MANCHESTER 1824
The University of Manchester

Can machine learning help secure power systems?

Panagiotis (Panos) Papadopoulos
Reader (Associate Professor) and UK Research and Innovation Future Leaders Fellow,
University of Manchester, UK

Contributors: Alinane Kilembe, Ifigeneia Lamprianidou, Luke Benedetti, Faheem Ul Haq, Robert Hamilton, Tabia Ahmad

DTU summer school, 19/05/2026

MANCHESTER 1824
The University of Manchester

Outline

- Background, overview and motivation
- Dynamics and challenges in converter dominated systems
- Explainability of ML models and functional approximations
  - Going beyond the notion that ML models are just black-box predictors (10 slides on SHAP)
- Functional approximations from ML models
  - One step even further: Getting functional approximations using symbolic regression and KANs
- Causal ML – Can we go beyond finding correlations?
- Decision support
  - Embedding detailed stability constraints in optimisation
- Control
  - Reinforcement learning for emergency control
- Topology: Spatio-Temporal Graph Neural Networks including the propagation of dynamics
- Bonus: Defining “small-signal system strength”

Blackouts - What do we know?

Panos Papadopoulos

Climate change

New devices

Decarbonisation

Natural events

Stability & dynamics

New instabilities

Cascading failures

Scale

Uncertainty

Complexity

The energy system

Human behaviour

How we heat our homes

How people use energy

How we charge our car

Expectations

Technical failures

Expected outages

Unexpected failures

Devices can fail

The role of humans

Using data driven methods & machine learning to better understand instabilities · Working in tandem with industry & policy makers.

Power system transformation on the way to achieving net zero

- Key driver for net zero
- System needs and services for secure and economic operation
- Enhanced monitoring, situational awareness and control
- Detailed dynamics
- Deal with complexity and uncertainty

- Measurements and machine learning
- As detailed as possible, real-time representation of power system dynamics
- Explainability/interpretability - Confidence and trust but also understanding
- Enable fast control and dynamics-aware optimisation
- Augment state of the art tools – what is the limit?

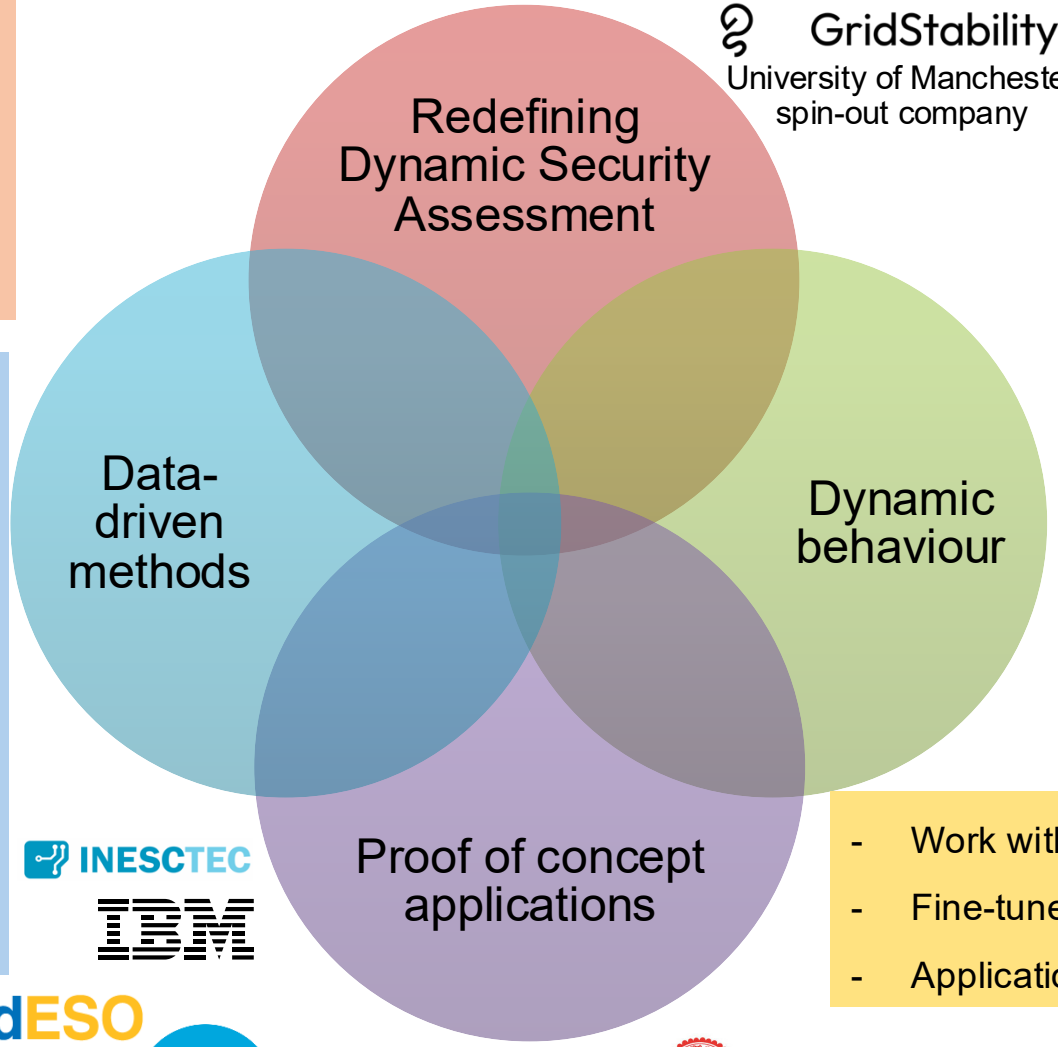
UK Research and Innovation

MANCHESTER PRIZE

Department for Science, Innovation & Technology

https://manchesterprize.org/grid-stability/

GridStability
University of Manchester spin-out company



- Appropriate models and simulation frameworks to represent increasingly complex dynamic behaviour
- Investigate new dynamic phenomena, instabilities and new control approaches (black box vendor models)
- Dynamics of active distribution networks
- Resilience – cascading events

- Work with industrial partners and ML experts
- Fine-tune and test methods and tools
- Applications in control room

Imperial College London
Manchester Metropolitan University
ULB Université Libre de Bruxelles
Aristotle University of Thessaloniki
Democritus University of Thrace
INESCTEC
IBM
SCOTTISHPOWER RENEWABLES
SP ENERGY NETWORKS
nationalgridESO
EPRI Electric Power Research Institute
Hitachi Energy
GE
Rte
The Alan Turing Institute
IIT Delhi
DTU
Virginia Tech
University of Strathclyde Engineering
University of Strathclyde Power Networks Demonstration Centre
National Technical University of Athens

Power system security under increasing complexity and uncertainty

The University of Manchester

Power system dynamic studies

- Power systems are dynamic nonlinear systems
  - Arguably one of most complex man-made systems
  - Up to hundreds of differential equations
- Numerical integration methods – simulations
  - Timescales from <ms up to tens of seconds
  - Computationally intensive

Dynamics increasingly important

- Machine learning and data-driven methods
  - increased visibility, enhanced control, decision support, automation
- Stability, security, and resilience improvement
  - Understanding and mitigating widespread events
- Maximise integration of low/zero carbon technologies
  - Cost efficient manner

Increasing Complexity

- Very different dynamic behaviour of new technologies and high number – need for detailed models
  - Power electronic interfaced (wind, solar, batteries, HVDC links, electric vehicles, electrolysers)
  - Smaller timescales – faster dynamic phenomena
  - Governed by control
  - Strong nonlinearities and hybrid dynamics – limiters, discrete control changes

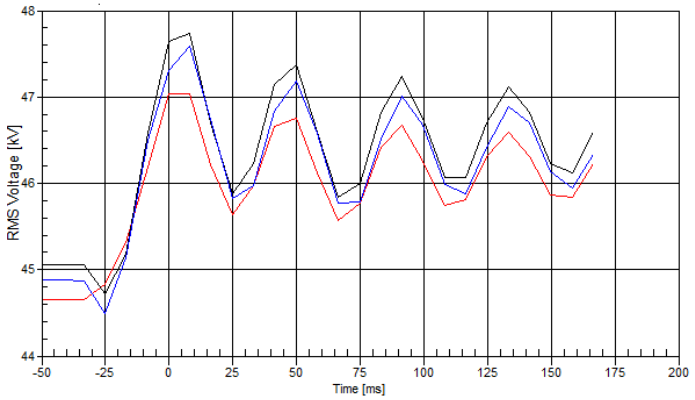
Increasing Uncertainty

- Millions of devices (solar PV, wind farms, batteries, EVs)
- “Exploding” search-space of possible operating conditions
  - Intermittent energy sources and social behaviour
  - Billions of cases if we want to do exhaustive search
  - Affecting operational and planning timescales

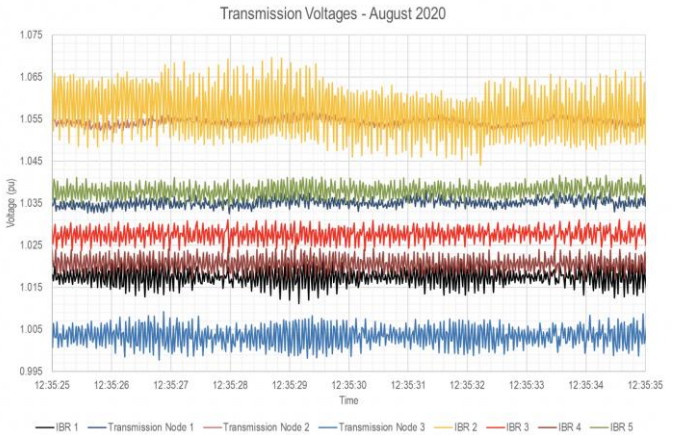
Power System dynamics under increasing complexity and uncertainty

- Modelling and representation of dynamics
  - Characterise systematically mechanisms of instability
  - Importance of investigating multiple operating conditions and understand sensitivities - data-driven and probabilistic approaches
  - Appropriate modelling frameworks (EMT, EMT-RMS, dynamic phasors, hybrid modelling) [4]
- New types of dynamic phenomena and oscillatory interactions
  - Control governing dynamic behaviour and challenges with black-box vendor models (grid forming converters)
  - Fundamental understanding: under what conditions, in which locations, what devices
  - How to mitigate? – damping control, discrete control modes (e.g. for weak grid), sync. comp., operational constraints – services
- Impact on system stability (transient, voltage and frequency)
  - Complex impact with respect to location, control parameters, etc.
  - Not straightforward to understand critical operating conditions and faults
  - Low inertia and locational aspects more pronounced

Canada, Hydro One
(20 Hz, PVs, current control) [2]



Australia, AEMO
(19 Hz, PVs, PLL-shunts) [2]

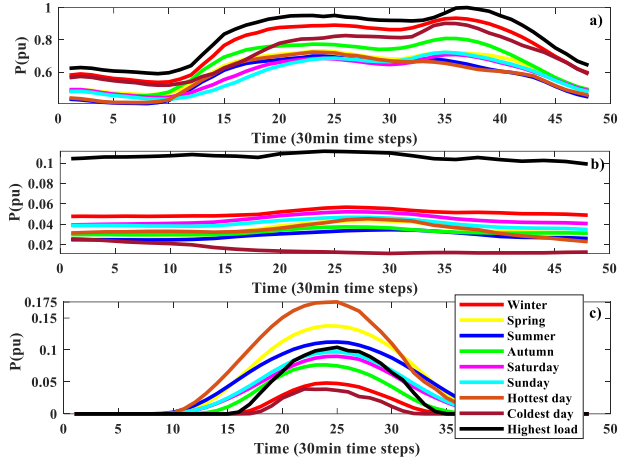


[1] U. Markovic, O. Stanojev, P. Aristidou, E. Vrettos, D. Callaway and G. Hug, "Understanding Small-Signal Stability of Low-Inertia Systems," IEEE Trans. Power Syst., vol. 36, no. 5, pp. 3997-4017, Sept. 2021.
[2] L. Fan *et al.*, "Real-World 20-Hz IBR Subsynchronous Oscillations: Signatures and Mechanism Analysis," *IEEE Trans. Energy Conversion*, 2022, doi: 10.1109/TEC.2022.3206795.
[3] L. I. Benedetti, A. Egea-Àlvarez, R. Preece and P. N. Papadopoulos, "Enabling Characterisation of Dynamic Interactions With Probabilistic Small-Signal Analysis in Converter-Integrated Power Systems," in IEEE Transactions on Power Systems, vol. 41, no. 2, pp. 942-954, March 2026, doi: 10.1109/TPWRS.2025.3612972.
[4] L. Benedetti, P. N. Papadopoulos and A. Egea-Àlvarez, "A Modal Contribution Metric for Quantifying Small-Signal Variability in Power Systems With Converter-Interfaced Generation," in IEEE Transactions on Power Systems, vol. 40, no. 3, pp. 2636-2647, May 2025, doi: 10.1109/TPWRS.2024.3500786.

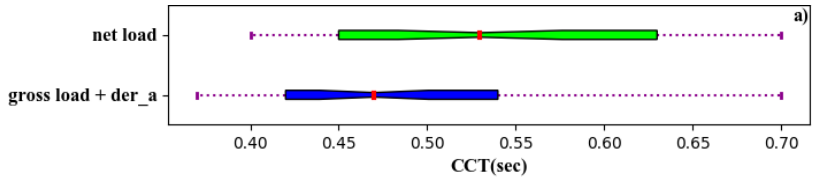
Dynamics and modelling needs for Active Distribution Networks

- Active Distribution Networks
  - Lack of situational awareness and equivalent model parameter sensitivity
  - Data-driven methods for extraction of representative parameters
  - Electric Vehicles (and Heating) changing the mix and dynamics of load (uncertainty due to social behaviour)
- Electric Vehicles and Data Centres
  - Typical dynamic load model structures might not be adequate
- Interaction and services at the interface with transmission
  - How can we confidently procure services from distributed resources (generation, EVs, flexible demand) without activating constraints
  - Providing aggregation information to the ESO for dynamic studies

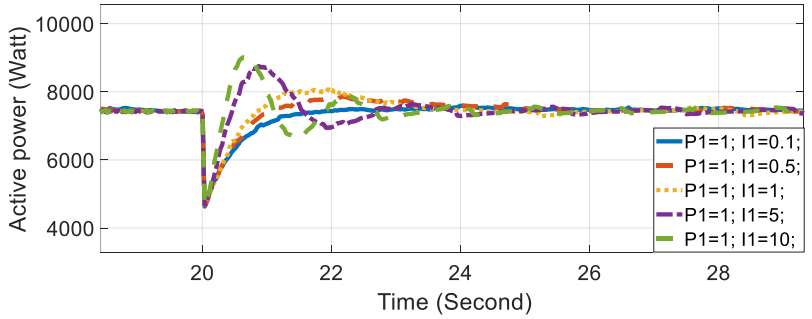
Typical profiles for a) load, b) wind, c) PVs



CCT for a typical day, modelled with and without accounting for ADN



EV responses for different control parameters

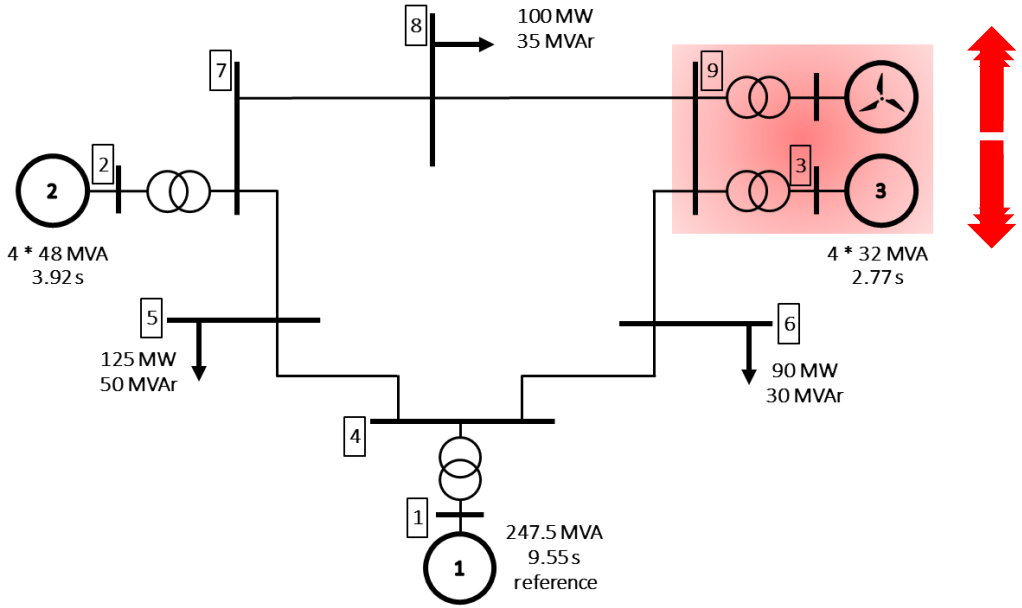


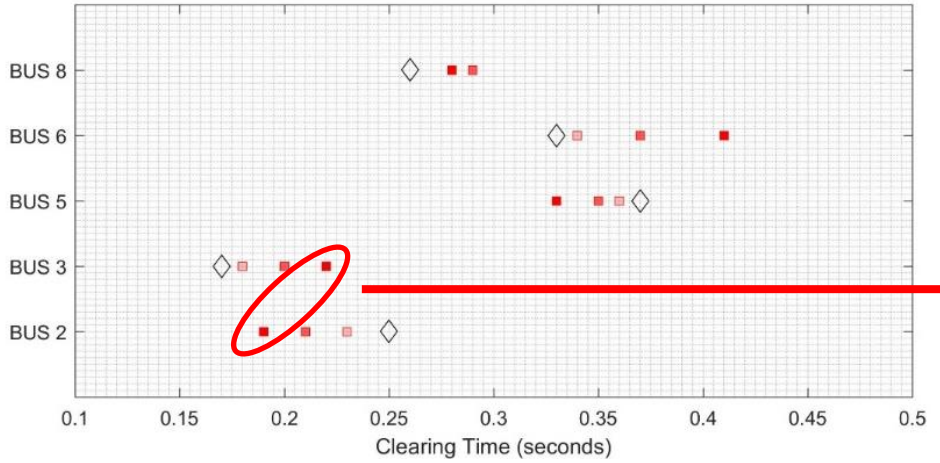
[1] I. S. Lamprianidou, R. I. Hamilton and P. N. Papadopoulos, "Impact of Active Distribution Network Modeling Approaches on Transient Stability Through Sensitivity Analysis," in IEEE Transactions on Power Systems, vol. 41, no. 1, pp. 572-584, Jan. 2026, doi: 10.1109/TPWRS.2025.3589751.

[2] G. A. Barzegkar-Ntovom, E. O. Kontis, T. A. Papadopoulos and P. N. Papadopoulos, "Methodology for Evaluating Equivalent Models for the Dynamic Analysis of Power Systems," in IEEE Transactions on Power Delivery, vol. 37, no. 6, pp. 5059-5070, Dec. 2022, doi: 10.1109/TPWRD.2022.3167136.

[3] Hengqing Tian, Eleftherios O. Kontis, Georgios A. Barzegkar-Ntovom, Theofilos A. Papadopoulos, Panagiotis N. Papadopoulos, “Dynamic modeling of distribution networks hosting electric vehicles interconnected via fast and slow chargers, International Journal of Electrical Power & Energy Systems, Volume 157, 2024, 109811, ISSN 0142-0615, https://doi.org/10.1016/j.ijepes.2024.109811.

A simple example





Change in critical bus

R. I. Hamilton, P. N. Papadopoulos, K. Bell, “An Investigation into Spatial and Temporal Aspects of Transient Stability in Power Systems with Increasing Renewable Generation”, International Journal of Electrical Power and Energy Systems, vol. 115, February 2020.

Can machine learning help keep the system secure?

