L) = r,$$

- Solution Approach: Coordinate Descent

$$\begin{cases} L^{(k)} = \arg \min_{rank(L)=r} \frac{1}{2}\|X-L-S^{(k-1)}\|_F^2 \\ S^{(k)} = \arg \min_{S} \frac{1}{2}\|X-L^{(k)}-S\|_F^2 + \lambda\|S\|_{21} \end{cases}$$

- Update L: Closed-form Bilateral Random Projection (BRP) enhanced by Power Scheme
- Update S with proximal method

$$Prox_{\lambda\|\cdot\|_{21}}(X[i,:]) = (1-\lambda\|X[i,:]\|_2)_+ X[i,:],$$

Algorithm 2 Proximal BRP (PBRP)

Input: $X \in \mathbb{R}^{n\times w_1}$, rank r, power factor Q, λ, ϵ
Output: L, S
1: Initialization: $L = S = 0$
2: **while** $\frac{\|X-L-S\|_F^2}{\|X\|_F^2} \geq \epsilon$ **do**
3: $L = BRP(X-S)$
4: $S = Prox_{\lambda\|\cdot\|_{21}}(X-L)$;
5: **end while**
6: **return** L, S

Numerical Results and Summary

F SCORES OF THREE ALGORITHMS ON THE TESTING DATASET

Statistics	OLAP	HOLAP	P-BRP
Precision	**0.8889**	0.8824	0.8881
Recall	0.8955	0.8955	**0.9478**
F1 Score	0.8922	0.8889	**0.9170**
Precision	0.8089	**0.8571**	0.8000
Recall	0.9478	0.9403	**0.9851**
F2 Score	0.9163	0.9224	**0.9415**

AVERAGE COMPUTATION TIME OF EVENT DETECTION ALGORITHMS OVER THREE-MINUTE TIME PERIOD

Number of PMUs		50	100	150
Computation Time (s) (partial/total)	HOLAP	61.78/68.46	181.50/189.25	336.27/344.58
	OLAP	7.53/15.01	9.58/17.33	16.99/24.79
	P-BRP	2.18/8.46	3.13/9.40	4.29/10.53

- Residual PMU data matrices during voltage events have distinctive sparsity structure
- Computationally efficient PBRP algorithm is proposed to decompose PMU data matrices
- The proposed online voltage event detection algorithm shows better accuracy and scalability

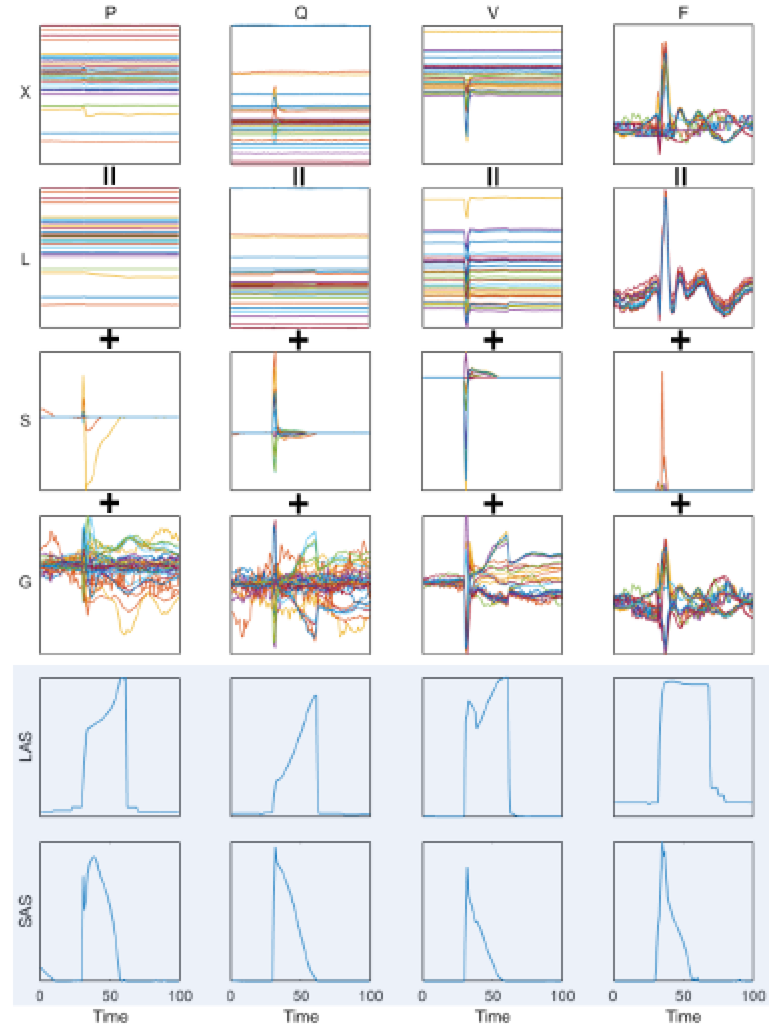
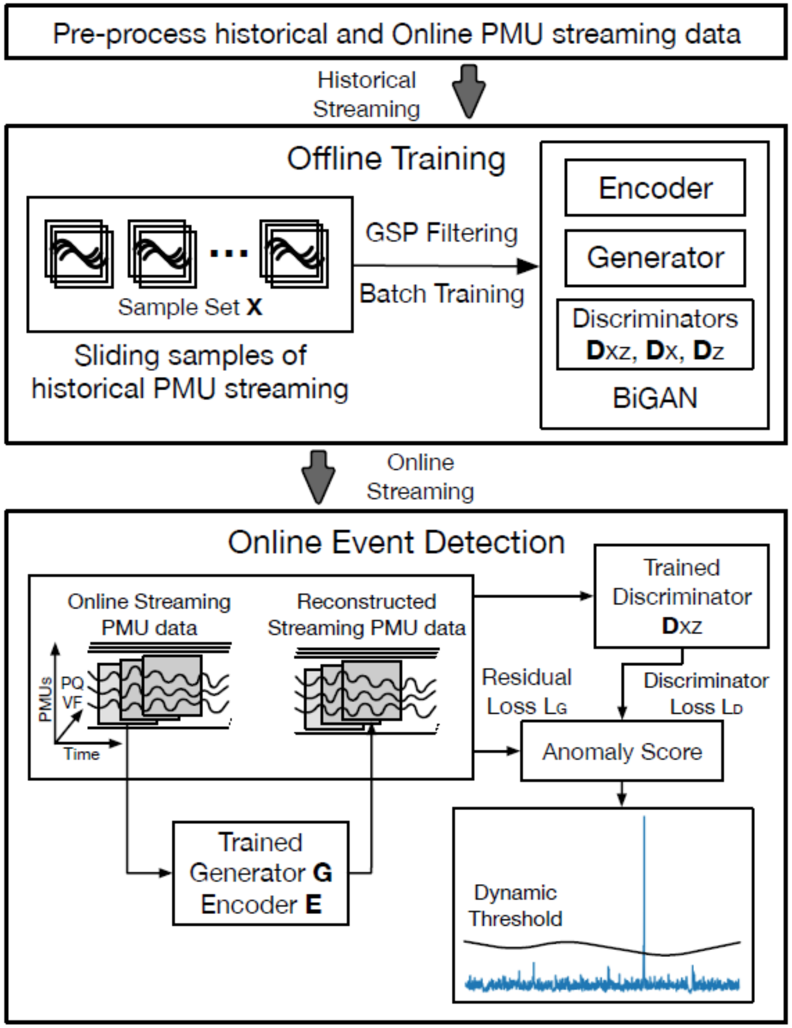


Fig. 5. An example of decomposition of streaming PMU data matrix X with corresponding anomaly scores for a voltage event.