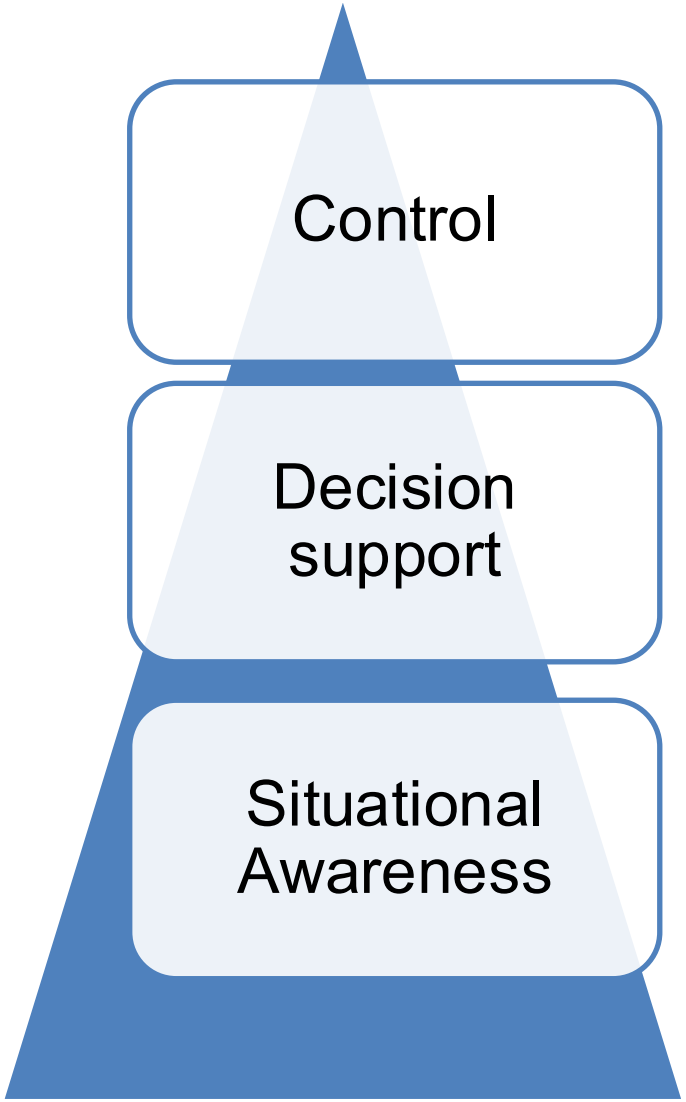
- Existing physics-based approaches
  - Based on first principles
  - Power Flow, Optimal Power Flow
  - Time domain simulations (RMS, EMT), eigen-analysis, bifurcation analysis
  - Analytical methods for stability: excellent understanding and insights but usually require simplifications/assumptions
- Increasing complexity and uncertainty
  - Renewables, Electric Vehicles, converter-interfaced – very different dynamic behaviour
  - New dynamic phenomena, not well understood, requiring a lot of modelling detail
  - A lot more scenarios to investigate – better understanding of risks and/or more cost-efficient operation
- What can Machine Learning do?
  - Speed up security/stability assessment – up to 1000s of times (with the downside of getting it wrong sometimes)
  - Help in getting insights into complex underlying behaviours

We can consider stability/dynamics when computational effort does not currently allow us

P. N. Papadopoulos, S. Chatzivasileiadis and A. Marot, "Can Machine Learning Help Keep the System Secure? Power Systems and Change - Addressing the Increasing Complexity and Uncertainty During the Energy Transition," in *IEEE Power and Energy Magazine*, vol. 22, no. 6, pp. 100-111, Nov.-Dec. 2024, doi: 10.1109/MPE.2024.3421388.

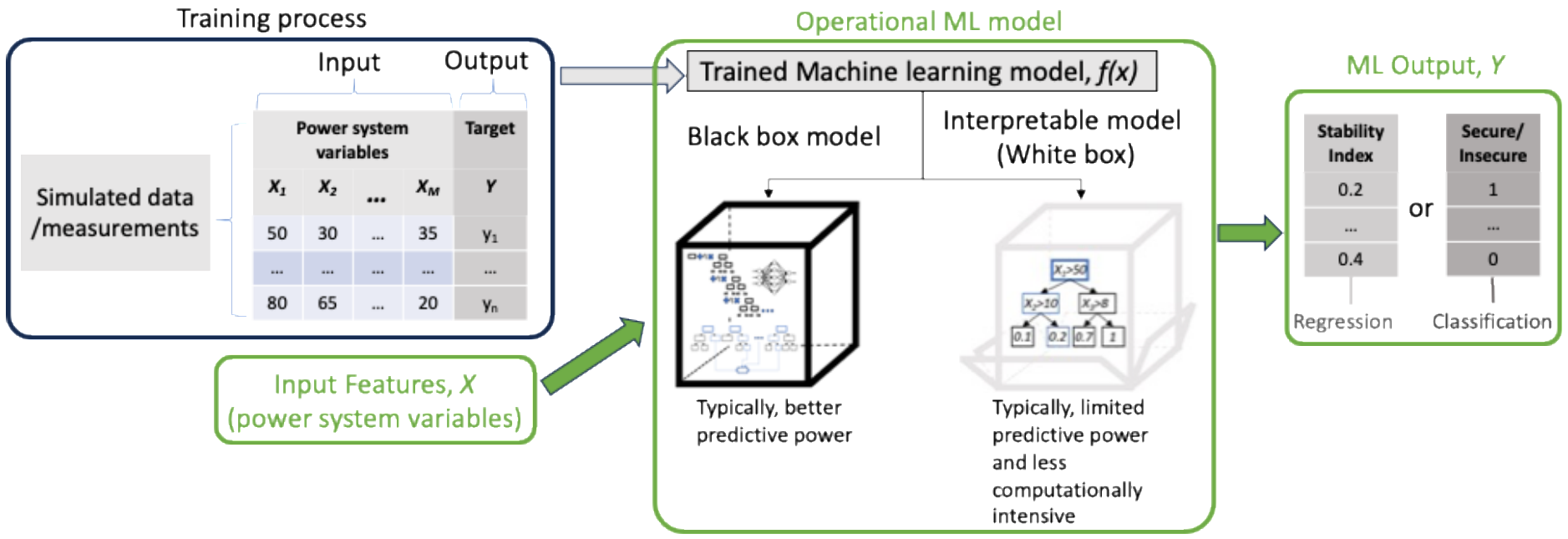
How can machine learning help security assessment?

- Key advantage is speed and fast screening (up to 1000s of times faster)
  - Planning stage – too many scenarios (not enough time)
  - Operational time (closer to real time) – not enough time (a lot of scenarios to run on given time)
  - Need to balance hidden risks or over-securing (which comes at a cost)
- Situational awareness
  - Fast calculation of stability status/metrics
- Decision support
  - Offer options/insights
  - Optimisation (while considering detailed dynamics)
- Control (in cases where we can't do much with current tools)
  - Millions of devices
  - Fast corrective actions



Stability/security assessment using machine learning

- Methods that take steady-state "snapshot" as input: Find a mapping between input operating conditions and specific stability metrics
  - Binary (safe/unsafe) or multiclass classification
  - Regression – calculation of a stability metric (e.g. critical clearing time or locational/regional RoCoF and nadir)

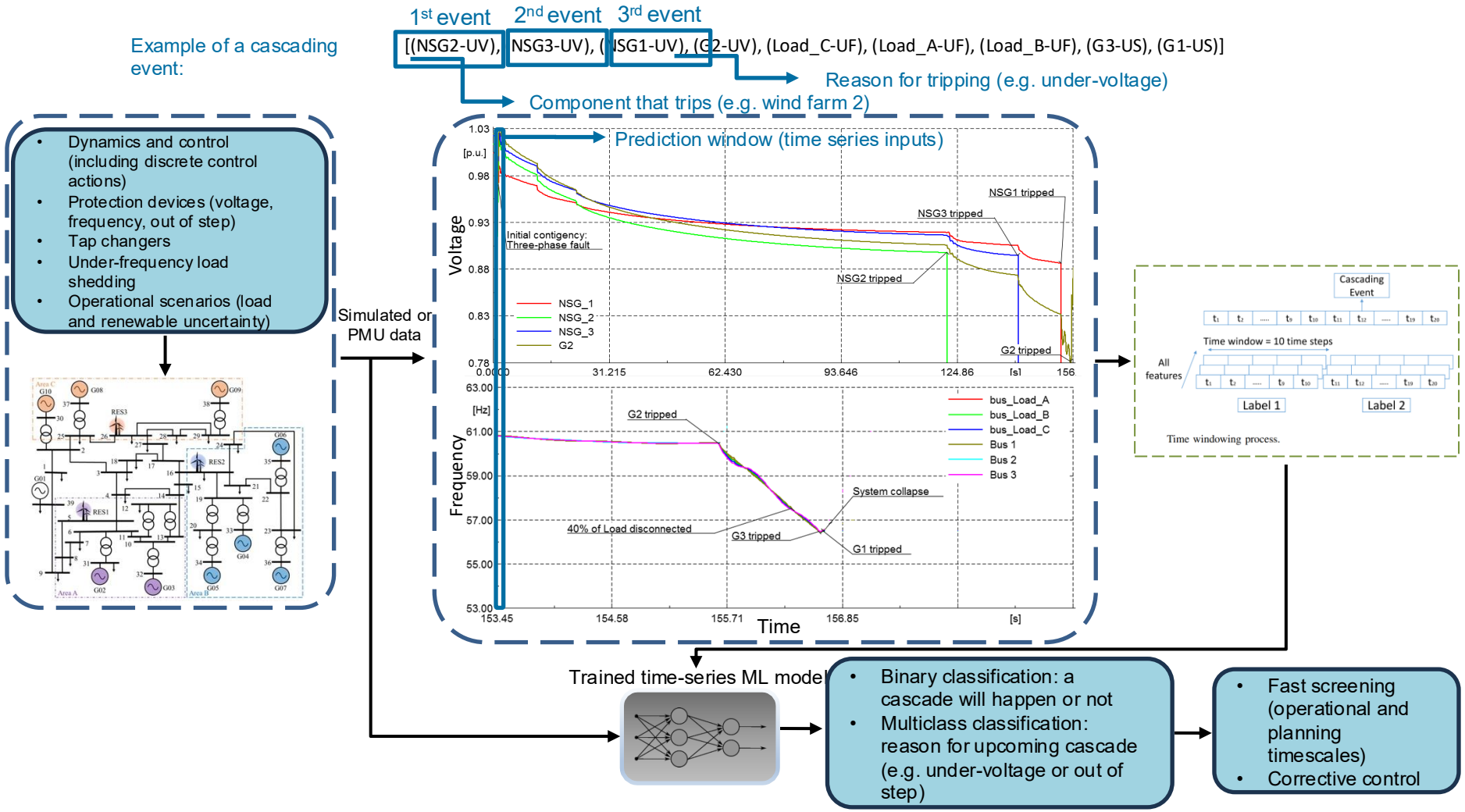


[1] L. Wehenkel, T. V. Cutsem and M. Ribbens-Pavella, "Artificial intelligence applied to on-line transient stability assessment of electric power systems," 1986 25th IEEE Conference on Decision and Control, Athens, Greece, 1986, pp. 649-650, doi: 10.1109/CDC.1986.267413.

[2] R. I. Hamilton and P. N. Papadopoulos, "Using SHAP Values and Machine Learning to Understand Trends in the Transient Stability Limit," in *IEEE Transactions on Power Systems* ,doi: 10.1109/TPWRS.2023.3248941.

Stability/security assessment using machine learning

- Methods that take a time window as input – closer to real time
  - Binary/multiclass classification
  - Useful for screening
  - (Mostly) corrective control
  - Special protection schemes
- Performing simulations faster
  - Physics-Informed NNs



[1] Georgios A. Nakas, Panagiotis N. Papadopoulos, "Online identification of cascading event sequences in power systems using deep learning," International Journal of Electrical Power & Energy Systems, Volume 169, 2025, 110717, https://doi.org/10.1016/j.ijepes.2025.110717.

[2] Jochen Stiasny, Baosen Zhang, Spyros Chatzivasileiadis, "PINNSim: A simulator for power system dynamics based on Physics-Informed Neural Networks," Electric Power Systems Research, Volume 235, 2024, 110796, https://doi.org/10.1016/j.epsr.2024.110796.

MANCHESTER 1824
The University of Manchester

CHARACTERISING NEW TYPES OF DYNAMIC PHENOMENA

Probabilistic small signal stability assessment – how do oscillatory modes behave across a range of different operating conditions?

[1] P. N. Papadopoulos and J. V. Milanović, "Probabilistic Framework for Transient Stability Assessment of Power Systems With High Penetration of Renewable Generation," in IEEE Transactions on Power Systems, vol. 32, no. 4, pp. 3078-3088, July 2017, doi: 10.1109/TPWRS.2016.2630799.

[2] L. I. Benedetti, A. Egea-Àlvarez, R. Preece and P. N. Papadopoulos, "Enabling Characterisation of Dynamic Interactions With Probabilistic Small-Signal Analysis in Converter-Integrated Power Systems," in IEEE Transactions on Power Systems, vol. 41, no. 2, pp. 942-954, March 2026, doi: 10.1109/TPWRS.2025.3612972.

[3] L. Benedetti, P. N. Papadopoulos and A. Egea-Àlvarez, "A Modal Contribution Metric for Quantifying Small-Signal Variability in Power Systems With Converter-Interfaced Generation," in IEEE Transactions on Power Systems, vol. 40, no. 3, pp. 2636-2647, May 2025, doi: 10.1109/TPWRS.2024.3500786.

Power system small-signal modelling

- Linearised models, analysis of which is valid in a small region around the operating point.
- Useful because it allows us to:
  - apply linear analysis techniques (we focus on eigenvalue analysis),
  - identify instabilities and/or problematic oscillations,
  - characterise oscillations/interactions/instability based on participation factors,
  - tune and/or design control systems (e.g., through parametric sweeps or more complex optimisation algorithms),
- Not useful for:
  - Nonlinear dynamics, limit cycles, hybrid dynamics

Eigen-analysis

Small-signal modelling

State-space representation

$$\frac{d\vec{x}}{dt} = \boldsymbol{A}\vec{x} + \boldsymbol{B}\vec{u} = f(\vec{x},\vec{u}) \qquad \vec{y} = \boldsymbol{C}\vec{x} + \boldsymbol{D}\vec{u} = g(\vec{x},\vec{u})$$

Linearise using Taylor approximation and discarding all terms above first order

$$\frac{d\Delta x_i}{dt} = \frac{\delta f_i}{\delta x_1}\Delta x_1 + \cdots + \frac{\delta f_i}{\delta x_N}\Delta x_N + \frac{\delta f_i}{\delta u_1}\Delta u_1 + \cdots \frac{\delta f_i}{\delta u_R}\Delta u_R$$

Eigenvalues

Eigenvalues of state matrix, $\boldsymbol{A}$, correspond to system poles

$$0 = \det(\boldsymbol{A} - \lambda \boldsymbol{I})$$

Real part gives 1/damping time constant and imaginary part gives oscillation frequency

$$\lambda_n = \sigma_n \pm j\omega_n$$

Stability requires all negative real parts

$$0 \geq (\Re\lambda = \sigma)$$

Eigenvectors

Right eigenvector (mode shape)*: On which states does a mode show up?

$$\boldsymbol{A}\vec{\phi}_n = \vec{\phi}_n\lambda_n$$

Left eigenvector: How much does the variation of a state excite each mode?

$$\overrightarrow{\psi_n^T}\boldsymbol{A} = \lambda_n\overrightarrow{\psi_n^T}$$

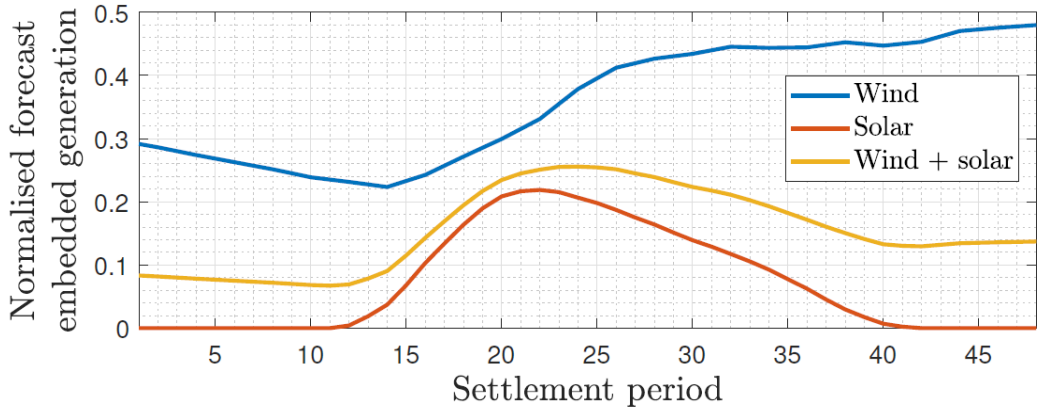
Participation factor: combined effects of left and right eigenvectors

$$p_{n,i} = \frac{|\phi_{n,i}||\psi_{n,i}|}{\sum_{n=1}^{N}|\phi_{n,i}||\psi_{n,i}|}$$

*Often used to determine which states are swinging with or against (in or out of phase to) each other. E.g., interarea vs. local modes

Small signal dynamics with high converter penetration

- Increase of variable renewable energy-sourced generation
- Increasing range, and uncertainty, of operating point
- Higher complexity, different control implementations (especially for GFM)
- Investigate multi-machine (small-signal) interactions and impact of varying system parameters
  - Different choices of controller architecture for GFM and GFL
- How do new types of oscillations (e.g. SSOs) behave or change?
- Deterministic linear analysis around a single operating point is insufficient
  - E.g., eigenvalues can vary significantly, and in complex patterns
  - Can miss instability, mischaracterise novel interactions, and misinform design choices

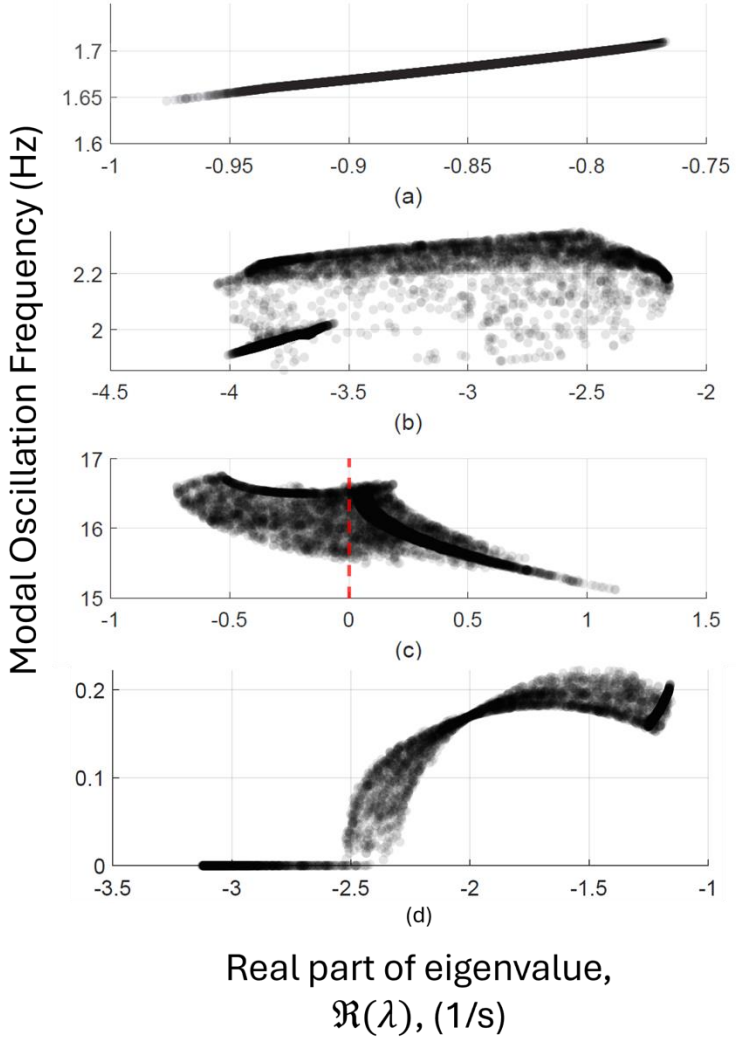


Embedded wind and solar generation over a single day (8th August 2024), taken from National Energy System Operator Data Portal [1].

[1] National Energy System Operator, NESO Data Portal, https://data.nationalgrideso.com/, accessed Jul. 17, 2024.