Power System Event Detection via Bidirectional Anomaly Generative Adversarial Networks

- Motivation
  - Detecting power system events with supervised machine learning algorithms requires a large amount of high quality training labels (confirmed events)
  - Event detection accuracy drops quickly as the number of training label reduces.
  - Develop a Bidirectional Anomaly Generative Adversarial Network (Bi-AnoGAN)-based event detection algorithm, which does not depend on a large amount of high quality event labels.
- Main Idea
  - Learn two mapping functions that project PMU data samples during normal operating conditions to the noise space and then back to the PMU data space.
  - A large reconstruction error and discriminator loss → it is very likely that the new PMU sample corresponds to a system event.
  - Improve computation efficiency with the design of Bidirectional GAN (BiGAN) by training an additional encoder network that can directly map a PMU data sample to the noise space.

Overview of System Event Detection Framework



Step 1: Pre-process historical and online streaming PMU data

Step 2: Offline training. Train an encoder E, generator G, and discriminators D using PMU data during normal operating conditions.

Step 3: Online event detection.

(1) Calculate the difference between original streaming PMU data and the reconstructed PMU data

(2) Calculate discriminator loss (Does incoming PMU sample come from normal operation periods?)

(3) Calculate anomaly score and compare it against a dynamic threshold

Offline Training of Bi-AnoGAN

- Training of Bi-GAN is formulated as a min-max problem

$$\min_{\boldsymbol{G},\boldsymbol{E}} \max_{\boldsymbol{D}} V(\boldsymbol{D},\boldsymbol{E},\boldsymbol{G}) = \mathbb{E}_{\boldsymbol{x}\sim p(\boldsymbol{x})}[log\boldsymbol{D}(\boldsymbol{x},\boldsymbol{E}(\boldsymbol{x}))] + \mathbb{E}_{\boldsymbol{z}\sim p(\boldsymbol{z})}[\log(1-\boldsymbol{D}(\boldsymbol{G}(\boldsymbol{z}),\boldsymbol{z}))]$$

  - Encoder, E improves computational efficiency by directly mapping PMU data samples to the noise space

- Training of BiGAN with Wasserstein loss

$$\min_{\boldsymbol{G},\boldsymbol{E}} \max_{\boldsymbol{D}\in\mathcal{D}} V_{\boldsymbol{xz}}(\boldsymbol{D}_{\boldsymbol{xz}},\boldsymbol{E},\boldsymbol{G}) = \mathbb{E}_{\boldsymbol{x}\sim p(\boldsymbol{x})}[\boldsymbol{D}_{\boldsymbol{xz}}(\boldsymbol{x},\boldsymbol{E}(\boldsymbol{x}))] - \mathbb{E}_{\boldsymbol{z}\sim p(\boldsymbol{z})}[\boldsymbol{D}_{\boldsymbol{xz}}(\boldsymbol{G}(\boldsymbol{z}),\boldsymbol{z})]$$

  - The 1-Lipschitz constraint on the discriminator function mitigates mode collapse problem and improves convergence of the training process

- Encourage cycle consistency by adding conditional entropy constraints

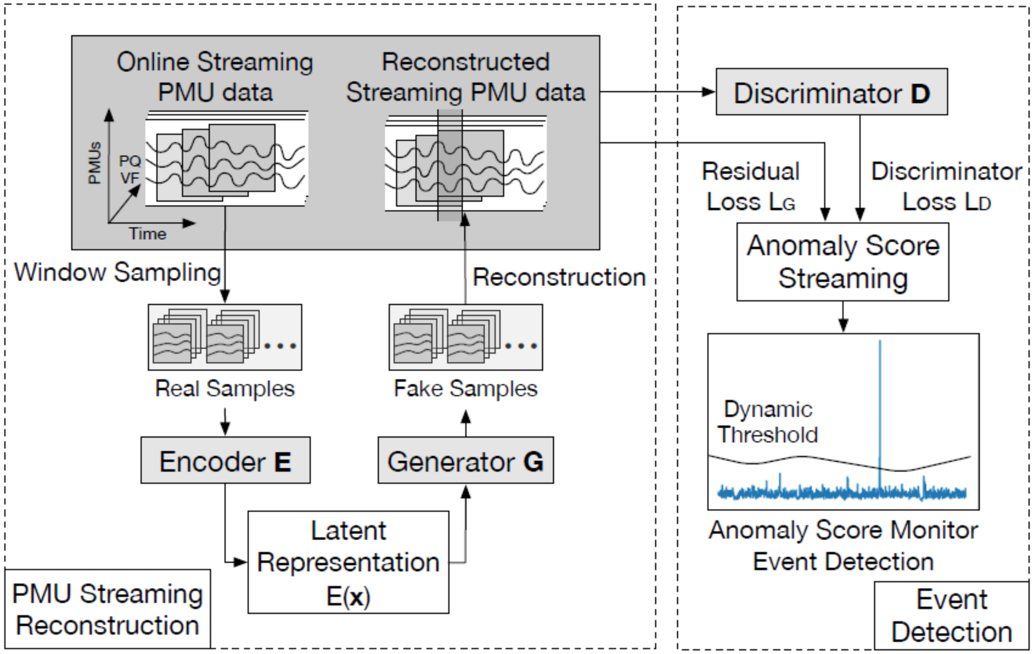
Add $V_x(\boldsymbol{D}_x,\boldsymbol{E},\boldsymbol{G}) = \mathbb{E}_{\boldsymbol{x}\sim p(\boldsymbol{x})}[\boldsymbol{D}_{\boldsymbol{x}}(\boldsymbol{x})] - \mathbb{E}_{\boldsymbol{x}\sim p(\boldsymbol{x})}\left[\boldsymbol{D}_{\boldsymbol{x}}\left(\boldsymbol{G}\big(\boldsymbol{E}(\boldsymbol{x})\big)\right)\right]$ to enforce $\boldsymbol{x} = \boldsymbol{G}(\boldsymbol{E}(\boldsymbol{x}))$

Add $V_z(\boldsymbol{D}_z,\boldsymbol{E},\boldsymbol{G}) = \mathbb{E}_{\boldsymbol{z}\sim p(\boldsymbol{z})}[\boldsymbol{D}_{\boldsymbol{z}}(\boldsymbol{z})] - \mathbb{E}_{z\sim p(\boldsymbol{z})}\left[\boldsymbol{D}_{\boldsymbol{z}}\left(\boldsymbol{E}\big(\boldsymbol{G}(\boldsymbol{z})\big)\right)\right]$ to enforce $\boldsymbol{z} = \boldsymbol{G}(\boldsymbol{E}(\boldsymbol{z}))$

- Final objective function

$$\min_{\boldsymbol{G},\boldsymbol{E}} \max_{\boldsymbol{D}_{\boldsymbol{xz}},\boldsymbol{D}_x,\boldsymbol{D}_z} [V_{\boldsymbol{xz}}(\boldsymbol{D}_{\boldsymbol{xz}},\boldsymbol{E},\boldsymbol{G}) + V_x(\boldsymbol{D}_x,\boldsymbol{E},\boldsymbol{G}) + V_z(\boldsymbol{D}_z,\boldsymbol{E},\boldsymbol{G})]$$

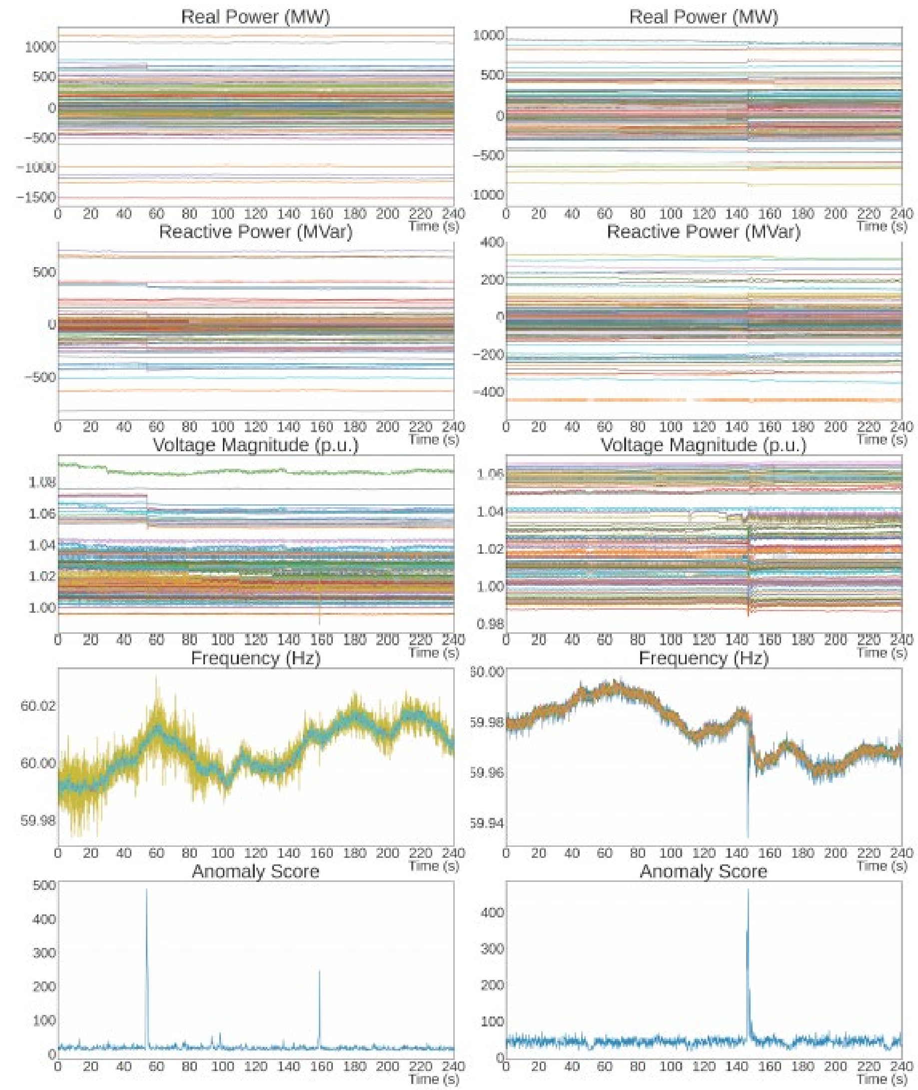
Online Event Detection



- Anomaly Score Calculation $\boldsymbol{L} = \lambda \boldsymbol{L}_G + (1-\lambda)\boldsymbol{L}_D$
  - PMU data reconstruction error $\boldsymbol{L}_G = \|\boldsymbol{x} - \boldsymbol{G}(\boldsymbol{E}(\boldsymbol{x}))\|_2$
  - The discriminator loss $\boldsymbol{L}_D = BCE(\boldsymbol{D}_{\boldsymbol{xz}}(\boldsymbol{x}, \boldsymbol{E}(\boldsymbol{x})))$. BCE: binary cross-entropy loss function.
- Dynamic Threshold for Anomaly Score
  - Threshold = $mean(L_{t-60:t-1}) + c \times std(L_{t-60:t-1})$
  - c is a hyper-parameter

Numerical Study Setup and Illustration

- PMU Dataset
  - 187 PMUs from Eastern Interconnection
  - May 2016 – December 2017
  - 807 voltage events, 82 frequency events
- Training Dataset for Bi-AnoGAN
  - First operation day's PMU data in a half year
- Size of training sample
  - A window size of 1 second
  - 3D tensor: 30 time stamps, 179 PMUs, 4 channels
  - Number of training samples in a day 86400
- Training Setup
  - Learning rate: 1e-4
  - Batch size 256
  - 8 hrs of training time on NVIDIA GeForce RTX 2080 Ti GPU



voltage event frequency event

Results and Summary

TABLE IV
ACCURACY OF DETECTION FOR VOLTAGE-RELATED EVENTS

	Bi-AnoGAN	OLAP	GSP-based	AnoGAN
True Positive	**584**	534	561	512
False Positive	**42**	67	138	77
False Negative	**23**	73	46	95
Precision	**93.29%**	88.89%	80.26%	86.92%
Recall	**96.21%**	89.55%	92.42%	84.34%
F_1 Score	**94.73%**	89.22%	85.91%	85.61%

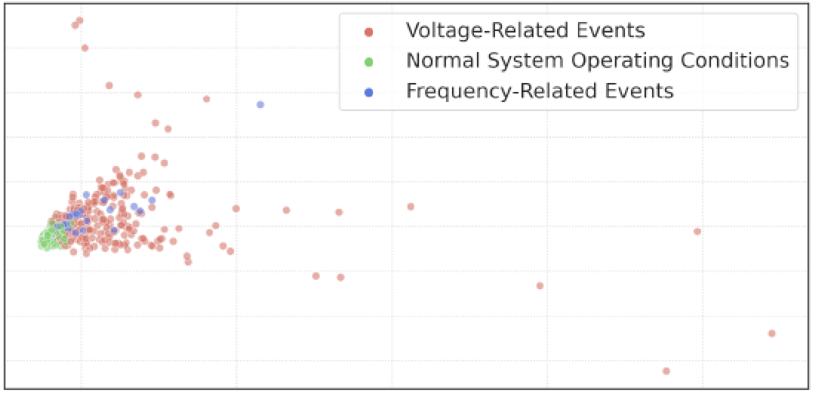


Fig. 7. Noise space representations of voltage-related events, frequency-related events, and normal system operation conditions.

TABLE V
ACCURACY OF DETECTION FOR FREQUENCY-RELATED EVENTS

	Bi-AnoGAN	OLAP	GSP-based	AnoGAN
True Positive	**82**	72	71	75
False Positive	**5**	56	19	46
False Negative	**0**	10	11	7
Precision	**94.25%**	53.33%	78.89%	61.98%
Recall	**100%**	88.89%	86.59%	91.46%
F_1 Score	**97.04%**	66.67%	82.56%	73.89%

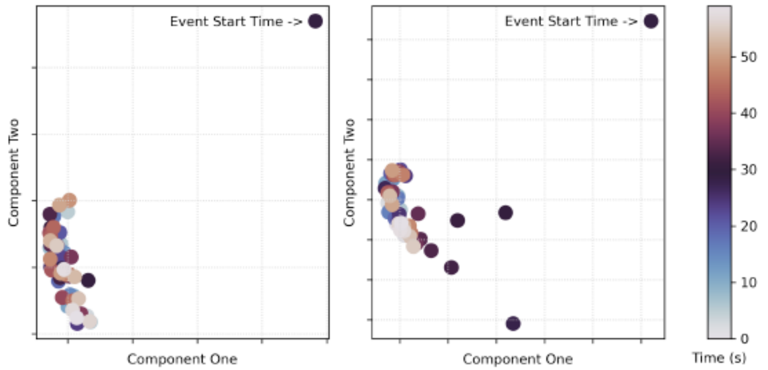


Fig. 8. Noise space representations of 1-minute samples surrounding the event. The left subfigure is a voltage-related event and the right subfigure is a frequency-related event.

TABLE VII
AVERAGE RUNTIME OF DIFFERENT ALGORITHMS FOR EVENT DETECTION

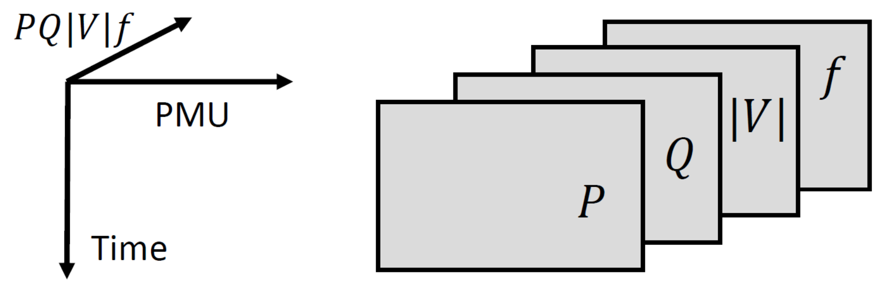
	Bi-AnoGAN	OLAP	GSP-based	AnoGAN
Voltage Events	13.59 s	21.75 s	7.16 s	876.58 s
Frequency Events	13.47 s	20.98 s	7.31 s	843.89 s