Small signal dynamics with high converter penetration

- Increase of variable renewable energy-sourced generation
- Increasing range, and uncertainty, of operating point
- Higher complexity, different control implementations (especially for GFM)
- Investigate multi-machine (small-signal) interactions and impact of varying system parameters
  - Different choices of controller architecture for GFM and GFL
- How do new types of oscillations (e.g. SSOs) behave or change?
- Deterministic linear analysis around a single operating point is insufficient
  - E.g., eigenvalues can vary significantly, and in complex patterns
  - Can miss instability, mischaracterise novel interactions, and misinform design choices

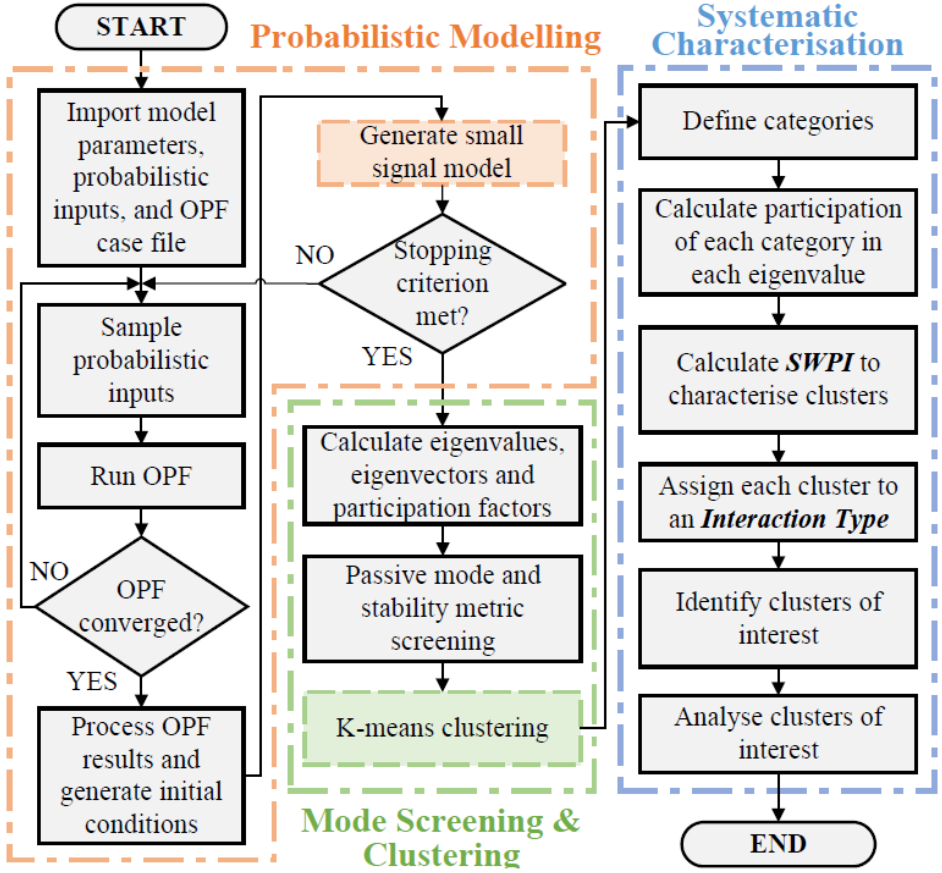


Eigenvalues varying with the operating point.

Probabilistic small-signal interaction analysis framework

- Monte-Carlo analysis
- Large, complex systems
- Huge amounts of data
  - E.g., eigenvalues, eigenvectors & participation factors

- A framework to condense the data into the ***key distinct dynamic phenomena*** (interactions) present on the system, across the full operating range.
  - Clustering based on participation factors.
- Furthermore, it enables ***characterisation and probabilistic analysis*** of the identified distinct dynamics.
  - Stability-weighted participation index (SWPI)



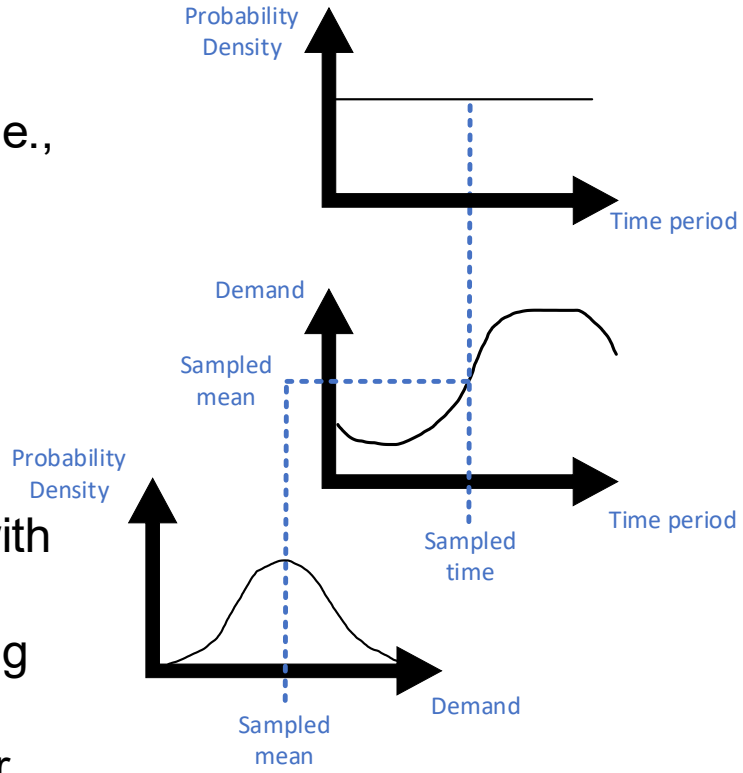
L. I. Benedetti, A. Egea-Àlvarez, R. Preece and P. N. Papadopoulos, "Enabling Characterisation of Dynamic Interactions With Probabilistic Small-Signal Analysis in Converter-Integrated Power Systems," in IEEE Transactions on Power Systems, vol. 41, no. 2, pp. 942-954, March 2026, doi: 10.1109/TPWRS.2025.3612972. – pre-print

MANCHESTER 1824
The University of Manchester

Uncertainty Sampling

- Sample uniform distribution: 0 – 47 (i.e., half hourly periods)
- Mean value of final distribution: value of demand curve at sampled time period
- Final distribution
  - Add Gaussian with $3\sigma = 0.1 \times \mu$
- Independent sampling

- (similar procedure for wind generation)



K-means Clustering

- Monte Carlo simulation -> too many modes to analyse!
- Mode tracking does not apply due to random variation of operating point
- Clustering based on participation factors (PFs)
  - i.e., modes within cluster have similar PF-based characteristics

- K-means (along with elbow method)
  - Only input data matrix (PF vectors for each mode) required.
  - Does not need heuristic tuning of parameters such as Euclidean distance (difficult with high dimensional data, i.e., PFs)
- Limitation: sensitive to initialisation of cluster centroids
  - Use *k-means*++ seeding
  - Use multiple-initialisations approach
    - Run with lowest within-cluster-sum-of-squares is chosen

MANCHESTER 1824
The University of Manchester

Systematic characterisation – defining categories of interest

Category-participation factors (PFs)

- Assign states to user-defined categories
- Sum of PFs of states assigned to those
  - dynamic phenomena,
  - types of generator
  - individual generators, controllers, etc.

$$p^*_{c,k} = \sum_{j=1}^{N_{cs}} (p_{\beta_j,k})$$

TABLE I: CUSTOM DYNAMIC PHENOMENA CATEGORIES

Categories	Elements
Active power - frequency (P-F)	Active power controller (APC), phase-locked loop (PLL), SG rotor, SG gov, SG rotor windings
Reactive power - voltage (Q-V)	Voltage magnitude controller (VMC)/reactive power controller (RPC), SG AVR/exciter, SG PSS
IVC-related (IVCr)	Inner voltage controller
ICC-related (ICCr)	Inner current controller
Measurement filter- & delay-related (MF&Dr)	PWM/control delay, input measurement filters (power filters are included in APC/VMC/RPC)
Network	Branch currents, capacitor voltages, VSC output RLC filters, SG stator

N_{cs} is the number of states in category c. $\vec{\beta}_{N_{cs}\times 1}$ is the vector of indices for the specific states.

MANCHESTER 1824
The University of Manchester

Systematic characterisation (SWPI)

Stability Weighted Participation Index (SWPI)

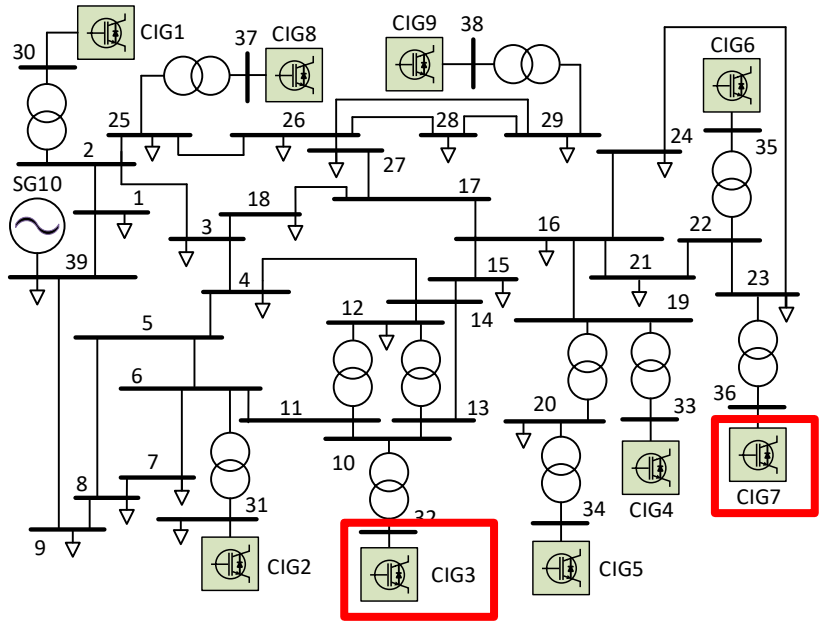
$$SWPI_c = \frac{\sum_{k=1}^{N_{mwc}} \left[p_{c,k}^{*} \cdot \Re(\lambda_k)' \right]}{\sum_{c=1}^{N_C} \left(\sum_{k=1}^{N_{mwc}} \left[p_{c,k}^{*} \cdot \Re(\lambda_k)' \right] \right)}$$

Numerator is the sum of weighted [by $\Re(\lambda_k)'$] category-PFs. The denominator is 2-norm normalisation (SWPI of all categories adds up to 1).

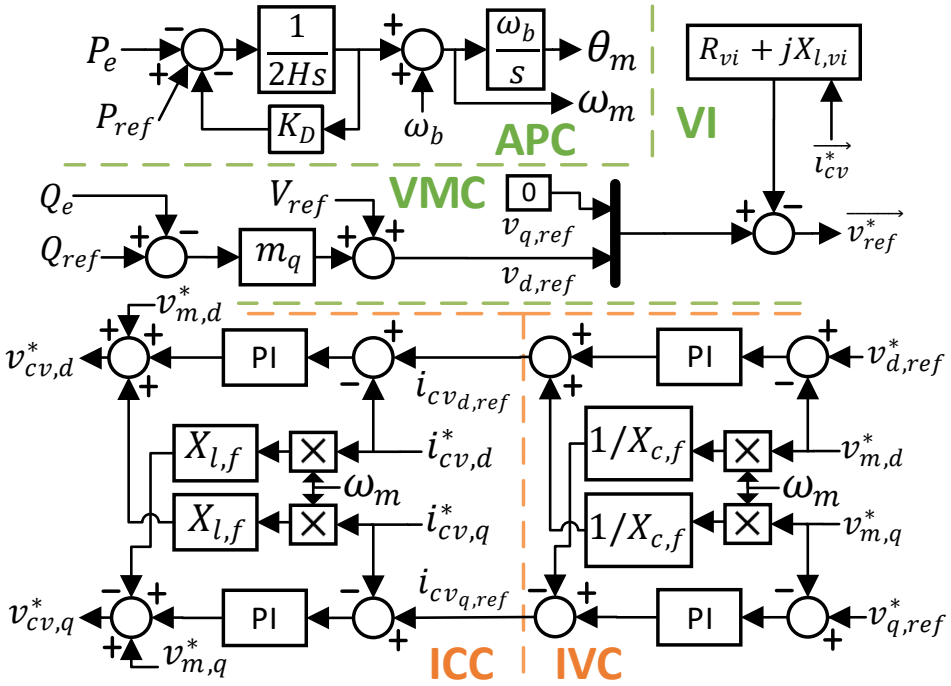
N_{mwc} is the number of modes assigned to the cluster being characterised. $\Re(\lambda_k)$ is the real part of the k^{th} mode in the cluster. N_C is the number of categories. The $\boxdot'$ operator is min-max normalisation.

Application of proposed framework

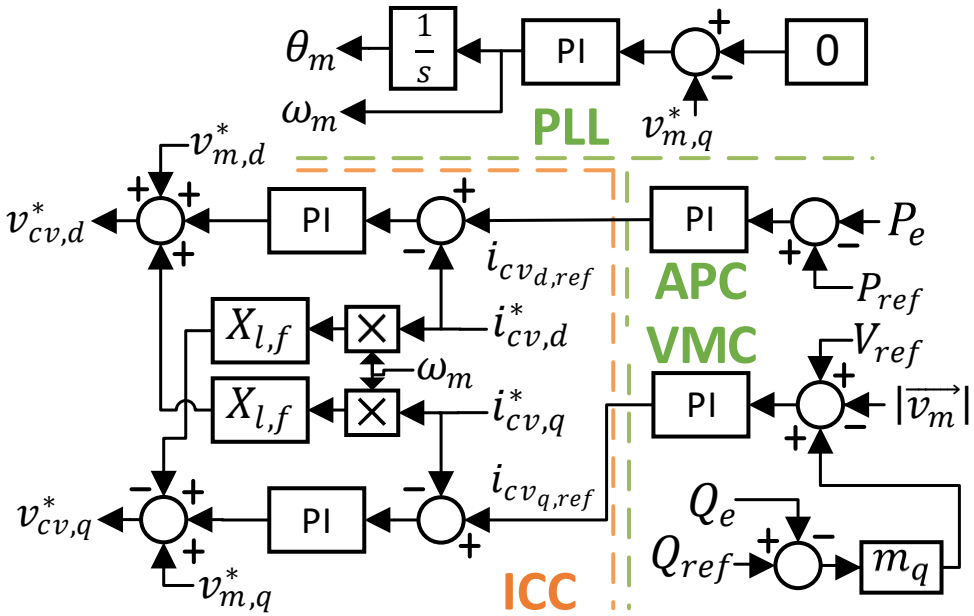
Modified IEEE 39-bus.
2 GFMs, 7 GFLs, 1 SG



Grid-forming (GFM) control

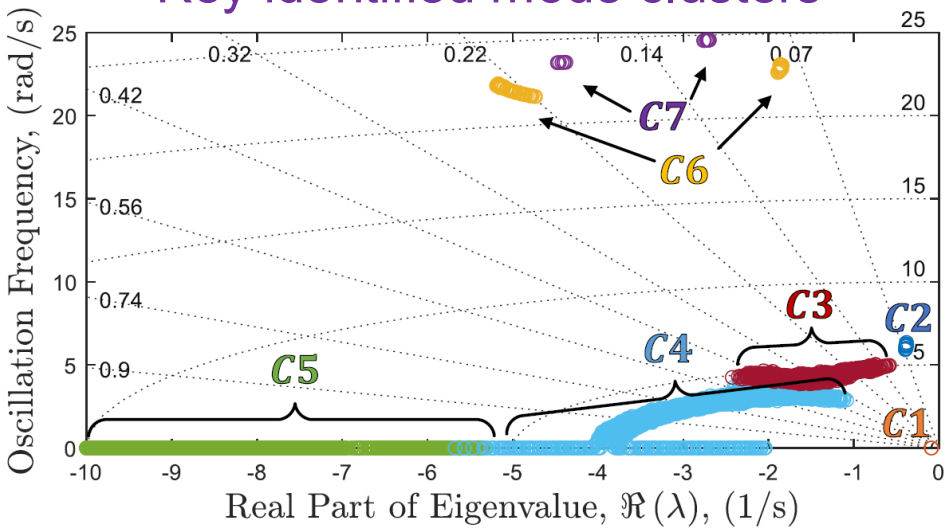


Grid-following (GFL) control



Application of the proposed framework

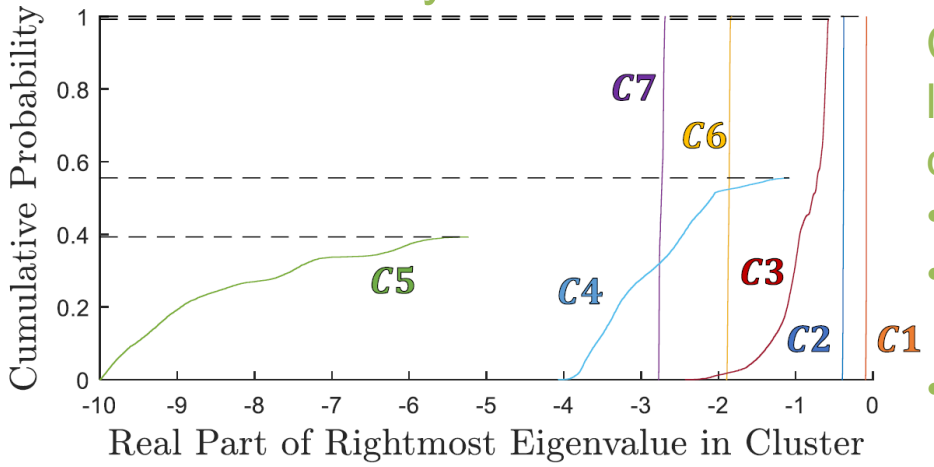
Key identified mode clusters



Modes grouped based on similar participation factor characteristics,

using K-means clustering

Probabilistic analysis of mode clusters



Can observe likelihood of a mode cluster to:

- Be unstable
- Be below a stability margin threshold
- Exist for any given operating point

Characterisation of mode clusters

	C1	C2	C3	C4	C5	C6	C7
GFL1	0.7522	1.579e-05	0.003302	0.01481	0.009972	0.0005114	3.303e-05
GFL2	0.01146	0.0007601	0.01797	0.06919	0.09328	0.002349	0.001032
GFM3	0.001521	0.6166	0.278	0.07127	0.05086	0.3101	0.6829
GFL4	0.01018	0.0002126	0.03572	0.1459	0.2198	0.003597	3.375e-06
GFL5	2.907e-05	0.0001439	0.01991	0.1019	0.1187	0.001511	1.387e-06
GFL6	0.007004	0.001113	0.02982	0.1034	0.09972	0.00524	0.0003117
GFM7	0.0009061	0.3789	0.4762	0.1134	0.08147	0.6701	0.3152
GFL8	0.1202	4.045e-05	0.01538	0.07443	0.0831	0.00179	9.289e-05
GFL9	0.02208	4.757e-05	0.03236	0.1664	0.2007	0.003212	7.971e-05
SG10	0.07448	0.002212	0.09134	0.1393	0.04238	0.001587	0.0003577