<2 ms for processing each snapshot of PMU data sample

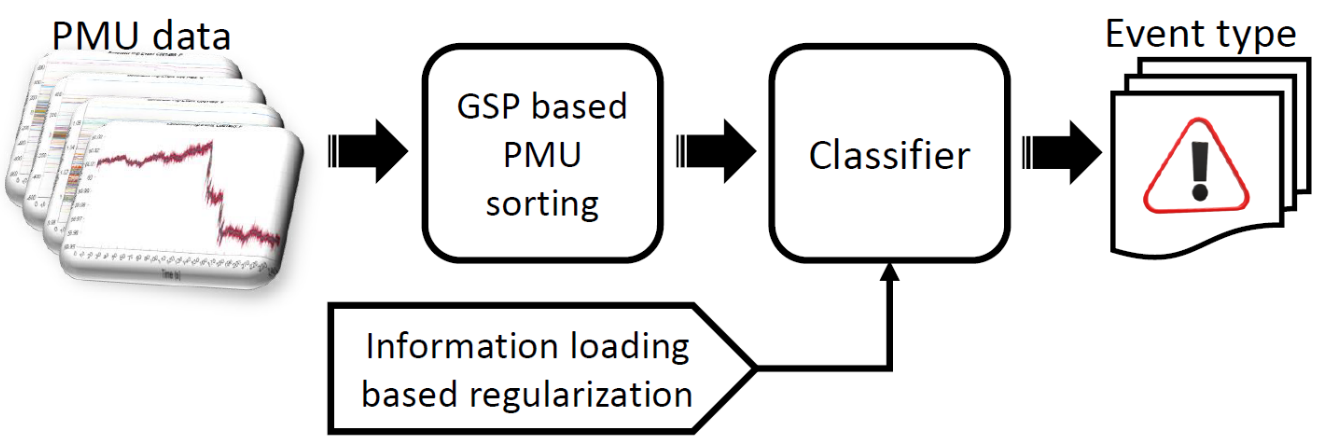
- Pros of Bi-AnoGAN: Computationally efficiency, do not need labels, high detection accuracy.
- Cons of Bi-AnoGAN: Network architecture needs to be appropriately designed to avoid non-convergence and instability.

System Event Identification: Overall Framework

- Formulated as a classification problem
  - Normal operation condition, line event, generator event, oscillation event
- Input: 3 dimensional tensor
  - Time, PMU ID, and _PQ|V|f measurement_



- Overall Framework



- Three key modules
  - CNN-based Classifier, GSP-based PMU Sorting, Info. Loading-based Regularization

Graph Signal Processing-based PMU Sorting

- Motivation
  - Make parameter sharing scheme of Convolutional Neural Classifier more effective
- Main Idea
  - Strategically place highly correlated PMUs close to each other
- Solution
  - Systematically rearrange PMUs in the input tenors with GSP-based PMU sorting algo.

$$\min_{\boldsymbol{d}} \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij} \left(d_i - d_j\right)^2$$

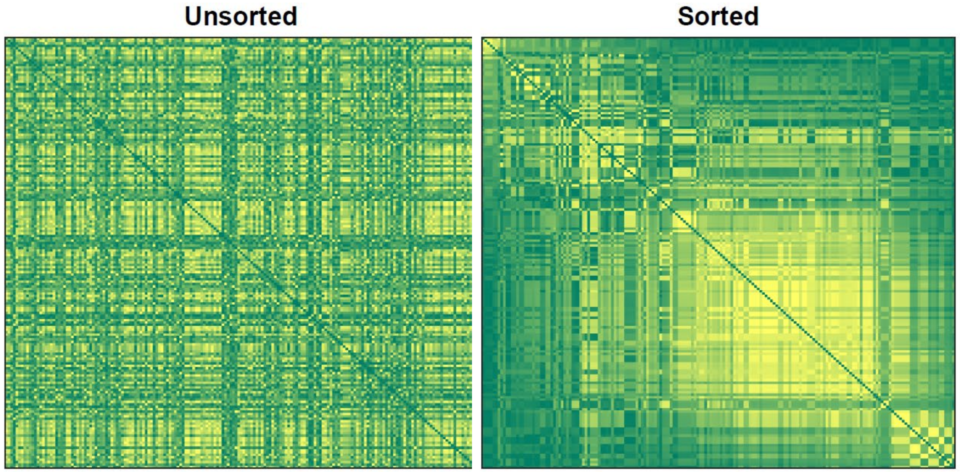
Subject to $\boldsymbol{d}^T \boldsymbol{d} = 1$

$\boldsymbol{d}^T \boldsymbol{1} = 0$

Algorithm 1: GSP based PMU sorting algorithm

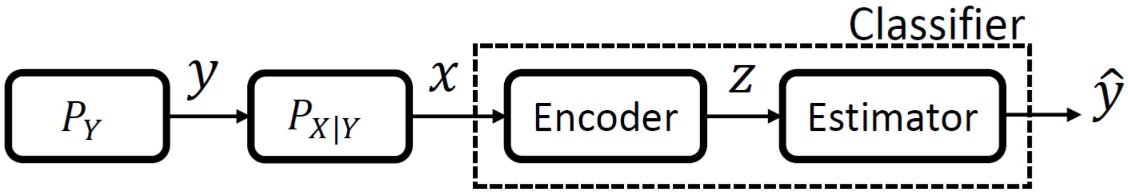
1 Obtain the Pearson correlation coefficients between PMUs;
2 Construct weight matrix W and Laplacian graph L;
3 Take eigendecomposition of L;
4 Sort PMUs according to the eigenvector corresponding to the second smallest eigenvalue of L;

Visualization of Spatial Correlation Matrix of PMU Measurements



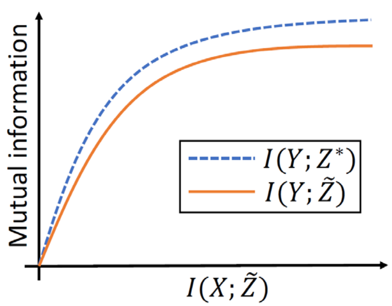
Information Loading-based Regularization

- Background
  - Abstract Representation of Deep Neural Network based Classifier

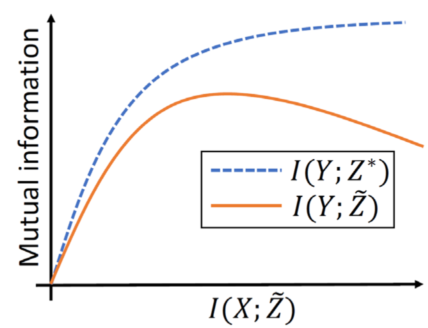


- Main Idea
  - Control the amount of information compression between the input layer and the last hidden layer of a deep neural network
  - Balance memorization and generalization