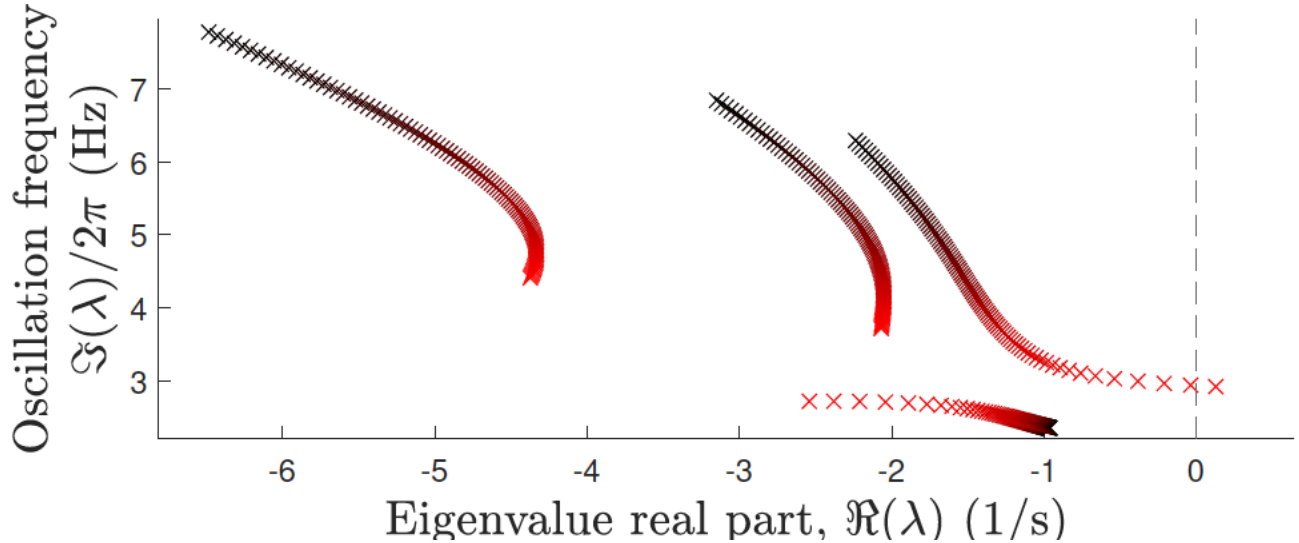
	C1	C2	C3	C4	C5	C6	C7
P-F	0.1009	0.9645	0.9591	0.9012	0.8937	0.0289	0.02174
Q-V	0.8989	0.0001242	0.01101	0.06008	0.02901	0.000538	0.0002329
IVCr	7.851e-05	0.01708	0.01657	0.02557	0.0498	0.4772	0.4768
ICCr	6.069e-07	0.01671	0.005308	0.001818	0.007437	0.4789	0.4801
MF&Dr	3.291e-09	4.483e-06	2.683e-05	3.903e-05	0.0001423	0.0006528	0.0006985
Network	0.0001174	0.001578	0.00801	0.01127	0.01993	0.01377	0.02042

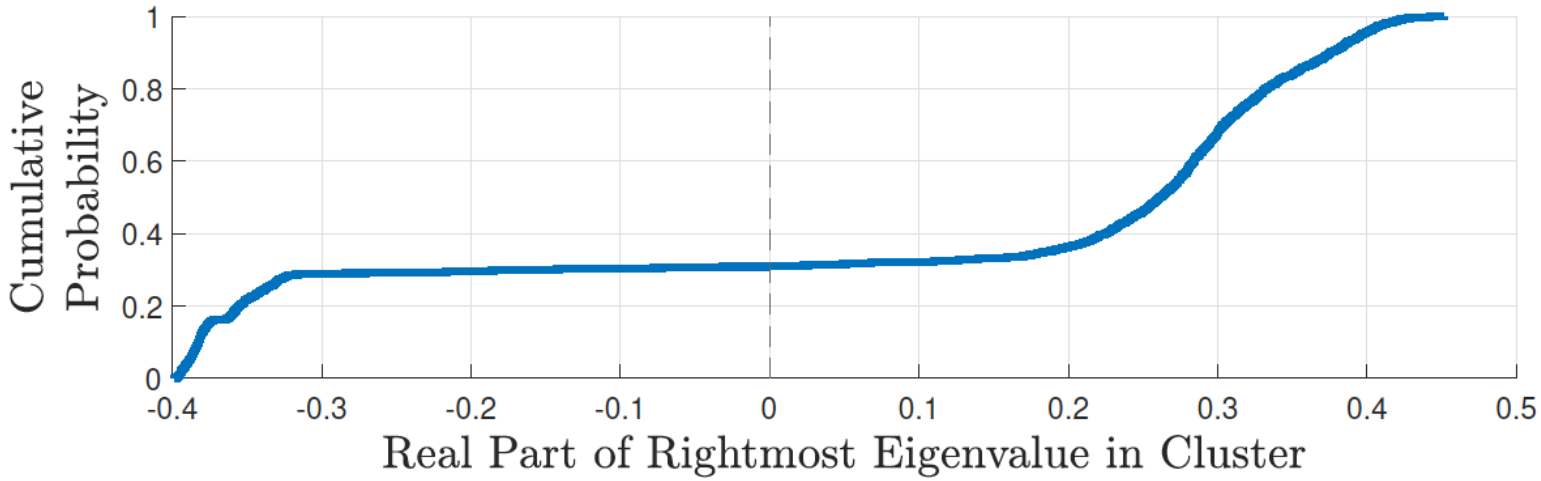
Stability-weighted participation index (SWPI) enables participation factor-like analysis of mode clusters.

What interactions do the clusters represent?

Probabilistic analysis of destabilising sub-synchronous oscillations

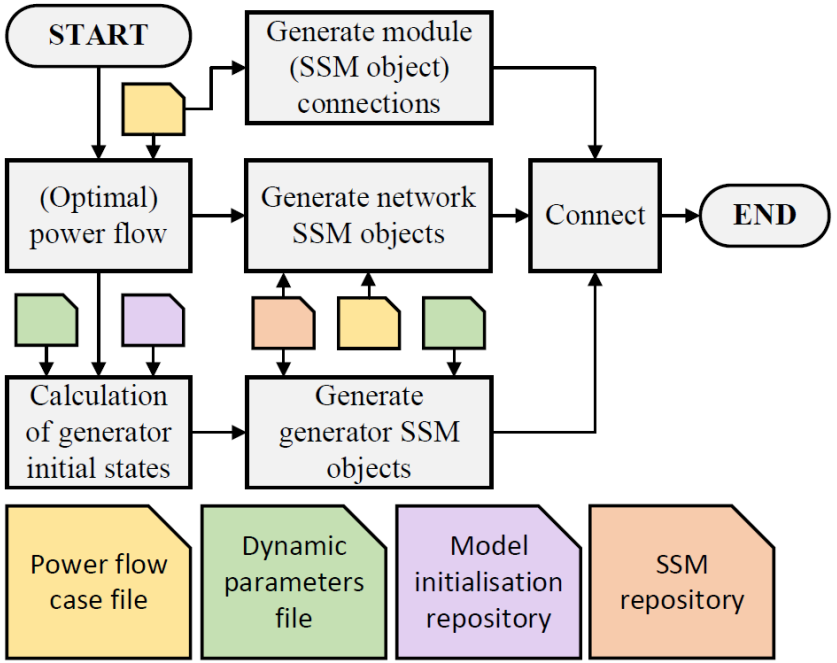
- SSO related to Inner Voltage Controllers of GFMs
- Closed-loop time constant of IVC 7.5ms- 10.44ms (black to red)
  - Realistic tuning
- Probabilistic analysis for 10.44ms
  - 69% probability of instability





Power system small-signal modelling: A MATLAB-based tool and dataset

- Linearised models, analysis in small region around an operating point.
  - identify instabilities and/or problematic oscillations,
  - characterise oscillations/interactions/instability based on participation factors,
  - tune and/or design control systems (e.g. through parametric sweeps or more complex optimisation algorithms),
- EMT-level detail in dq0 (balanced)
- Inputs required: MATPOWER case file and Dynamic parameters file.
- Automatic (initialisation and library):
  - Can quickly initialise and generate small-signal models
  - A repository/library of components is provided (and users can add and/or edit as desired), including detailed grid-following and grid-forming controllers.
- Modular:
  - Easy to add or remove components to a system model.



Dataset: Benedetti, Luke; Egea, Agusti; Papadopoulos, Panagiotis (2024). Results Data: A Modal Contribution Metric for Quantifying Small-Signal Variability in Power Systems with Converter-Interfaced Generation. University of Manchester. Dataset. https://doi.org/10.48420/26412331.v1.

MANCHESTER 1824
The University of Manchester

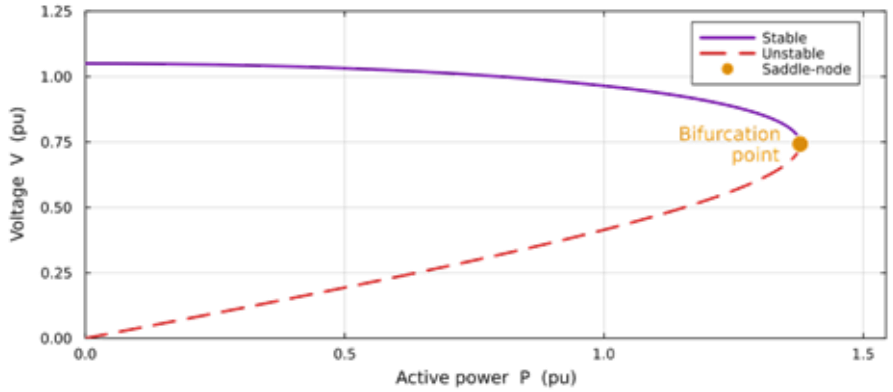
NONLINEAR DYNAMICS AND BIFURCATION ANALYSIS

Grid-forming converter-induced sub-synchronous oscillations

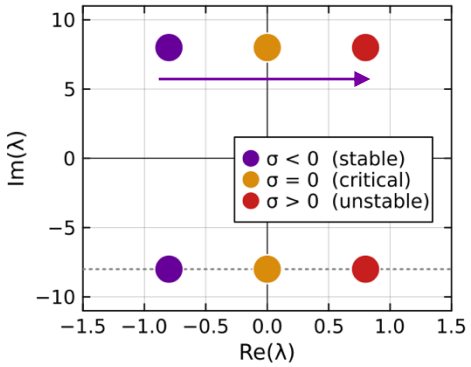
L. Benedetti, R. Preece and P. N. Papadopoulos, “Bifurcation Analysis of Sub-Synchronous Oscillations Related to Grid-Forming Converter Inner Controllers," accepted at PowerUp conference 2026.

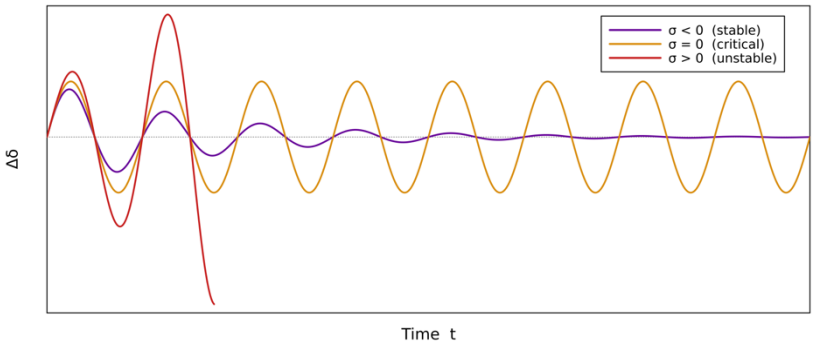
MANCHESTER 1824
The University of Manchester

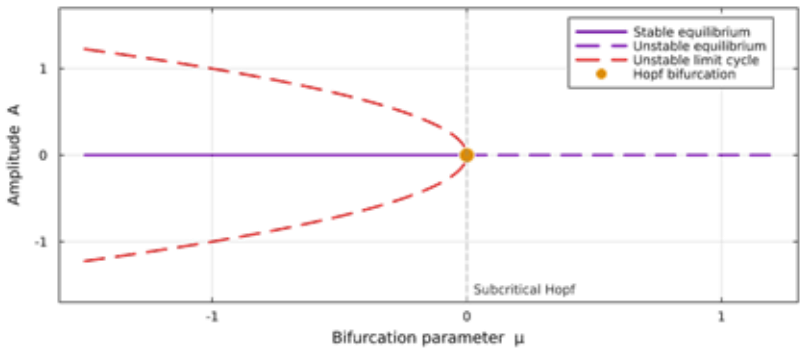
What is a bifurcation?

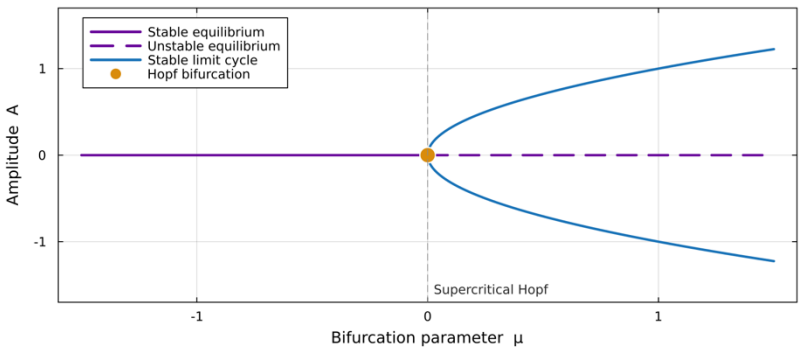


- A qualitative change in system dynamics as parameters vary
  - E.g., equilibrium point stability change
  - Eigenvalue(s) crossing imaginary axis
- Saddle-node bifurcation
  - Stable and unstable equilibrium collide and disappear
  - PV curve
  - Real-valued eigenvalue crossing
- Hopf bifurcation
  - Oscillatory instability
  - Complex conjugate pair of eigenvalues crossing
  - Subcritical
  - Supercritical









Continuation

- Nonlinear dynamical system in ODE form:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{p})$$

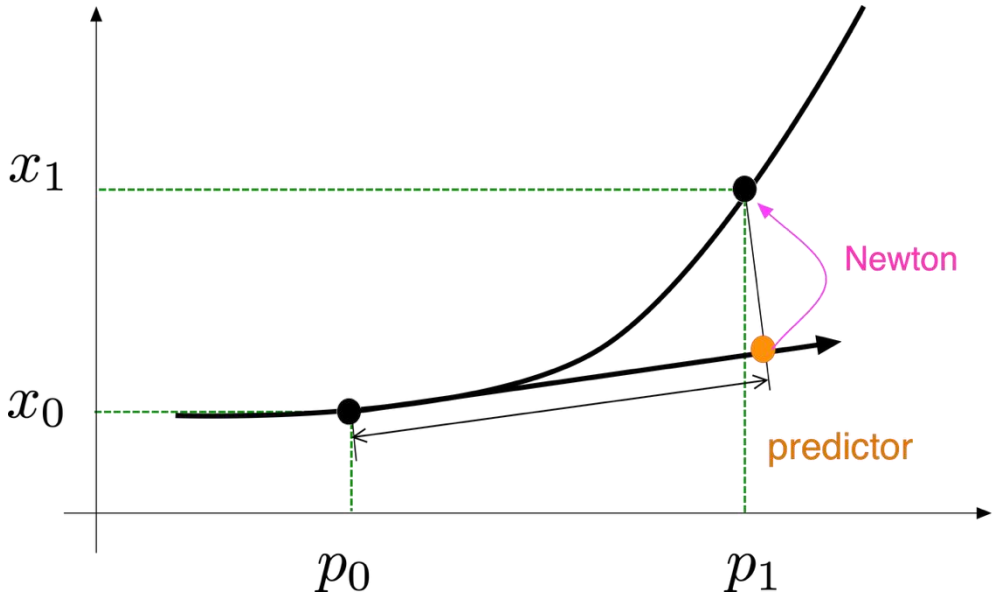
where $\boldsymbol{f} : \mathbb{R}^N \times \mathbb{R}^{N_p} \to \mathbb{R}^N$ denotes a sufficiently smooth vector field with N dynamic states, $\boldsymbol{x}$, and N_p parameters, $\boldsymbol{p}$

- The equilibrium point is

$$0 = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{p})$$

- We want to trace the curve of solutions of the equilibrium point
  - And identify any bifurcations along the way!

Aside: Continuation is much faster than defining a range of parameter values and initialising from scratch every time!



Source: BifurcationKit.jl documentation

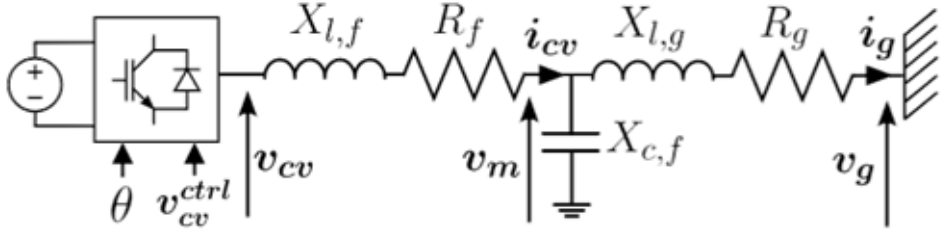
After solving for the initial equilibrium point $\boldsymbol{x_0}$:

- Predictor:
  - Jacobian → linear estimate of how $\boldsymbol{x}$ will change for small changes of $\boldsymbol{p}$
- Corrector:
  - Newton refines the guess

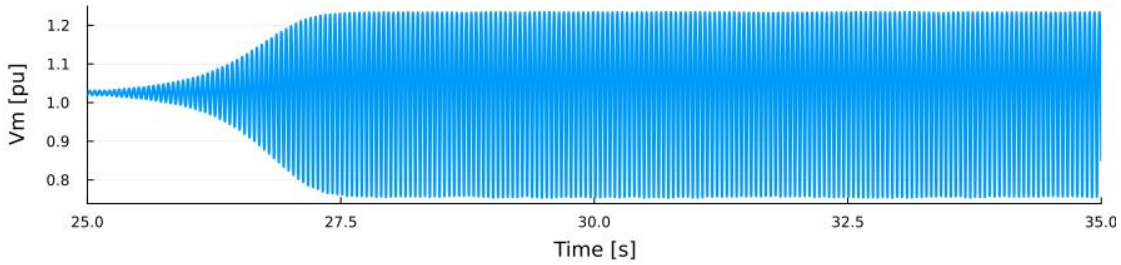
Actual algorithms are a little more sophisticated and robust, especially when continuing past saddle nodes, e.g., pseudo arclength continuation (PALC)

Case study: grid-forming converter strong grid instability

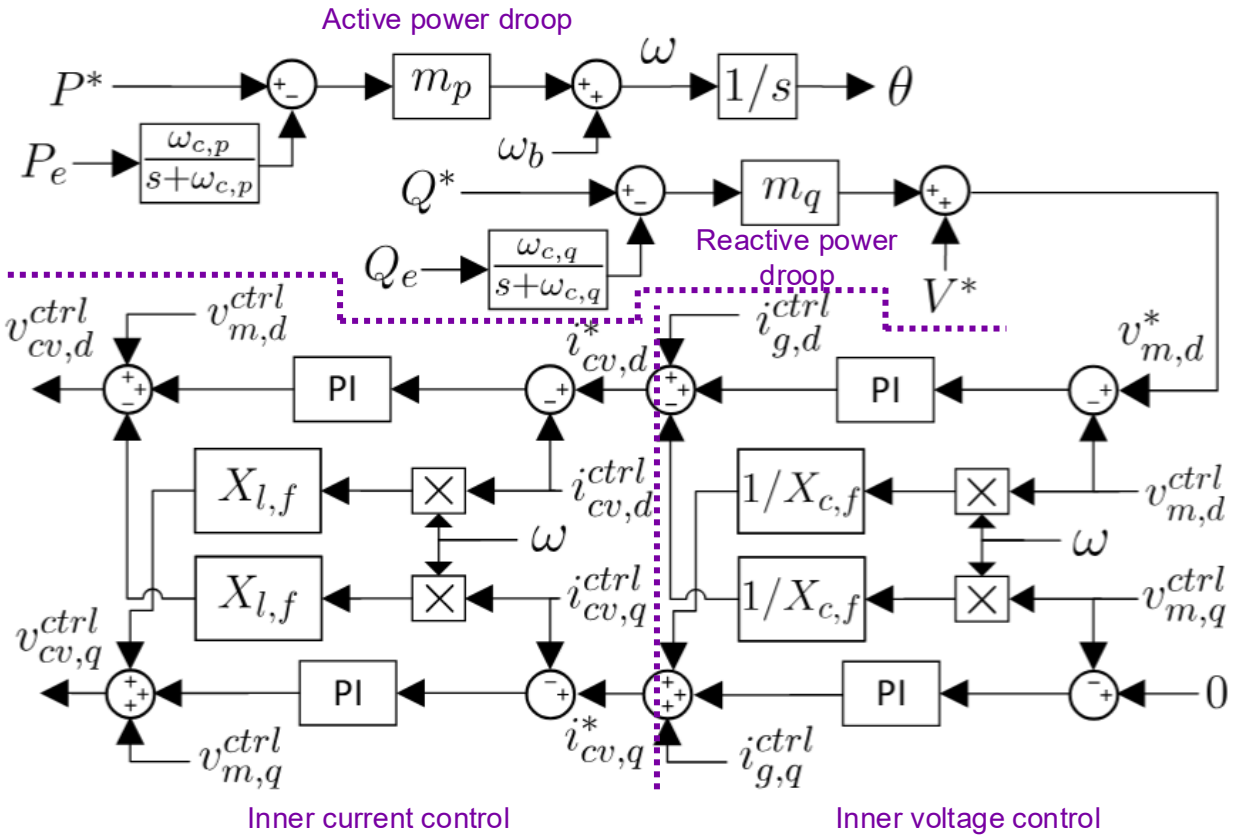
- Single-machine infinite bus system



- Sub-synchronous oscillation (≈ 20 Hz in this case)