Low entropy input feature space



High entropy input feature space

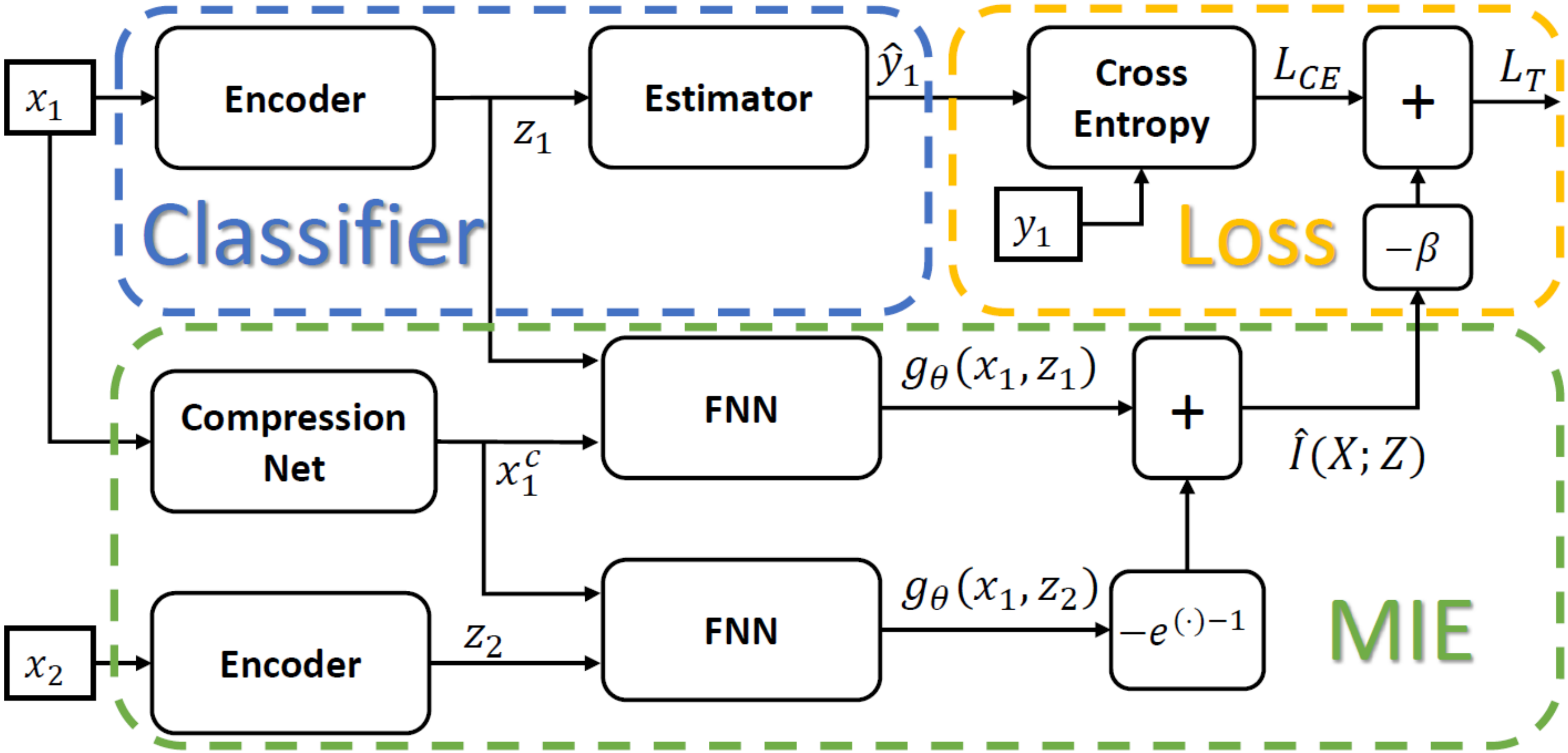


- Algorithm
  - Augment the typical cross-entropy loss function with estimated mutual information between the input layer and the hidden representation

$$L_T = L_{CE} - \beta \hat{I}(X;Z)$$

Overall Neural Network Architecture*

- Neural Classifier, Mutual Information Estimator, Loss Function Augmentation



* J. Shi, B. Foggo, and N. Yu, "Power System Event Identification based on Deep Neural Network with Information Loading," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5622-5632, Nov. 2021.

Numerical Study Results

- Dataset Description
  - 2 years of PMU data from Eastern Interconnection
  - 1247 labeled Events, 187 PMUs (Training, Validation, Testing)
- Data Augmentation

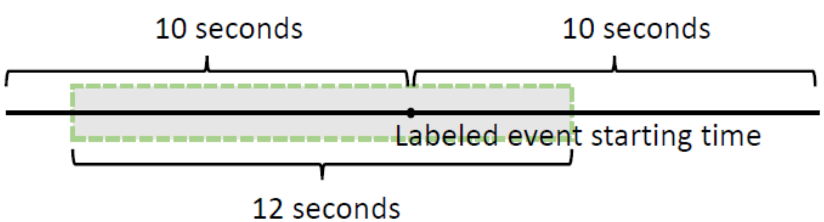
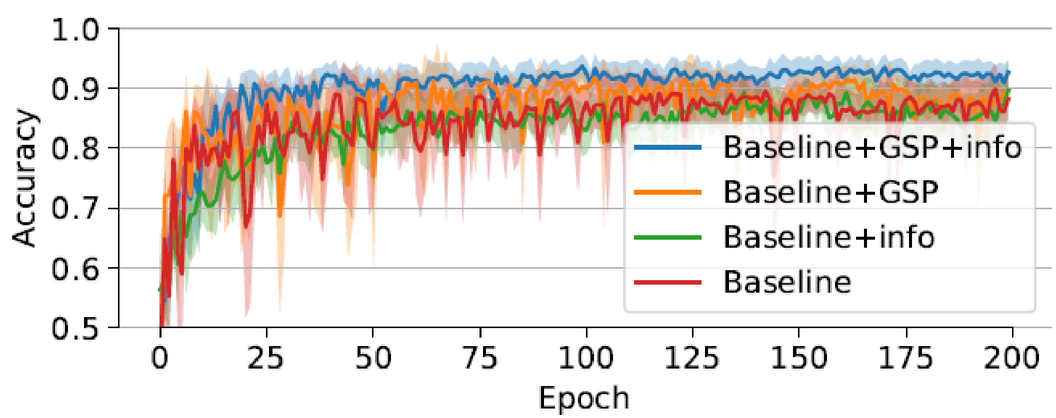


Illustration of Sub-tensor Sampling

TABLE I: Distribution of $PQ|V|f$ snapshots

Class	*Non-event*	*Line-event*	*Generator-event*	*Oscillation-event*
# of snapshots	720	825	756	708

- Performance on Validation Data



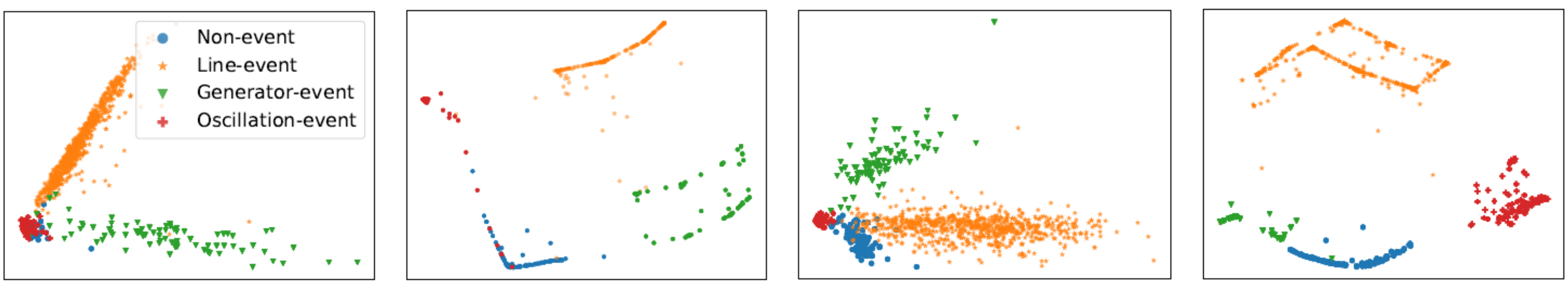
Testing Results and Learned Representation

- F1 Scores on Testing Dataset

	Non-event	*Line-event*	*Generator event*	*Oscillation event*
Baseline	0.554	0.879	0.881	0.208
Baseline+info	0.596	0.928	0.924	0.205
Baseline+GSP	0.894	0.937	0.907	0.922
Baseline+GSP+info	0.973	0.971	0.962	0.986