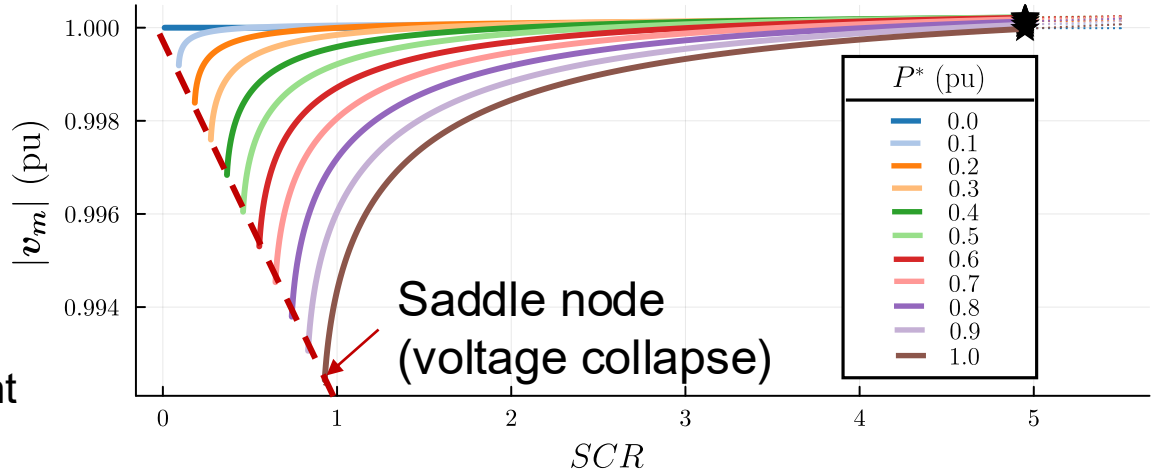
- Droop-based grid-forming converter



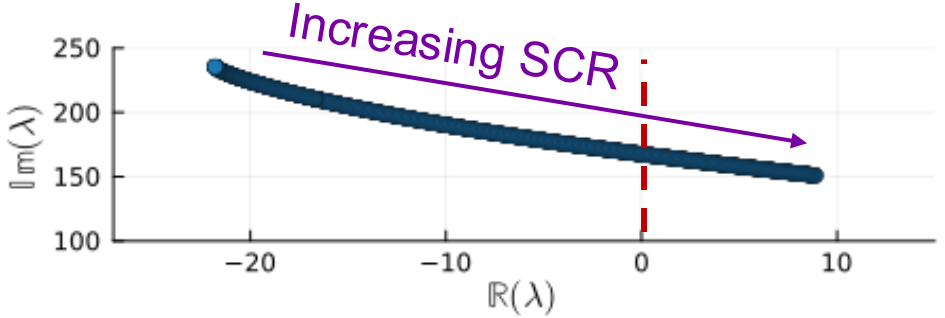
Case study: grid-forming converter strong grid instability

- Continuation parameter: short circuit ratio

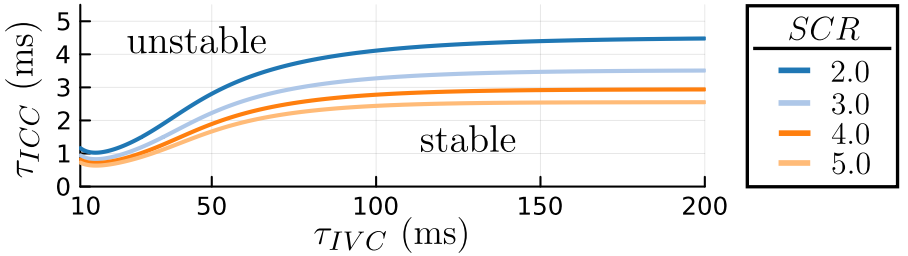
$$SCR = 1/\sqrt{\left(R_g^2 + X_{l,g}^2\right)}$$ for Thevenin equivalent with $V = 1$ pu



- Strong grid Hopf bifurcation
  - Participation factors tell us the corresponding oscillation is related to inner voltage and current controllers of grid-forming converter



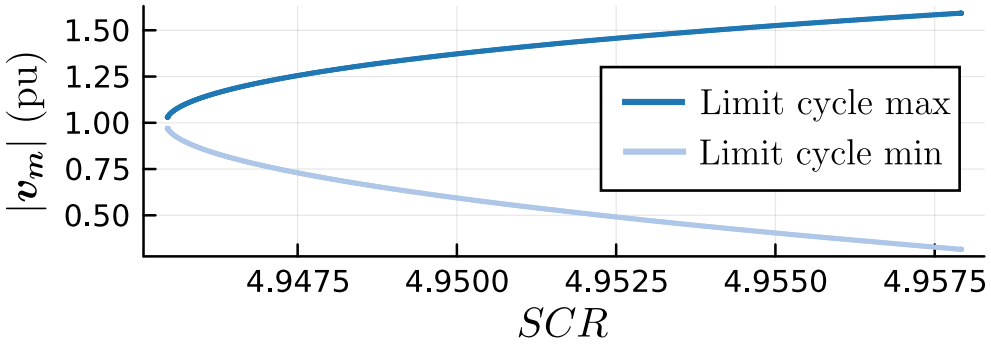
- We can continue the Hopf point as 2 parameters vary
  - Gives stability boundary (provided no other bifurcation is present in the study region)



τ_{ICC} = inner current controller 5% settling time,
τ_{IVC} = inner voltage controller 5% settling time

Case study: continuation of the limit cycle (oscillation)

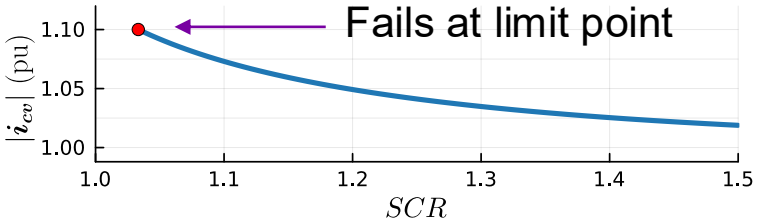
- The strong grid Hopf is supercritical
  - This can be determined using the normal form and centre manifold theory (out of scope)
- Stable limit cycle emerges
- We can continue the limit cycle to observe its behaviour



- Even for very small (further) increase of SCR, the oscillation grows rapidly

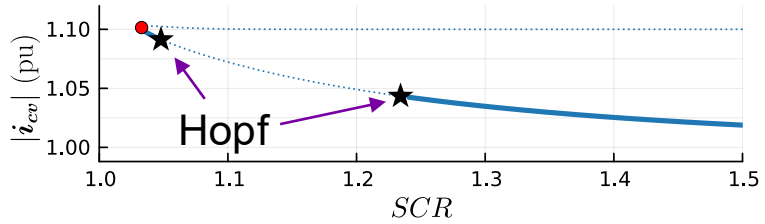
Current limiting and non-smooth dynamics

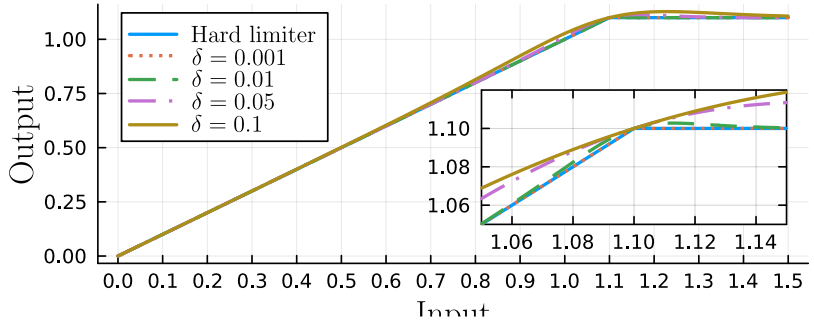
- Non-smooth elements such as current limiters are not immediately compatible with standard continuation tools
  - Smooth approximations

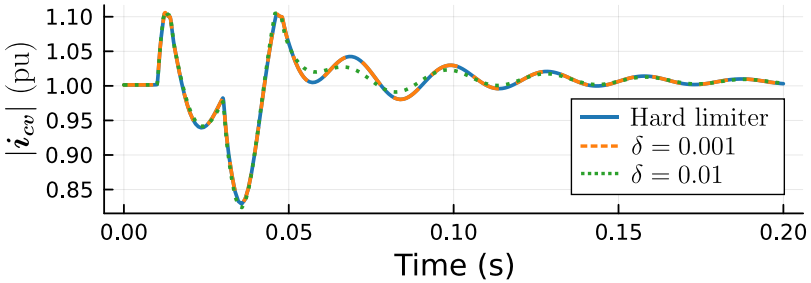


Smooth approximation enables continuation past limit point

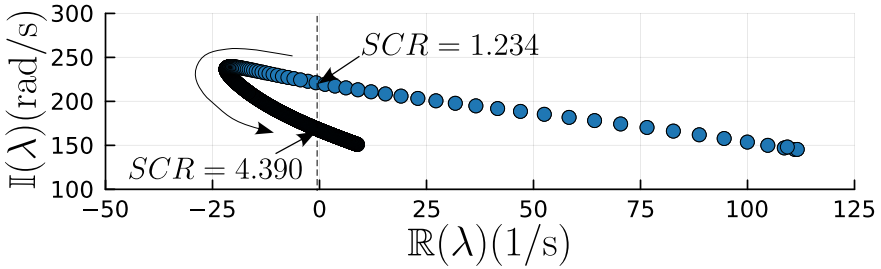
But introduces spurious Hopf bifurcation(s) at low SCR







For this example, a back-calculation-based anti-windup control is also included



Quick Summary

- Power system dynamics are complex, even more challenging with converter connected units
  - Especially in large systems with several units
- New phenomena with unknown behaviours
- But also complex effect of converter dynamics on “traditional” stability aspects/metrics

MANCHESTER 1824
The University of Manchester

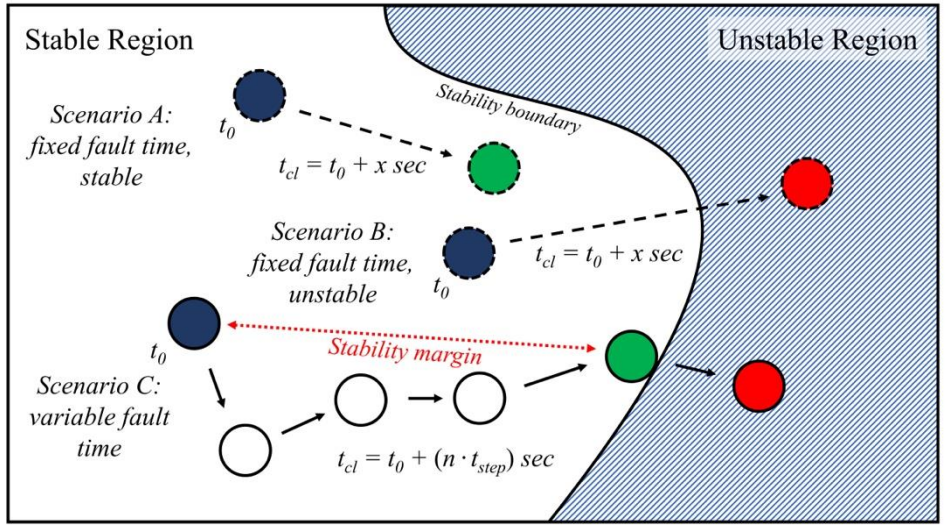
EXPLAINABLE MACHINE LEARNING AND POWER SYSTEM DYNAMICS

Gaining trust, insights and improving situational awareness

Stability assessment under increasing complexity and uncertainty

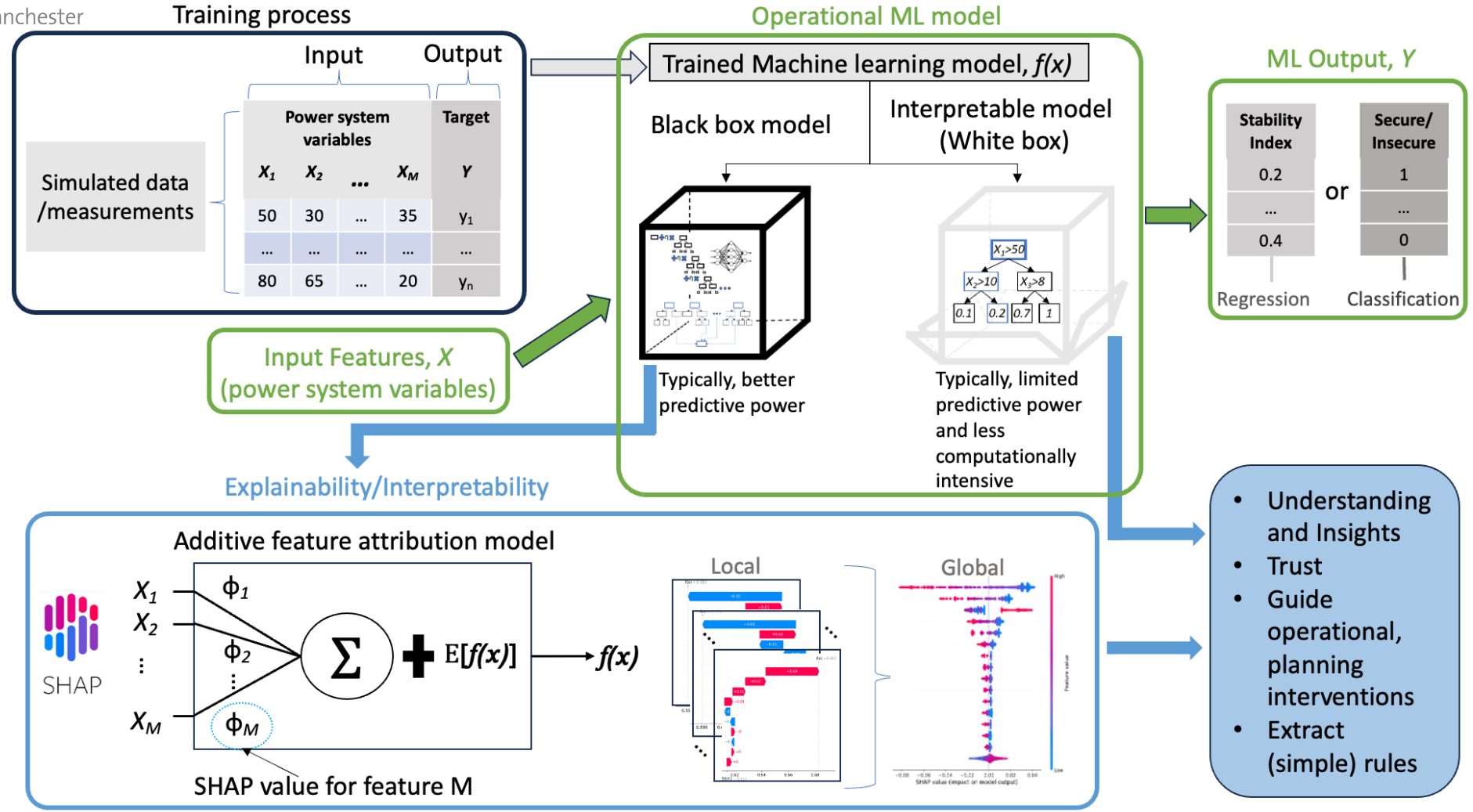
Complex phenomena, not easy to identify clear trends

- Time-domain simulation
  - Linear/nonlinear/hybrid dynamics
  - RMS: models the system dynamic components with differential-algebraic equations, solving iteratively in time
  - EMT: captures additional transient details
- Analytical methods
  - Less time consuming, very good insights but requires simplifications
- Explainable Machine Learning
  - Fast and accurate, but often ‘black boxes’ – tradeoff between accuracy/interpretability
  - Has the potential to offer speed-up of assessment and also useful insights/understanding through explainability/interpretability
- Transient/frequency stability but also other stability aspects
  - Locational aspects are becoming increasingly important



Going beyond the notion that ML is just a powerful black-box predictor

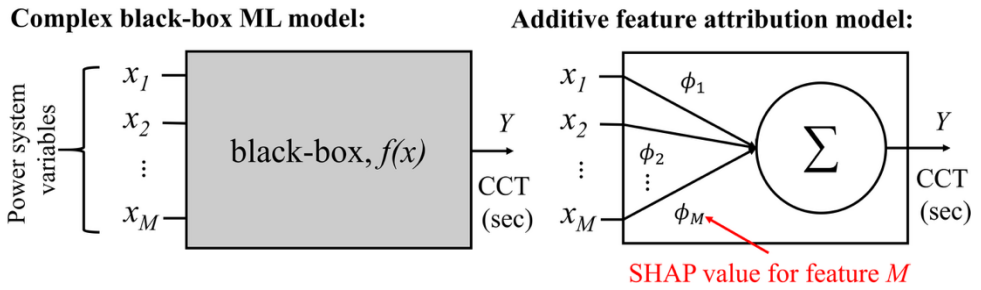
The University of Manchester



[1] P. N. Papadopoulos, S. Chatzivasileiadis, A. Marot, "Can Machine Learning Help Keep the System Secure?," accepted in IEEE Power and Energy Magazine.
[2] R. I. Hamilton and P. N. Papadopoulos, "Using SHAP Values and Machine Learning to Understand Trends in the Transient Stability Limit," in *IEEE Transactions on Power Systems*, doi: 10.1109/TPWRS.2023.3248941.
[3] R. I. Hamilton, J. Stiasny, T. Ahmad, S. Chevalier, R. Nellikkath, I. Murzakhanov, S. Chatzivasileiadis, P. N. Papadopoulos , "Interpretable Machine Learning for Power Systems: Establishing Confidence in SHapley Additive exPlanations", ICLR 2024 workshop on climate change and arXiv, 2022. doi: 10.48550/ARXIV.2209.05793. (https://arxiv.org/abs/2209.05793)
[4] R. I. Hamilton, P. N. Papadopoulos, W. Bukhsh and K. Bell, "Identification of Important Locational, Physical and Economic Dimensions in Power System Transient Stability Margin Estimation," in *IEEE Transactions on Sustainable Energy*, vol. 13, no. 2, pp. 1135-1146, April 2022, doi: 10.1109/TSTE.2022.3153843.

Shapley Additive Explanations (SHAP)

- SHAP builds a simpler (linear) explanation model of the original (non-linear) black-box model (model agnistic)
  - Uses (approximate) Shapley values defined as the average marginal contribution of a feature to all feature coalitions with that feature
  - *Provides feature <u>effects</u>*
  - <u>*Local*</u> (for one operating scenario) and <u>*global*</u> (for all operating scenarios) explanations
- Reveals tendencies (no guarantee of causal relations)



i's Shapley value

S's weight

i's marginal contribution

$$\varphi_i(v) = \sum_{S \subseteq D \setminus \{i\}} \frac{|S|!\,(|D| - |S| - 1)!}{|D|!} \left(v(S \cup \{i\}) - v(S)\right)$$

i works with group S

i does NOT work with group S

All possible subsets S to which *i* does not belong

- How is this concept extended to explain machine learning model predictions?
  - Focus on explaining the prediction $f(x)$ for a single sample x
- How do we remove the absent features?
  - Baseline values
  - Distributional values (Marginal/conditional Shapley values)
- How do we calculate Shapley values?
  - Approximation or assumption based

Shapley value concept for explaining ML model predictions

The University of Manchester

Goal: Explain the prediction $f(x^{ie})$ for a single instance $\boldsymbol{x^{ie}}$ by attributing contributions to all feature values.

1. The mismatch

ML model

$f\colon \mathbb{R}^{|D|} \to \mathbb{R}$

Takes vectors as inputs (all features)

$\not\equiv$

Coalitional game

$v\colon 2^{|D|} \to \mathbb{R}$

Takes sets of players as inputs

2. The key challenge & idea

We cannot remove features (input size is fixed).

We simulate 'absence' using a reference mechanism.

$$v_{x^{ie}}(S) = f(h_{x^{ie}}(S))$$

3. Hybrid input $h_{x^{ie}}(S)$

$$\forall \text{ feature } j\colon h_{x^{ie}}(S)_j \begin{cases} x_j^{ie}, & j \in S \\ \text{reference value}, & j \notin S \end{cases}$$

Keep features in S at their observed values x^{ie}; replace absent features (not in S) with a reference.