- Learned Representation

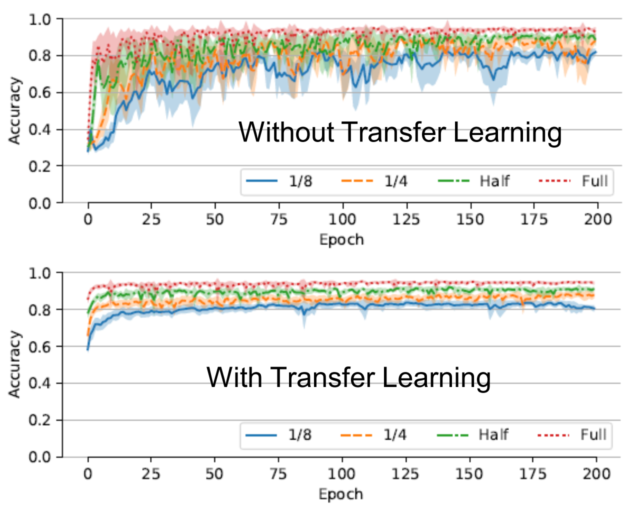
Comparison of representations of different ML methods after linear dimension reduction



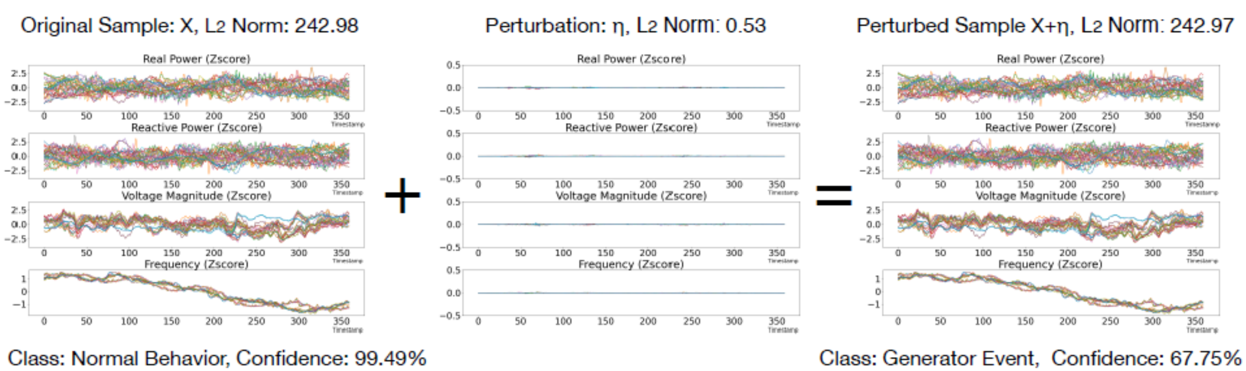
(a) *Baseline.* (b) *Baseline+info.* (c) *Baseline+GSP.* (d) *Baseline+GSP+info.*

Summary and Extensions

- Summary
  - Off-the-shelf ML algorithms often do not work well
  - Physical-domain knowledge + deep learning is needed
  - Information loading helpful in balancing memorization and generalization
- Extensions
  - Transfer Learning: PMU and event log data from one electric grid provide useful info. in analyzing the behavior of another electric network
  - Adversarial attacks and defense: easy to add tailored noise signal to fool event classifier



Example of a small perturbation computed by DeepFool that make the model misclassify from normal operation behavior to generator event



Synthetic Power System Event Data Creation

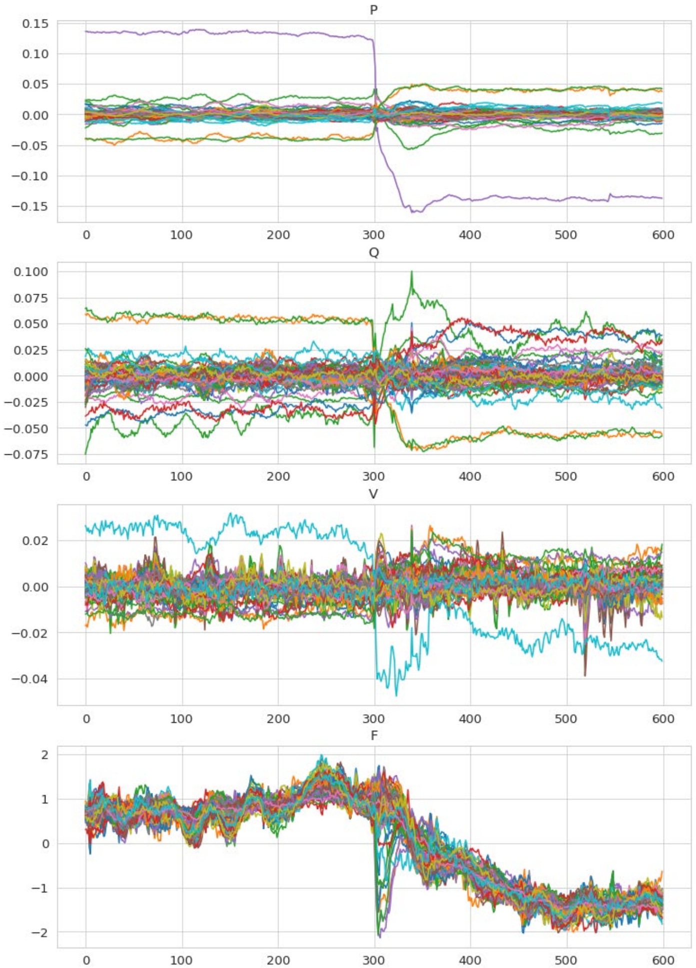
- Why do we need synthetic PMU dataset?
  - Researchers/developers of machine learning algorithms for transmission system always identify the lack of large-scale and realistic PMU data set as a bottleneck for innovation
    - Security concerns, common problem for both academia and industry
  - Benchmarking across algorithms is hard when they’re all tested on different data
- Is PMU data generated from dynamic simulation sufficient?
  - Advantages
    - PMU data generated is consistent with simulated dynamic system
    - Simulation model can be configured to answer any hypothetical research questions
  - Disadvantages
    - IEEE dynamic test cases can not match the complexity of real-world transmission systems
    - Parameterization of generic models (e.g. renewables) are extremely difficult to match observed dataset
    - Lack realistic details (PMU data in response to real-world events often can not be easily emulated by dynamic models, noise, missing values, outliers)

pmuBAGE: The Benchmarking Assortment of Generated PMU Events*

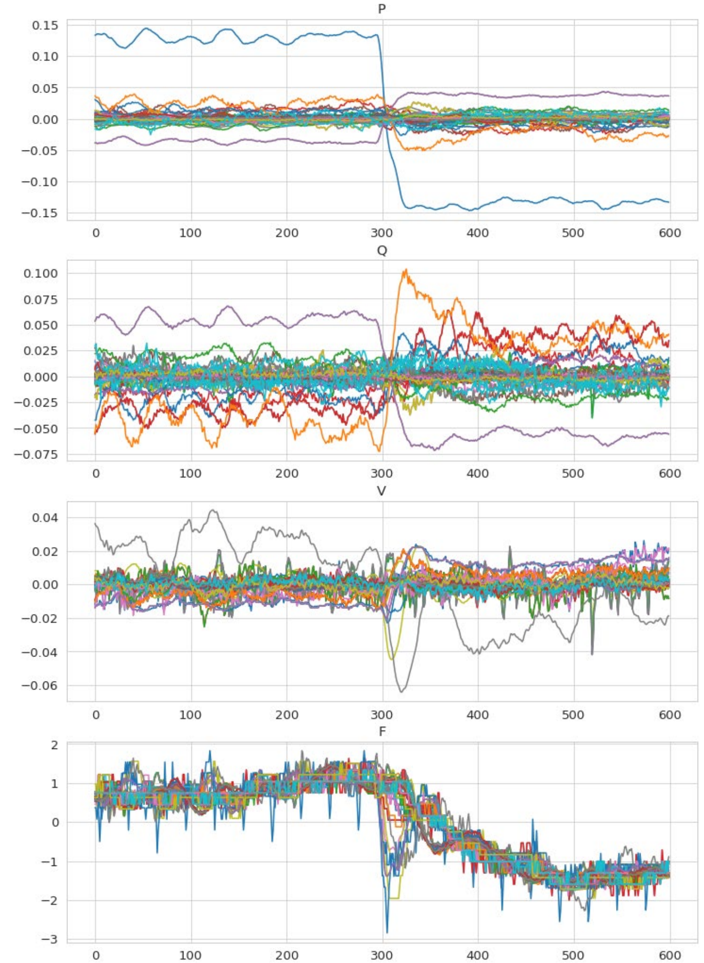
- pmuBAGE: the result of training a generative model on ~1,000 real-world power system events in the Eastern Interconnection.
- Publicly available at https://github.com/NanpengYu/pmuBAGE
- Advantages: accessibility, homogeneity of results & unprecedented level of realism
- Contains 84 synthetic frequency events and 620 synthetic voltage events
- 4 channels (PQ|V|F), 20 seconds event window length, 100 PMUs
- Key Ideas
  - Decompose PMU data during an event into: *Event Signatures* and *Participation Factors*
  - Event signatures can be separated into two types: inter-event and intra-event
  - Physical event signatures are PMU private and are used directly
  - Statistical participation factors are synthesized with generative model

* B. Foggo, K. Yamashita, N. Yu, "pmuBAGE: The Benchmarking Assortment of Generate PMU Events – Part I and II" https://arxiv.org/abs/2204.01095

pmuBAGE – Frequency Events



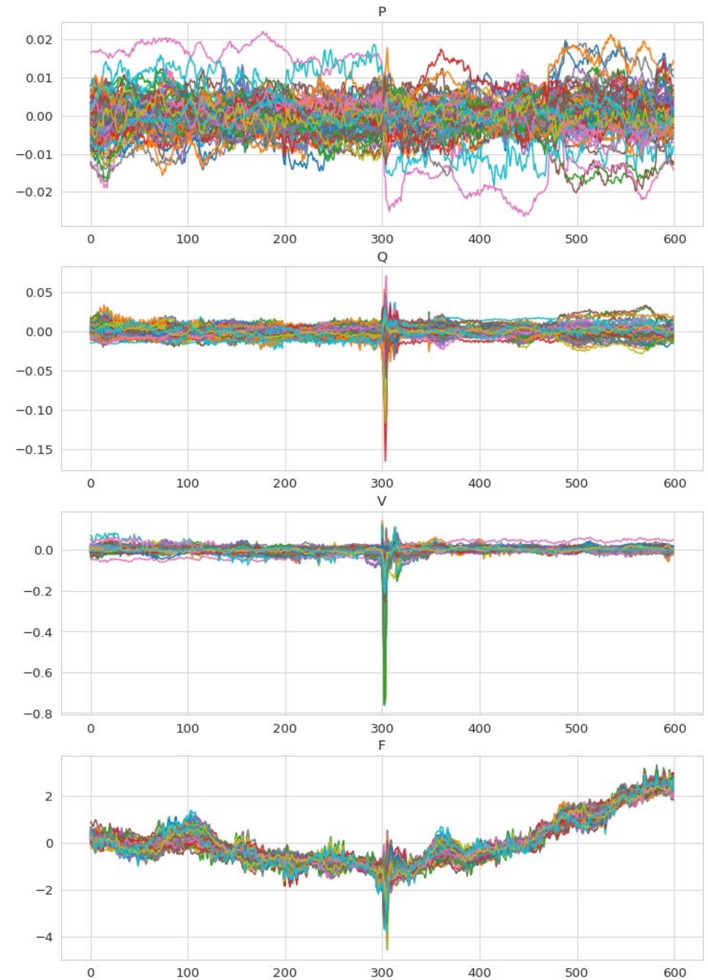
pmuBAGE frequency event



An actual frequency event

The interval between two time indices is 1 / 30 seconds.
The presented data is scaled to per unit values.

pmuBAGE – Sample Voltage Event



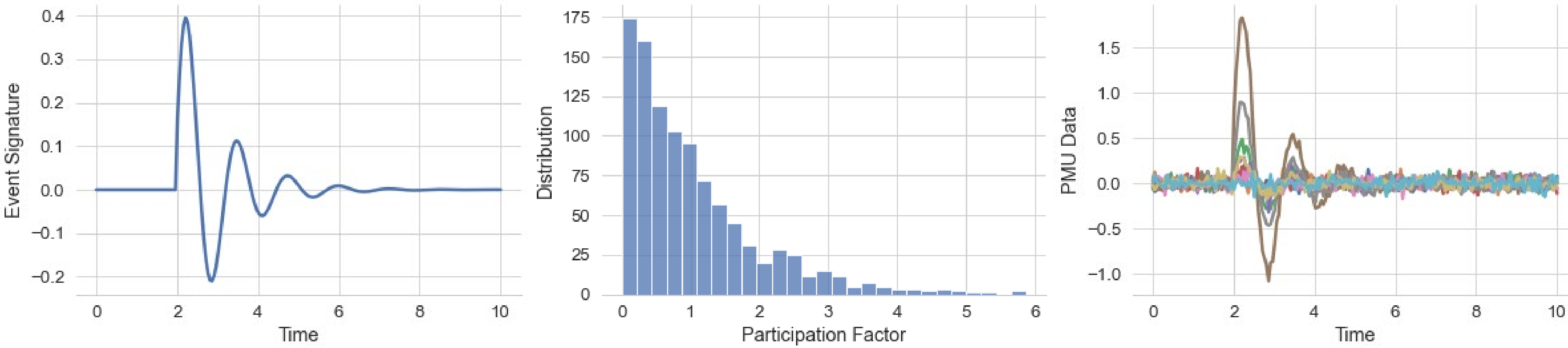
An actual voltage event

pmuBAGE voltage event

The interval between two time indices is 1 / 30 seconds.
The presented data is scaled to per unit values.

The Event-Participation Decomposition*

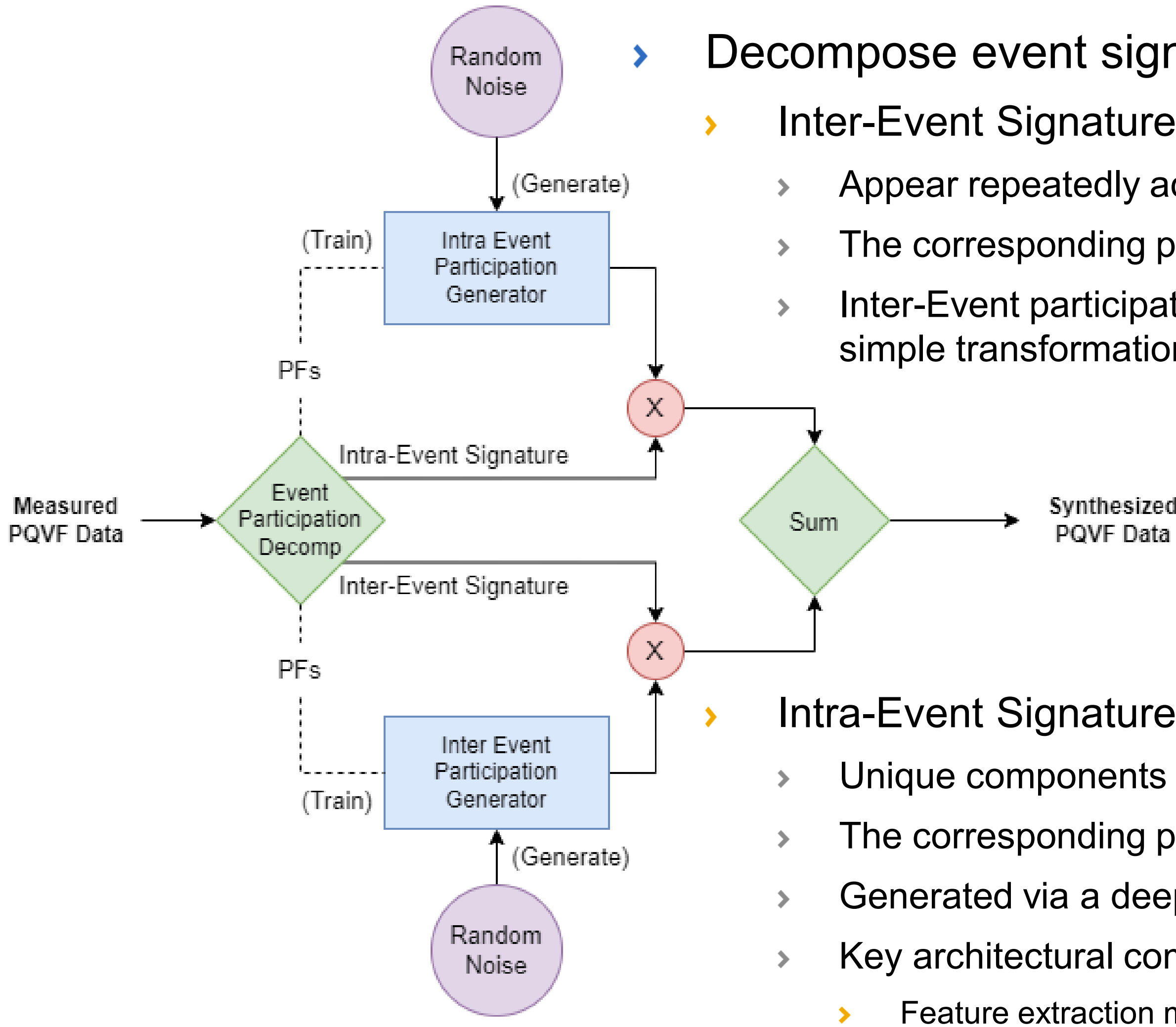
- Decomposes PMU data in an event window into
  - A dynamic component shared by all PMUs – the Event Signature
  - A static component which varies by PMU – the Participation Factor