4. Two ways to choose the reference for absent features

Baseline values

(deterministic)

$$h^{base}_{x^{ie}}(S)_j \begin{cases} x_j^{ie}, & j \in S \\ x_j^b, & j \notin S \end{cases}$$

Reference: *fixed baseline* $x^b \in \mathbb{R}^{|D|}$

Distributional values

(probabilistic)

$$h^{dist}_{x^{ie}}(S)_j \begin{cases} x_j^{ie}, & j \in S \\ X_j \sim P_j^*, & j \notin S \end{cases}$$

Reference: *draw absent features from a distribution* P_j^*

Takeaway: By constructing $v_{x^{ie}}(S) = f(h_{x^{ie}}(S))$ via a reference mechanism we turn the ML model into a set function over feature value subsets.

Defining absent features with a baseline value

The University of Manchester

Training dataset
Original dataset with all features

x_1	x_2	x_3	x_4
1	5	6	1
4	8	1	4
3	7	6	5
1	5	4	7
4	9	1	8

Instance being explained
Row whose prediction is explained

$x^{ie} =$

x_1	x_2	x_3	x_4
1	5	6	1

2^D Possible subsets

$2^4 = 16$

Subset S	
x_1 x_2 x_3 x_4	$S:=D$ †
x_1 x_2 x_3	$S:=\{x_1, x_2, x_3\}$
x_1 x_2	$S:=\{x_1, x_2\}$
⋮	
	$S:=\{\emptyset\}$ †

† $v(D) = f(x^{ie})$

Present features (in S)

$S:=$
$\{x_1, x_2\}$

Absent features (not in S)

$D \backslash S =$
$\{x_3, x_4\}$

† $v(\emptyset) = f(x^b)$

Baseline
Fixed reference (here: all zeros)

$x^b =$

x_1	x_2	x_3	x_4
0	0	0	0

Corresponding 'game'

$v_{x^{ie}}(S) = f(h_{x^{ie}}(S))$

$h_{x^{ie}}(S) \sim$

x_1	x_2	x_3	x_4
1	5	0	0

The model is evaluated with the observed values x^{ie} for present features in S, while absent features (not in S) are **replaced** according to the baseline value x^b

Baseline choices & trade-offs

- An all-zeros baseline
- An average across features
- A baseline drawn from a uniform distribution

✓ Simple to implement

✗ Baseline heavily influences attributions; no neutral baseline, especially for tabular data

✗ Off-manifold evaluations — $h_{x^{ie}}(S)$ may lie outside the data distribution

Defining absent features with distributional values

Training dataset
Original dataset with all features

x_1	x_2	x_3	x_4
1	5	6	1
4	8	1	4
3	7	6	5
1	5	4	7
4	9	1	8

Instance being explained
Row whose prediction is explained

$x^{ie} =$

x_1	x_2	x_3	x_4
1	5	6	1

2^D Possible subsets
$2^4 = 16$

Subset S	
x_1 x_2 x_3 x_4	$S:=D$ †
x_1 x_2 x_3	$S:=\{x_1, x_2, x_3\}$
x_1 x_2	$S:=\{x_1, x_2\}$
⋮	
	$S:=\{\emptyset\}$ †

† $v(D) = f(x^{ie})$

Present features (in S)

$S:=$
$\{x_1, x_2\}$

Absent features (not in S)

$D \backslash S =$
$\{x_3, x_4\}$

† $v(\emptyset) = E[f(X)]$

'Games': $v_{x^{ie}}(S) = E\left[f(h^{dist}_{x^{ie}}(S))\right]$ — Absent feature values drawn from distribution P_J^*

Marginal — Interventional

Absent feature values sampled independently from present
$[P_J^*: P(X_{D\backslash S})]$

$h_{x^{ie}}(S) \sim$

x_1	x_2
1	5

X

x_3	x_4
6	1
1	4
6	5
4	7
1	8

Present feature values

independence

$$v(S) = E\left[f(\mathrm{X}) \middle| do\left(X_S = x_S^{ie}\right)\right]$$

Conditional — Observational

Absent features sampled given present features
$[P_J^*: P(X_{D\backslash S} \mid X_S = x_S^{ie})]$

$h_{x^{ie}}(S) \sim$

x_1	x_2	x_3	x_4
1	5	6	1
1	5	4	7

$$v(S) = E\left[f(\mathrm{X}) \middle| X_S = x_S^{ie}\right]$$

Marginal vs Conditional SHAP

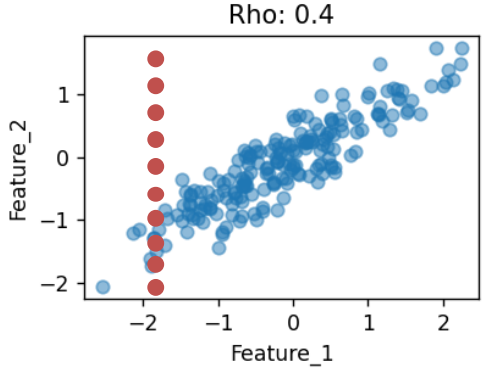
The University of Manchester

Marginal — Interventional

Absent features sampled independently of the present ones.

✗ Unrealistic instances — off-manifold evaluations

In power systems → may violate physical laws (e.g., OPF)



✓ Only gives credit to the features used by the model

$$\varphi_1 \neq 0$$
$$\varphi_2 = 0$$

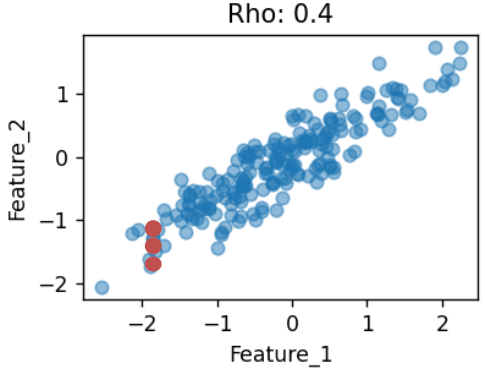
Appropriate for understanding the functional form of the model f.

Conditional — Observational

Absent features sampled given the present ones.

✓ Respects the data distribution — on-manifold evaluations

In power systems → faithful to feasible operating points



✗ Spreads credit among *correlated* features even if some of them are irrelevant to the model

$$\varphi_1 \neq 0$$
$$\varphi_2 \neq 0$$

Appropriate for understanding the underlying mechanism in data.

If model: $f(x_1, x_2) = x_1$
and x_1, x_2: correlated

Not for corrective actions	Only as good as the model	Approximation accuracy	Computational cost
SHAP shows associations, not interventions (re-dispatching SG9 is informed by, but not prescribed by, a SHAP value).	Confidently-wrong explanations from poor models (hence the Sobol validation).	Generic estimators are approximate; the choice of estimator matters.	Exact Shapley is exponential in features; practical SHAP needs a smart estimator.

MANCHESTER 1824
The University of Manchester

SHAP inherits fairness axioms

1 Efficiency

Per-feature attributions sum to the prediction's deviation from the mean.

The SHAP values for all features sum exactly to the gap between the instance's prediction and the expected prediction.

$$\sum_i \varphi_i(v) = v(D) - v(\emptyset) = f(x^{ie}) - E[f(X)] \quad \dagger$$

Attributions are on the scale of the output — directly readable as contributions to the prediction.

2 Symmetry

Equal influence, equal attribution. Feature order is irrelevant.

Two features that influence the prediction identically across every coalition receive the same SHAP value.

$$\text{If } v(S \cup \{i\}) = v(S \cup \{k\}) \text{ for all } S \subseteq D\backslash\{i,k\}, \text{ then } \boldsymbol{\varphi}_i = \boldsymbol{\varphi}_k$$

Symmetry ensures correct interpretation of SHAP values order, eg, when ranking features using SHAP importance (sum of absolute SHAP values per feature).

3 Dummy

Features that do not influence the prediction receive zero SHAP value.

A feature whose addition never changes the model's prediction, in any coalition, receives a SHAP value of zero.

$$\text{If } v(S \cup \{i\}) = v(S) \text{ for all } S \subseteq D\backslash\{i\}, \text{ then } \boldsymbol{\varphi}_i = 0$$

No spurious attribution to features the model never uses.

4 Additivity

Decompose the model, decompose the explanation

If a model is the sum of two component models, each feature's SHAP value is the sum of its values in each component.

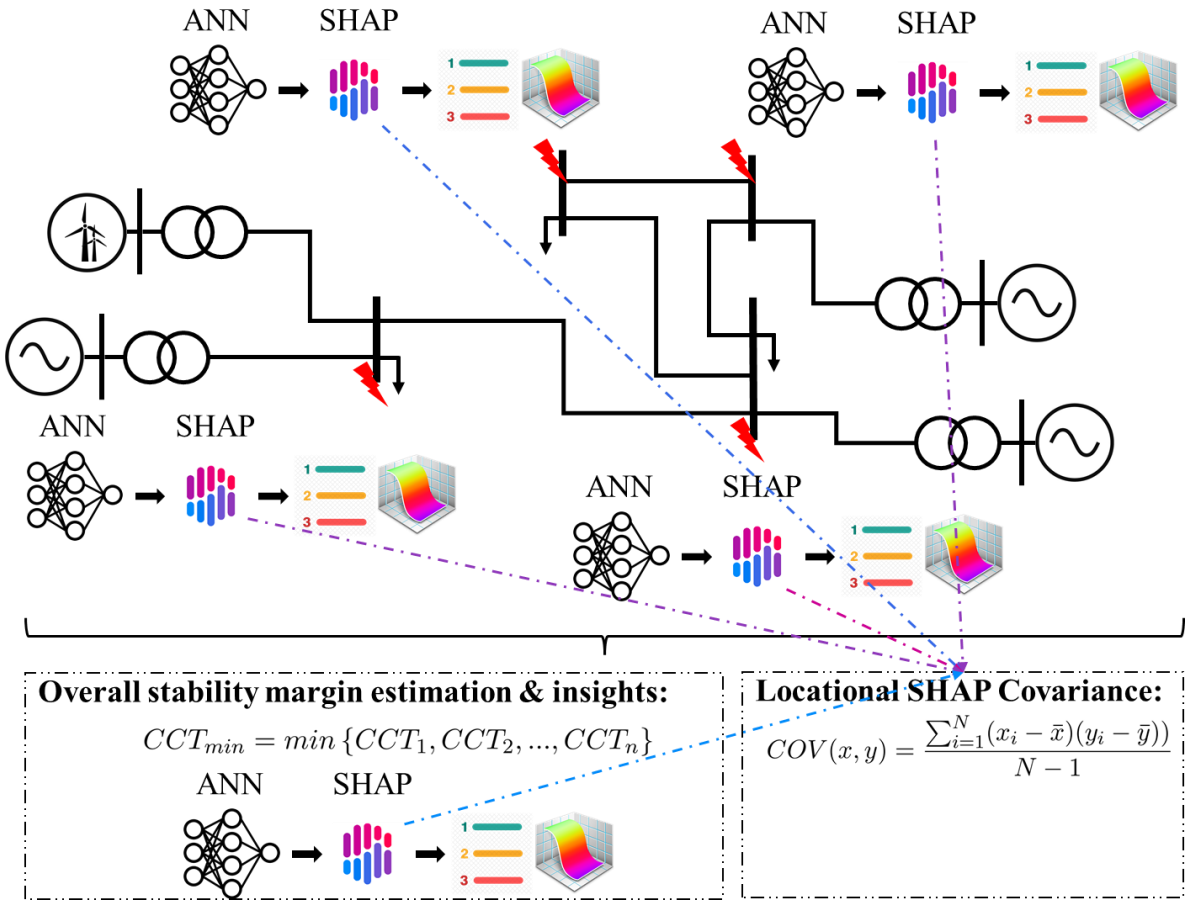
$$\varphi_i(v_1 + v_2) = \varphi_i(v_1) + \varphi_i(v_2)$$

For additive ensembles (eg, XGBoost), SHAP values are computed per tree and summed across the model. This is a key reason Tree SHAP is tractable.

★ **Because $v_{x^{ie}}(S) = f(h_{x^{ie}}(S))$ defines a valid coalitional game on the feature set, all four Shapley axioms transfer to SHAP by construction** *(Lundberg & Lee, 2017; Štrumbelj & Kononenko, 2014).*

† ***Convention:*** *we follow Lundberg & Lee with $v(\emptyset) = E[f(X)]$. Equivalently (Molnar, 2023), one may centre the value function $v(S) := f(h_{x^{ie}}(S)) - E[f(X)]$, giving $v(\emptyset) = 0$. Both conventions yield identical φ_j.*

Explainable ML (SHAP) for transient stability

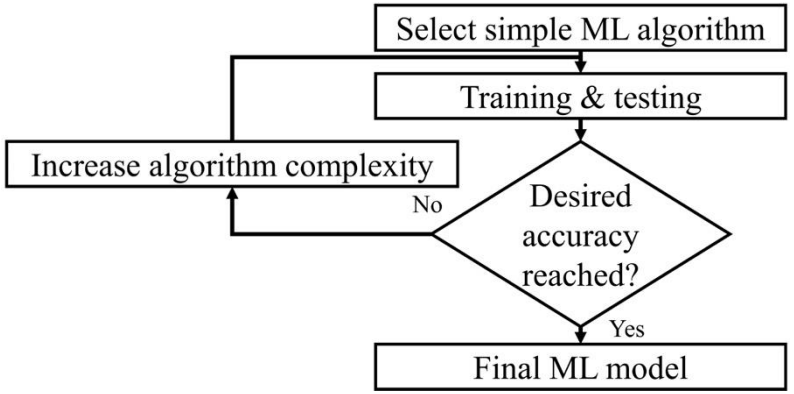


- Locational ML models for CCT estimation
  - Focus on accuracy and minimization of maximum over-/under- estimation
- Covariance between features and SHAP values can reveal locational aspects

R. I. Hamilton and P. N. Papadopoulos, "Using SHAP Values and Machine Learning to Understand Trends in the Transient Stability Limit," in *IEEE Transactions on Power Systems* ,doi: 10.1109/TPWRS.2023.3248941.

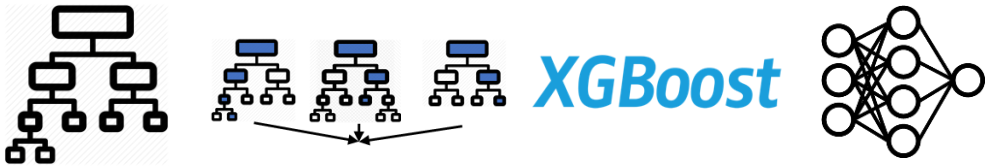
Locational Accuracy

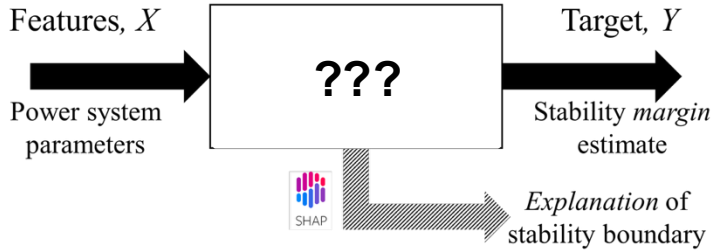
- Mostly motivated by improvement in maximum errors



IMPACT OF ML ALGORITHM ON ACCURACY METRICS FOR ALL LOCATIONS

Performance Metric	DT	RF	XGBoost	ANN
Average RSQ				
Average MSE (sec^2)				
Average RMSE (sec)				
Max MOE (sec)				
Min MUE (sec)				
Max MOE < 0.3 (sec)				
Min MUE < 0.3 (sec)				

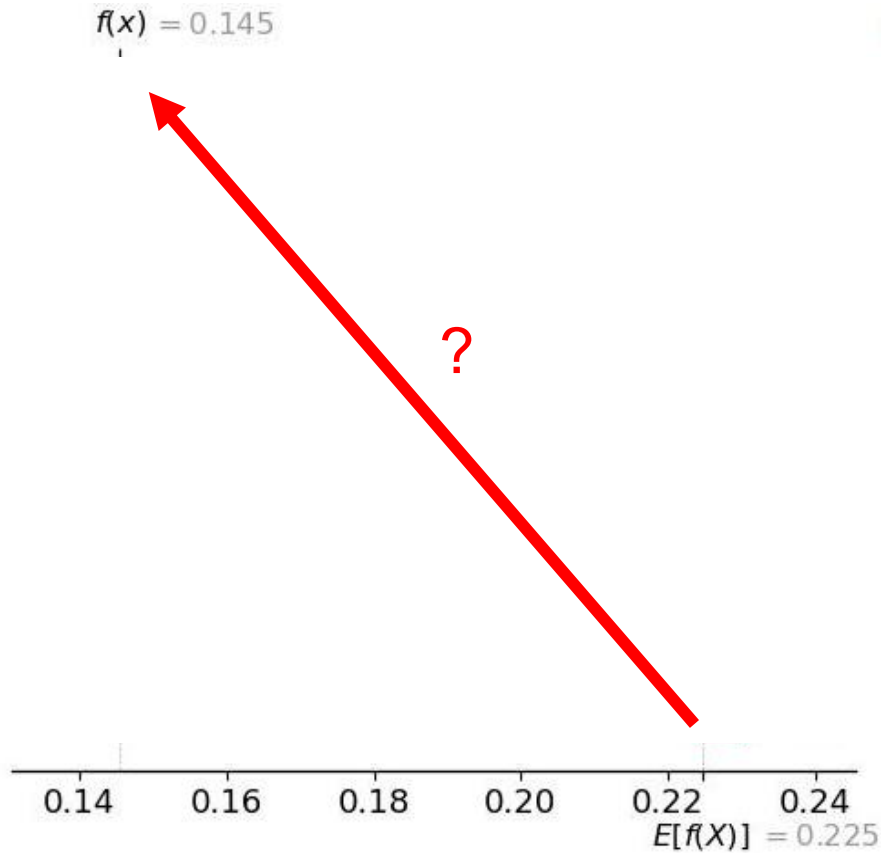




MANCHESTER 1824
The University of Manchester

SHAP – local explanation

Prediction, $f(x)$ (sec)



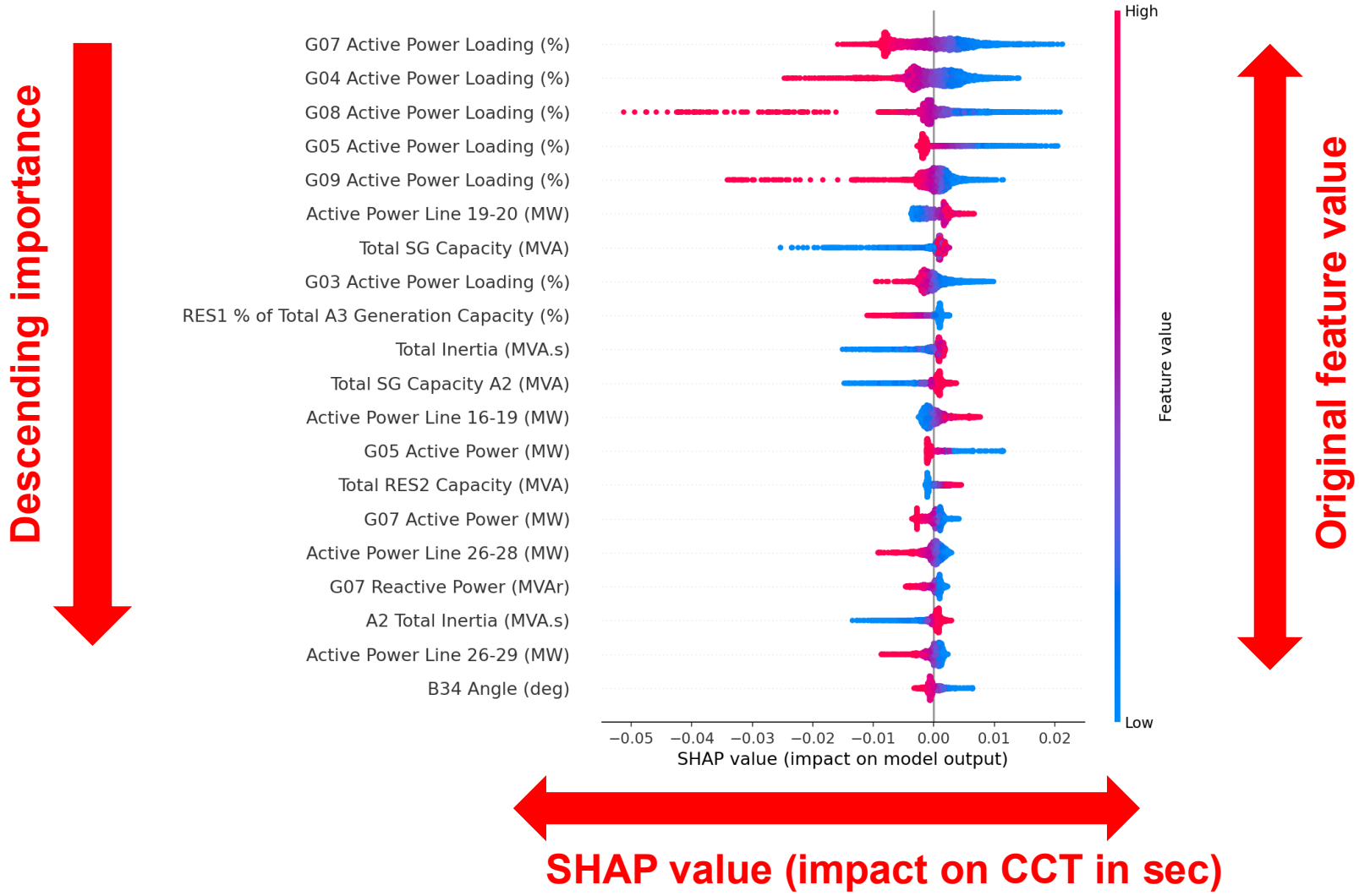
SHAP values, ϕ_i (sec)