- Properties of Physical Event Signatures
  - Depend on all PMUs, but don't depend much on any *single PMU*.
  - Event signatures are PMU private and can be used directly to generate synthetic PMU data.
- Properties of Statistical Participation Factor
  - Participation factors are not PMU private by definition.
  - They must be synthesized

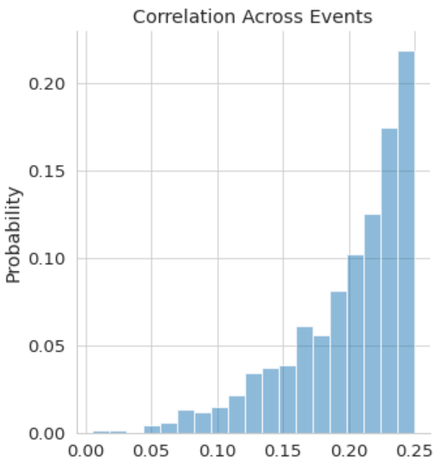
* B. Foggo and N. Yu, "Online PMU Missing Value Replacement via Event-Participation Decomposition," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 488-496, Jan. 2022.

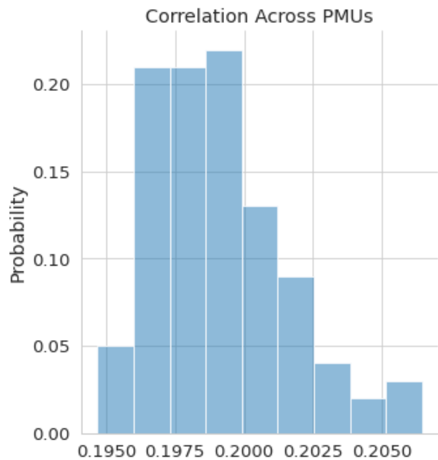
Overall Framework: Generating synthetic PMU data



- Decompose event signatures into 2 types
  - Inter-Event Signature
    - Appear repeatedly across events with little variation
    - The corresponding participation factors are statistically simple
    - Inter-Event participation factors ~ Multivariate Gaussian after simple transformation
  - Intra-Event Signature
    - Unique components of an event
    - The corresponding participation factors are more complicated
    - Generated via a deep generative probabilistic program
    - Key architectural components
      - Feature extraction maps with cascaded convolutional network
      - Loss function with feature mapping and quantile loss

Correlation Analysis and Inception-Like Scoring





- Max correlation between synthetic and real events is 0.25
- No historical events used to train the model are compromised
- Max correlation between synthetic and real PMU measurements is 0.205.
- No PMUs used to train the model are compromised

- Quality of generated PMU data samples measured by “Inception-like score”
  - Train a standard ResNext model to classify event types of labels “frequency” and “voltage”
  - 200 epochs of training with a batch size of 50 with Binary Cross Entropy loss function

Training-Testing	Accuracy	F1	F2
Synthetic-Synthetic	99.9%	94.3%	93.3%
Synthetic-Measured	**94.3%**	**94.2%**	**92.8%**
Measured-Measured	99.8%	94.4%	91.2%
Measured-Synthetic	**93.2%**	**94.3%**	**92.7%**

- No significant degradation in F1 or F2 scores in cross-comparison compared to self comparisons.
- pmuBAGE may serve the community as a standard benchmarking tool for event detection and classification.

Lessons Learned and Next Steps

- Lessons Learned
  - Off-the-shelf machine learning models are often not sufficient
  - Physics-based machine learning is the key to developing breakthrough technology in power system data analytics.
  - The availability of real-world (synthetic) power system data is critical to the accelerated development and benchmarking of data-driven algorithms.
- Next Steps
  - Pilot demonstrations with partner institutions (EPRI and EPG)
  - Deeper integration of physical power system model with machine learning algorithms
  - Interpretable machine learning models for PMU data analytics
  - Making artificial intelligence algorithms actionable in bulk power system
  - Safety and robustness of ML in critical infrastructure systems (bulk power system)
  - Closer collaboration between artificial and operator intelligence

Publications

- Accepted

1. J. Shi, B. Foggo, X. Kong, Y. Cheng, N. Yu, and K. Yamashita "Online Event Detection in Synchrophasor Data with Graph Signal Processing," *IEEE SmartGridComm*, pp. 1-7, 2020.
2. J. Shi, B. Foggo, and N. Yu, "Power System Event Identification based on Deep Neural Network with Information Loading," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5622-5632, Nov. 2021.
3. B. Foggo and N. Yu, "Online PMU Missing Value Replacement via Event-Participation Decomposition," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 488-496, Jan. 2022.
4. X. Kong, B, Foggo, and N. Yu, "Online Voltage Event Detection Using Optimization with Structured Sparsity-Inducing Norms," *IEEE Transactions on Power Systems*, 2022. DOI: 10.1109/TPWRS.2021.3134945.
5. J. Shi, K. Yamashita, and N. Yu, "Power System Event Identification with Transfer Learning Using Large-scale Real-world Synchrophasor Data in the United States," *IEEE ISGT North America*, pp. 1-5, 2022.
6. Y. Cheng, N. Yu, B. Foggo, and K. Yamashita, "Online Power System Event Detection via Bidirectional Generative Adversarial Networks," to appear in *IEEE Transactions on Power Systems*, 2022.
7. X. Kong, K. Yamashita, B. Foggo, and N. Yu, "Dynamic Parameter Estimation with Physics-based Neural Ordinary Differential Equations," *IEEE Power and Energy Society General Meeting*, pp. 1-5, 2022.
8. Y. Cheng, K. Yamashita, and N. Yu, "Adversarial Attacks on Deep Neural-Network-based Power System Event Classification Models," *IEEE ISGT Asia*, pp. 1-5, 2022.

- Under Review and Preparation

1. pmuBAGE: The Benchmarking Assortment of Generate PMU Events – Part I and II
2. A dynamic Behavior-based Bulk Power System Event Signature Library with Empirical Clustering
3. Short-term Forecasting of PMU Data by Attentional Seq2Seq LSTM with Prior Knowledge Matrix and Magnitude Direction Coupling

Contact Information:
Nanpeng Yu, nyu@ece.ucr.edu

Thank You

DOE: Sandra Jenkins, Brian Mollohan, Carol Painter

PNNL: James Follum, Jeffery Banning

UCR: Brandon Foggo, Koji Yamashita, Jie Shi, Yuanbin Cheng, Xianghao Kong, Eamonn Keogh

EPG: Neeraj Nayak, Song Xue, Vikram Budhraja

MTU: Chee-Wooi Ten