Expectation, $E[f(x)]$ (sec)

$$f(x) = E[f(x)] + \sum_{i=1}^{M} \phi_i$$

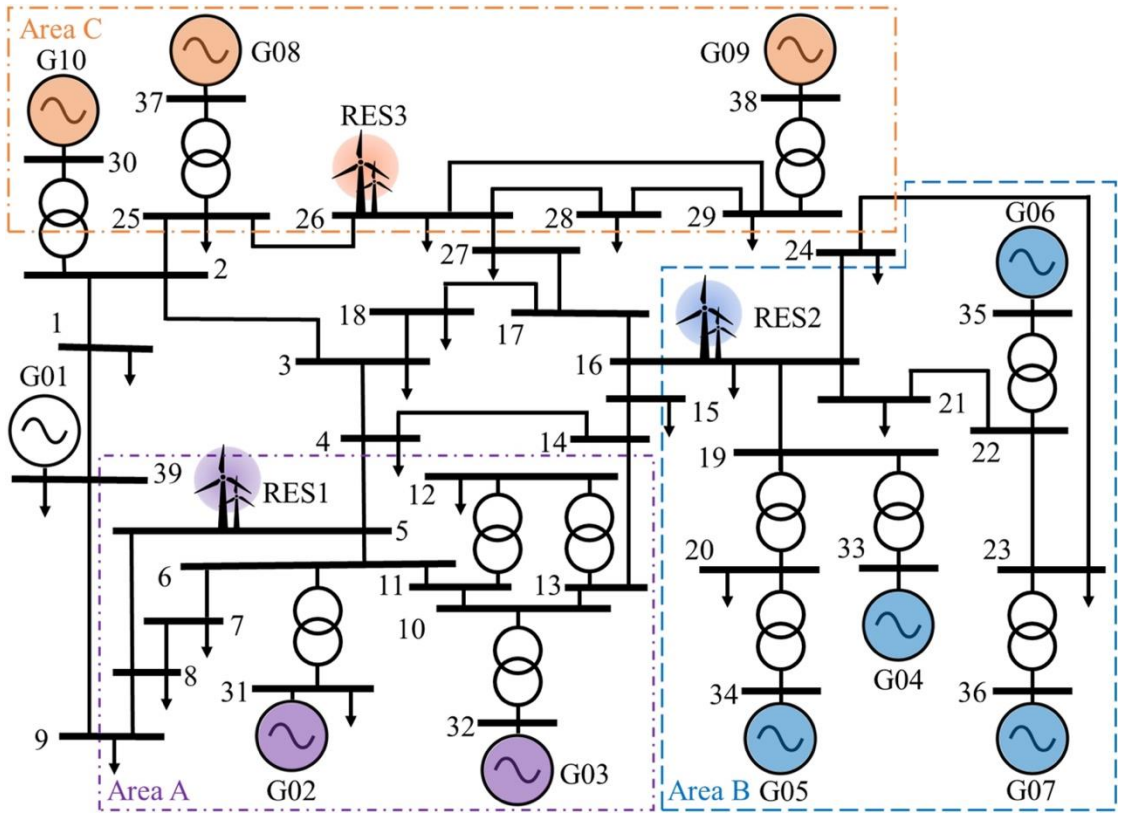
MANCHESTER 1824
The University of Manchester

SHAP – global explanation



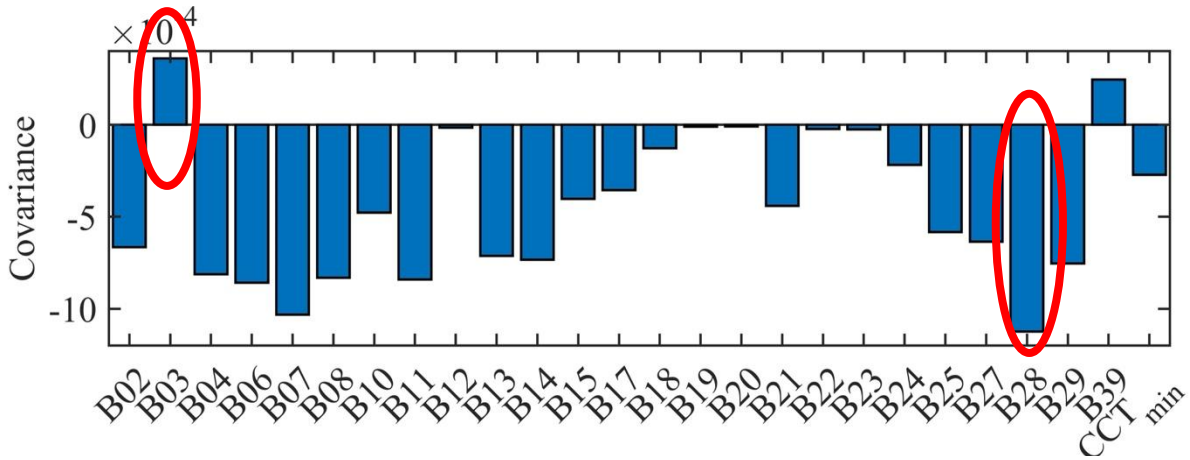
MANCHESTER 1824
The University of Manchester

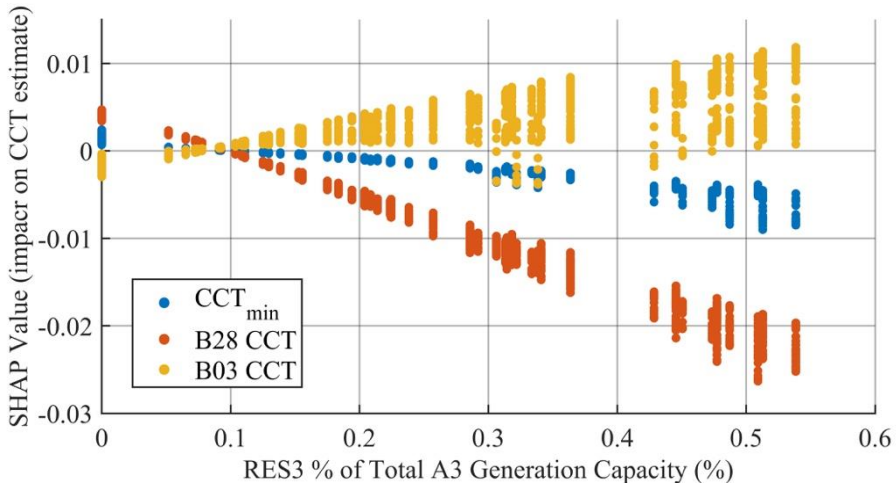
SHAP and locational aspects

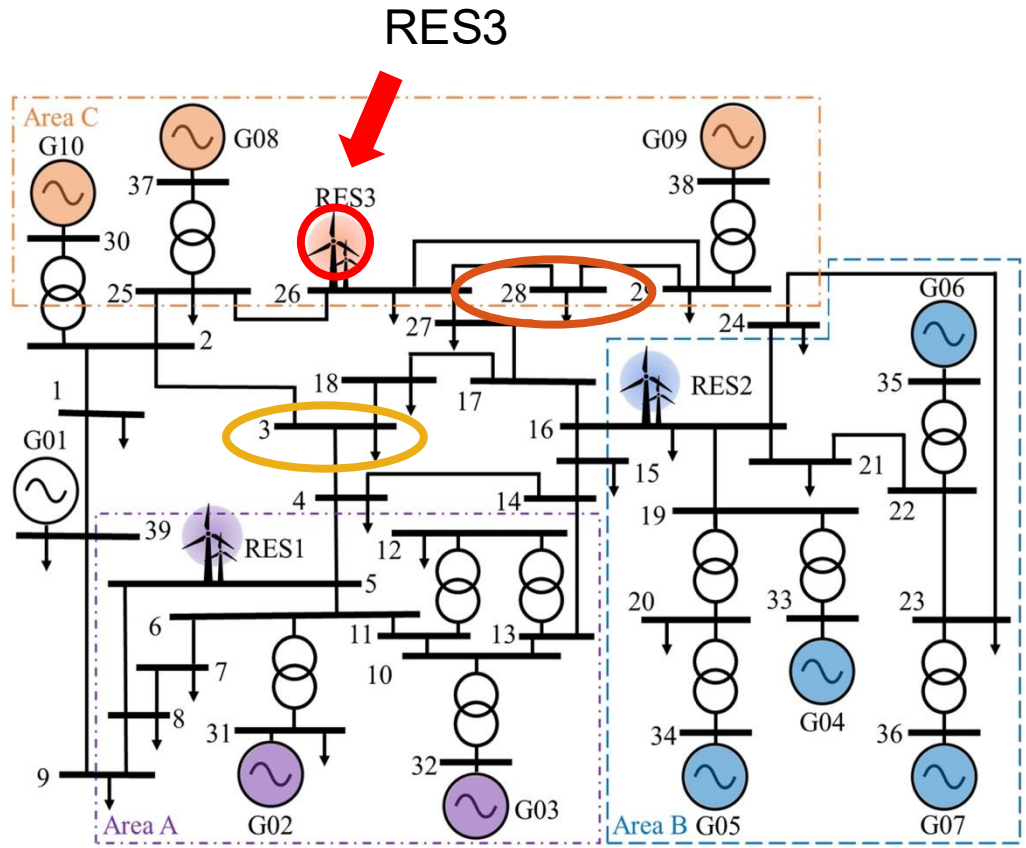


Example result of locational explanaibality

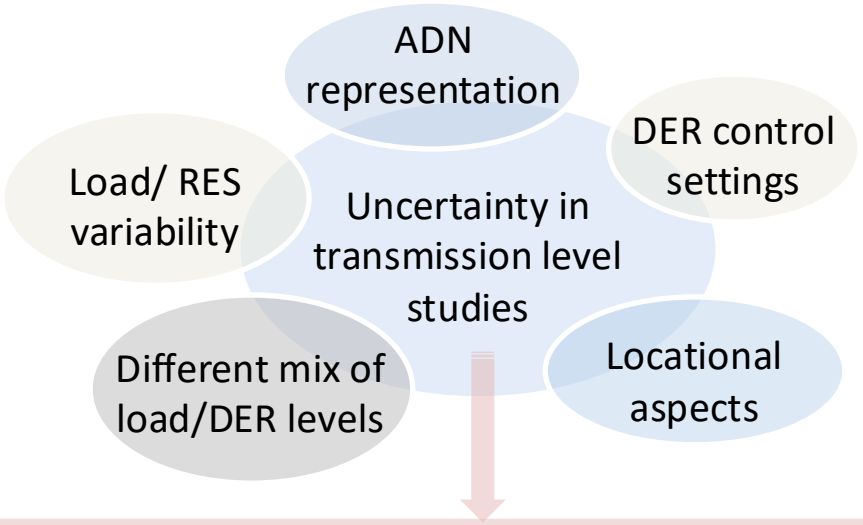
- Covariance between features and SHAP values can reveal locational aspects
- Increase in wind generation in area 3 (RES3) causes local increase of CCT in Bus 3 but decrease in Bus 28





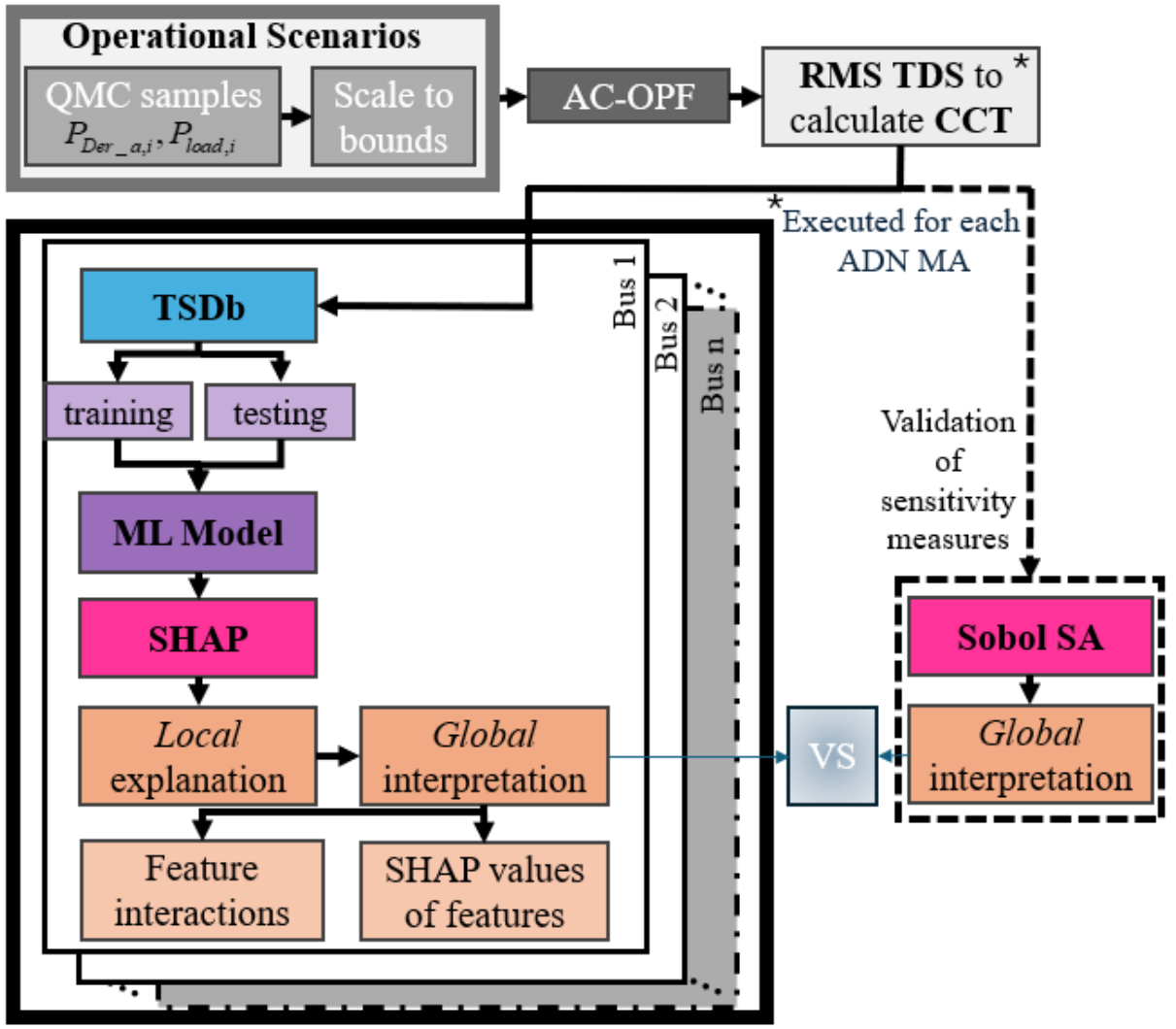


ML and Explainability (SHAP) for Active Distribution Networks



Increasing need for **efficient sensitivity analysis frameworks** that can handle non-linear complex systems characterized by many interacting variables

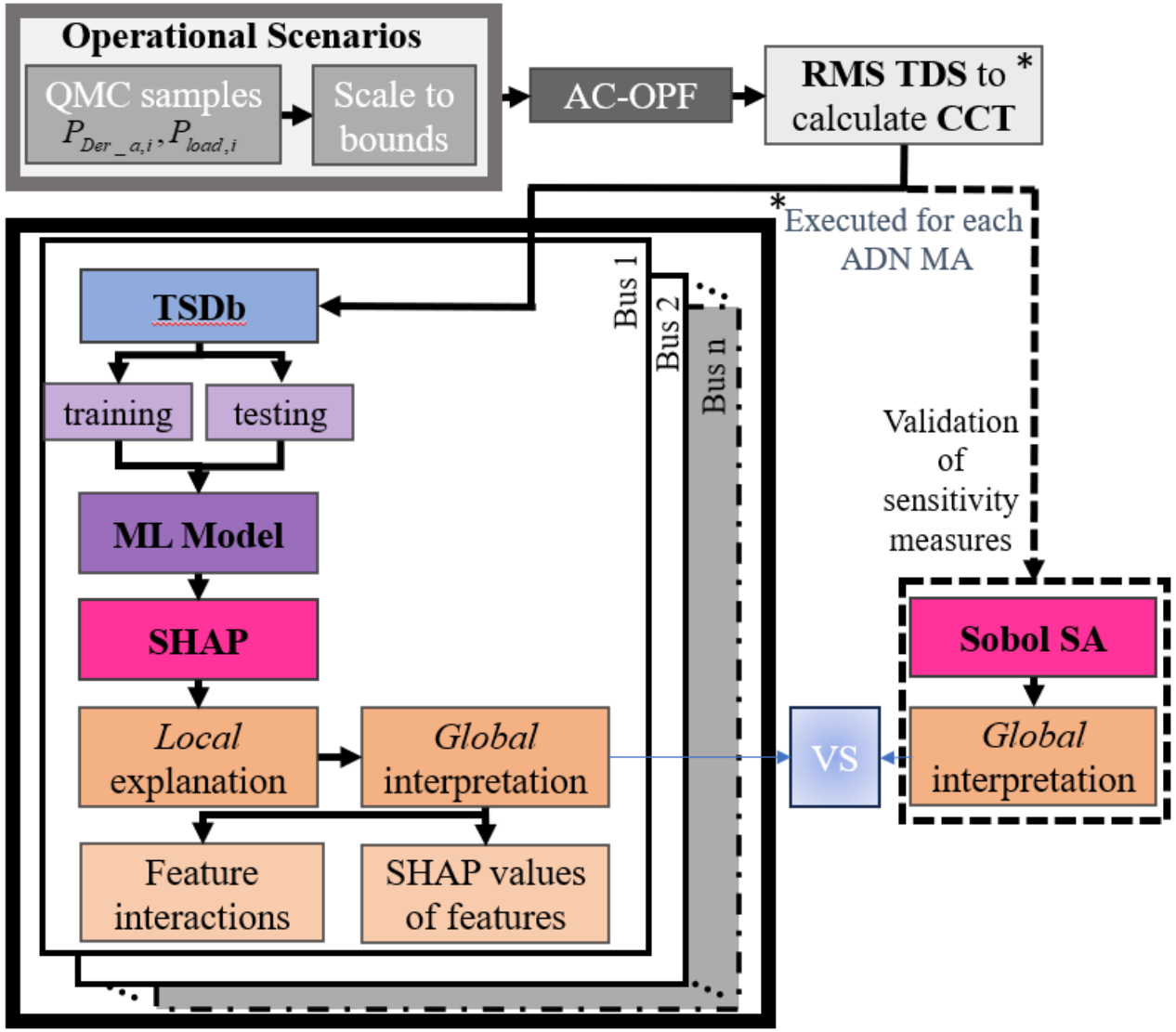
- How different modelling approaches can affect Critical Clearing Times across the network?



I. S. Lamprianidou, R. I. Hamilton and P. N. Papadopoulos, "Impact of Active Distribution Network Modeling Approaches on Transient Stability Through Sensitivity Analysis," in IEEE Transactions on Power Systems, vol. 41, no. 1, pp. 572-584, Jan. 2026, doi: 10.1109/TPWRS.2025.3589751.

Impact of Active Distribution Network Modelling Assumptions under operational uncertainties

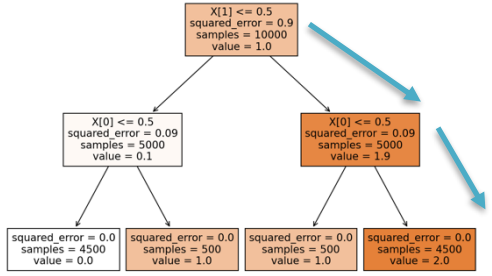
Test System & Case study design	
System	Modified 39-Bus system – 10 ADN locations
Operational sampling	Quasi-MC (Sobol); efficient coverage
Loads/DERs	$P_{load} \in [0.6, 1.3]$ pu / $P_{DER} \in [0.01, 1.1]$ pu
SG inertia	100 % / 75 % / 50 % / 25 % of S_{rated}
ADN MAs	11 MAs: static→dynamic representations (Net Load→der_a + Load)
DER control variations	voltage ride-through settings, tripping time, dynamic voltage support
Stability index	CCT for self-clearing 3-φ faults (binary search)
ML model	XGBoost; Bayesian-optimised hyper parameters



Explainer: Path-Dependent Tree SHAP

Estimation strategy	Tree (path-dependent traversal)
Removal approach	Conditional
Removal variant	Empirical: uses tree's split structure
Bias-free	No: depth-position bias for correlated features
Variance-free	Yes : iterate through all possible subsets (exact)
Computational time	$O(TLD^2)$→polynomial

[T=number of trees, L=number of leaves, D=maximum depth of any tree]



Conditional

$$v(S) = E\left[f(\mathrm{X}) \middle| X_S = x_S^{ie}\right]$$

***Approximated** empirically by weighted descents through the tree, using each split's training-sample frequencies as conditional probabilities.*

Why Path-Dependent Tree SHAP?

- Faithfull to original data distributions – **preserves physical constraints in power systems**
- Doesn't give credit to irrelevant features or spread importance among correlated features (unlike true observational)
- Deterministic: no MC sampling or background dataset needed
- Efficient for large-scale tree-based models (like XGBoost)

Caveat — depth-position bias

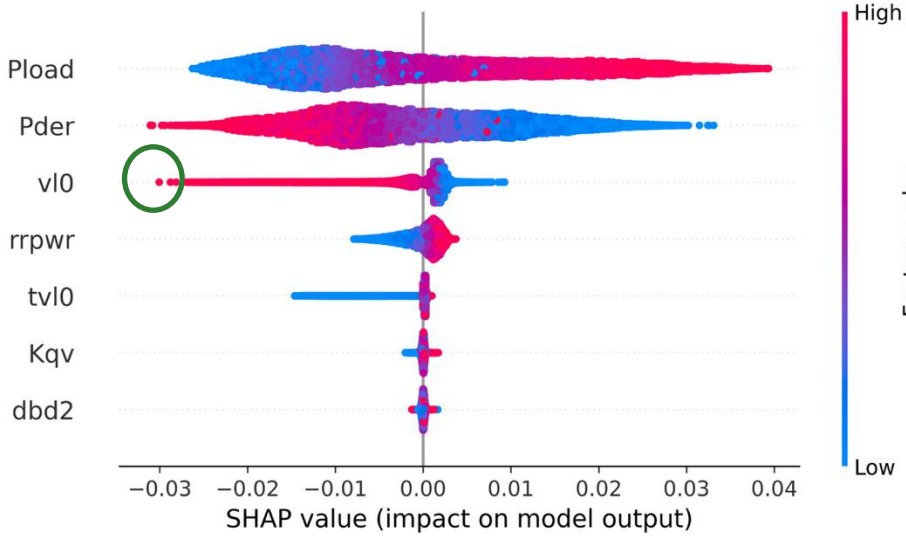
Among **correlated features**, the algorithm gives more credit to features splitting near the root and penalizes those near the leaves — a consequence of inconsistent traversal weighting when an upstream feature is present vs. absent in the evaluated subset.

Mitigation: *feature engineering — among each correlated pair keep only the most physically meaningful variable (Pearson screen + domain-driven merging, e.g., from the pair Pload-Qload→ keep only Pload).*

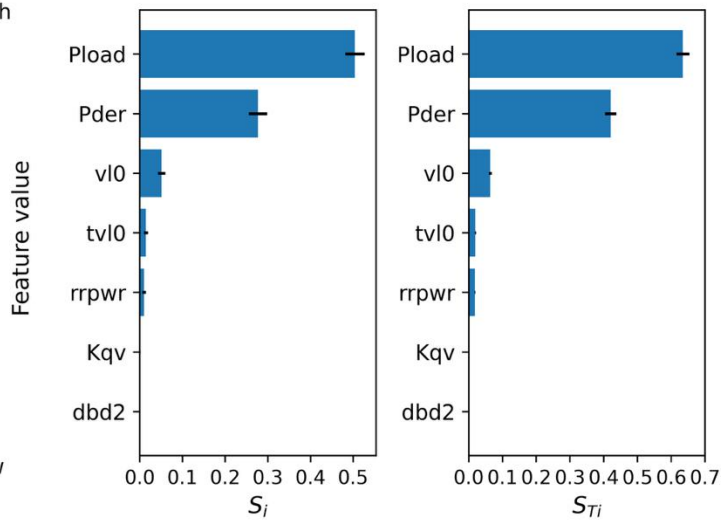
SHAP as sensitivity analysis tool -comparison with Sobol

Controlled study: D=7 independent factors · sampled scenarios growing as T=N(D+2) · faults on Bus 28. Identical operating data feed both methods.

SHAP



Sobol



99 % *simulation-budget reduction*

Sobol convergence: **73 728** TDS ≈ **21.9 days**

SHAP convergence: **512** TDS ≈ **0.152 days**

Confidence intervals as percentages of key indices

T	S_{Pload}	S_{Pder}	S_{vl0}	S_{TPload}	S_{TPder}	S_{Tvl0}
4608	±17%	±26%	±67%	±14%	±17%	±23%
73728	±3.9%	±6.4%	±15%	±3.1%	±4%	±6%

Agreement

Both methods place Pload, Pder, vl0 as the dominant drivers. Minor reordering only among low-impact controls.

SHAP provides effects

SHAP reveals the SIGN of the effect: high Pload → +CCT, high Pder → −CCT (not visible in variance-only SA).

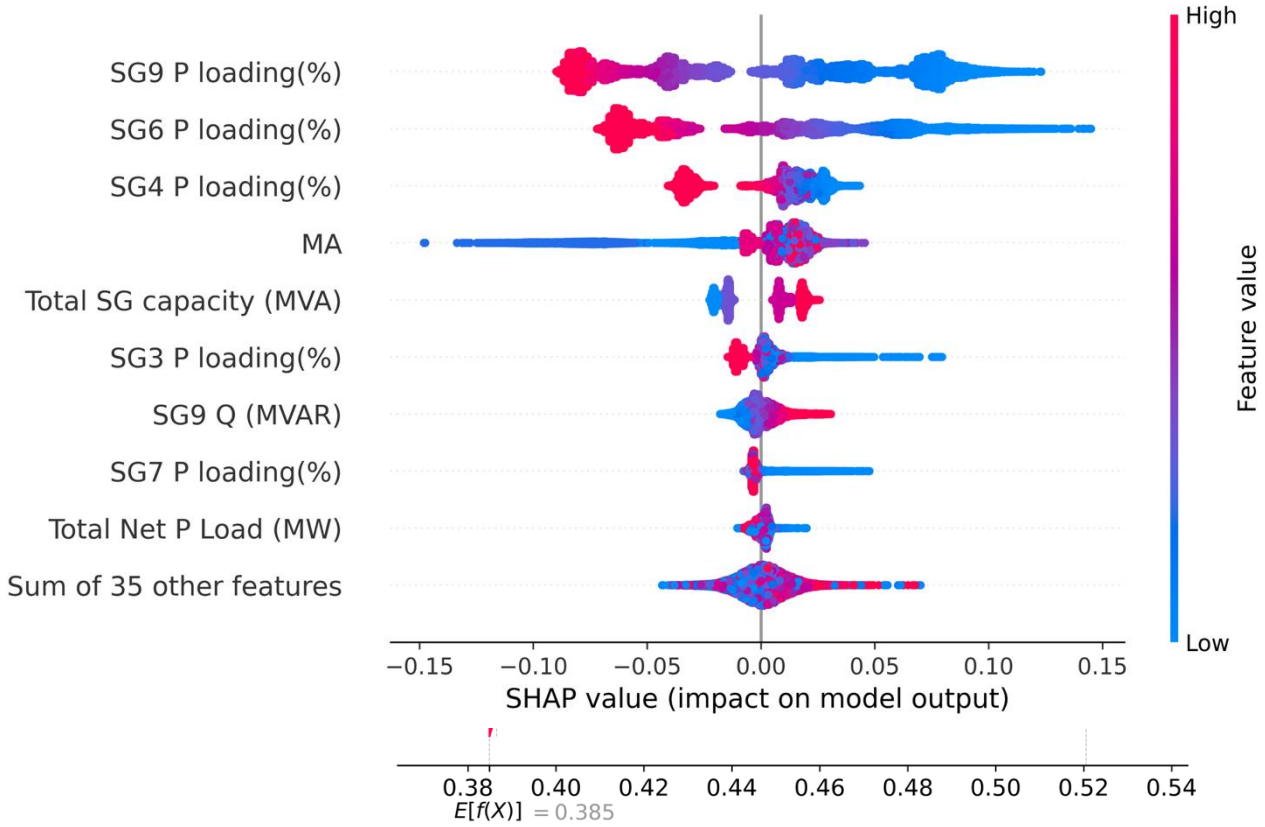
SHAP - local and global

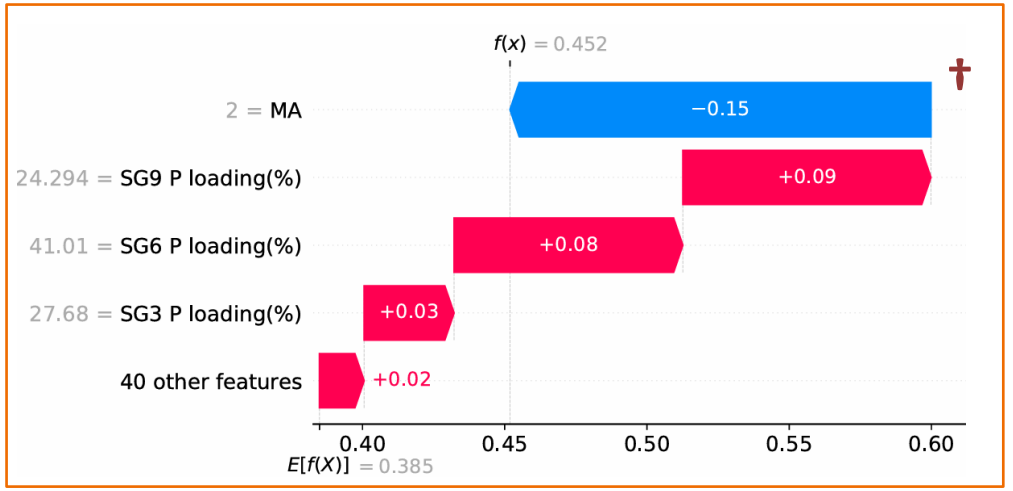
Per instance sensitivities in the natural units of the target (CCT)

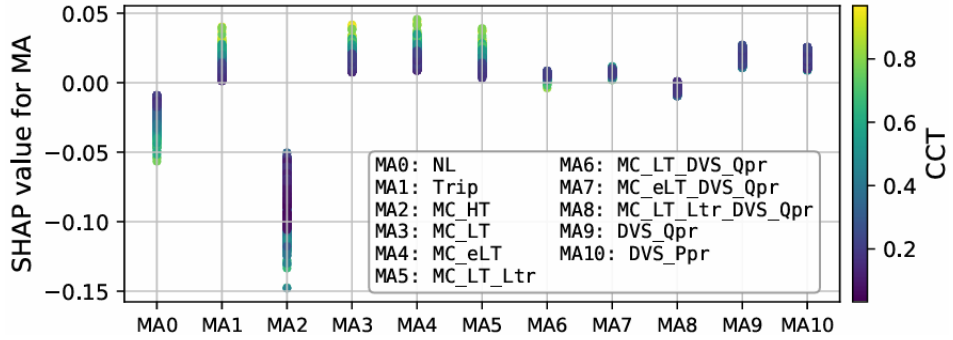
Surface rare but impactful effects

Impact of ADN modelling approach

44 features · 3 754 operating points · 11 ADN modeling approaches · faults on Bus 29 (closest SG: SG9).







MA0 Net-Load
Missing DER dynamics → misleading stability estimates.

MA2 legacy MC-HT †
High-voltage momentary cessation cancels SG6 + SG9 contribution: −150 ms in CCT.

MA3–5 broader ride-through
Voltage threshold is the most influential control — flips many cases unstable → stable.

DVS schemes
Help, but effectiveness depends on DER location relative to fault & critical SG (CSG).

Practical aspects and scalability

Two-phase workflow for operators

PHASE 1 · Risk prioritisation

ML scans **1024 forecast scenarios** in seconds.

Flags **44 high-risk cases** (CCT < 100 ms) — **≈ 4 %**

Local SHAP confirms **DER MA** as a recurrent risk driver, even when settings are not visible to the TSO.

PHASE 2 · Engineering validation & mitigation

Full RMS TDS run only on the **44 flagged cases**, under conservative MA2 settings.

Mitigation explored offline: SG re-dispatch, SG9 maintenance reschedule, DER re-parameterisation.

Order-of-magnitude reduction in compute, complementing — not replacing — standard contingency analysis.

Scalability stress test

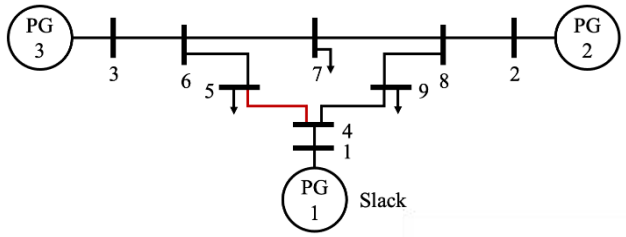
2 000-bus system · 149 features · 2 048 scenarios

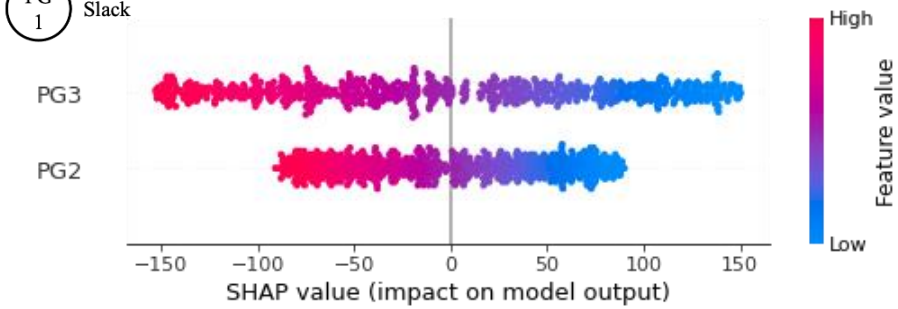
5.9 h	10.1 m	2.25 s
RMS TDS (10 cores, 3 s window)	XGBoost training	SHAP inference (entire test set)

Bottleneck = simulation. *SHAP cost is essentially free relative to the dynamic-simulation budget — and the ML+SHAP layer is independent of system size: parallelising TDSs, reducing the simulation window scales the framework further.*

“Learning” system physics through explainable machine learning

- Application to DC power flow
  - Train individual XGBoost models to predict each line flow
  - Input features: PG2 and PG3, output: specific line flow
- SHAP values are equivalent to Power Transfer Distribution Factor (PTDF)
  - Sensitivity of line flows to power injections
  - Learning physical behaviour purely from the trained ML model





SHAP values features Mean value of training set

$$\phi_i(f, \boldsymbol{x}) = \boldsymbol{w}_i(\boldsymbol{x}_i - \mathbb{E}[\boldsymbol{X}_i]),\ i \neq 0 \rightarrow \frac{\partial \phi_i}{\partial \boldsymbol{x}_i} = \boldsymbol{w}_i.$$

Relates the sensitivity of line flow to i^{th} injection, hence SHAP derivative is equivalent to PTDF

TABLE I
9-BUS 3-GENERATOR TEST NETWORK PTDF & SHAP.

	True physical PTDF, $\boldsymbol{D}$		SHAP-based PTDF, $\hat{\boldsymbol{D}}$	
Line	Bus 2	Bus 3	Bus 2	Bus 3
Line 1-4	-1.0000	-1.0000	-0.9999	-0.9999
Line 4-5	-0.3613	-0.6152	-0.3613	-0.6151
Line 5-6	-0.3613	-0.6152	-0.3613	-0.6151
Line 3-6	0	1.0000	0.0000	0.9999
Line 6-7	-0.3613	0.3848	-0.3613	0.3848
Line 7-8	-0.3613	0.3848	-0.3613	0.3848
Line 8-2	-1.0000	0	-1.0000	0.0000
Line 8-9	0.6387	0.3848	0.6386	0.3848
Line 9-4	0.6387	0.3848	0.6386	0.3848

R. I. Hamilton, J. Stiasny, T. Ahmad, S. Chevalier, R. Nellikkath, I. Murzakhanov, S. Chatzivasileiadis, P. N. Papadopoulos , ”Interpretable Machine Learning for Power Systems: Establishing Confidence in SHapley Additive exPlanations”, ICLR 2024 workshop on climate change and arXiv, 2022. doi: 10.48550/ARXIV.2209.05793. (https://arxiv.org/abs/2209.05793)

MANCHESTER 1824
The University of Manchester

BEYOND EXPLAINABILITY – EXTRACTING FUNCTIONAL APPROXIMATIONS FROM ML MODELS

Symbolic Regression and Kolmogorov-Arnold Networks

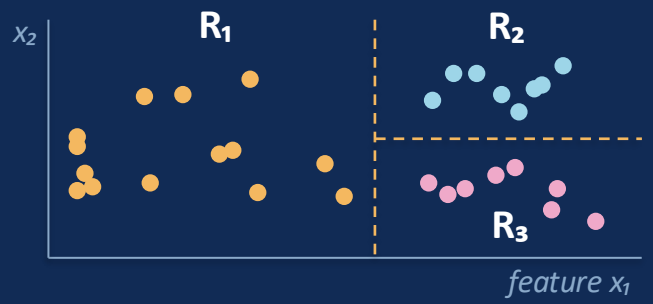
I. Lamprianidou, F. Fernandes, R. J. Bessa, and P. N. Papadopoulos, “Symbolic Explainer of Power System Dynamics”, PSCC 2026, Cyprus.

The University of Manchester

Piecewise Symbolic Regression for stability assessment

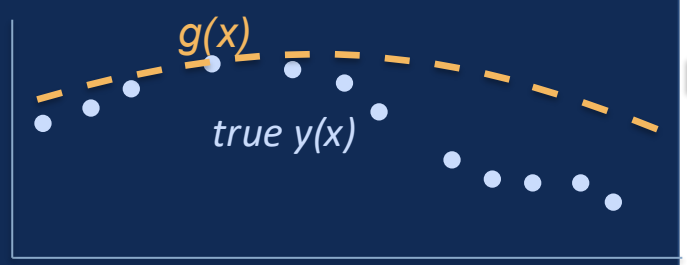
Application to calculating a Stability Metric (Critical Clearing Time – CCT in this case).

Partitioning of the input space



Operating space split into meaningful regimes/regions

Functional approximations



Symbolic regression · compact interpretable surrogates

Analytic understanding

```
if  SGKE9 > 8.1 :
  CCT  =  0.012 (δ5 · DP25)
    + (δ1 + 1.33)/(δ9+SGP9)
else  if  ...
```

Regime-conditioned equations revealing system dynamics

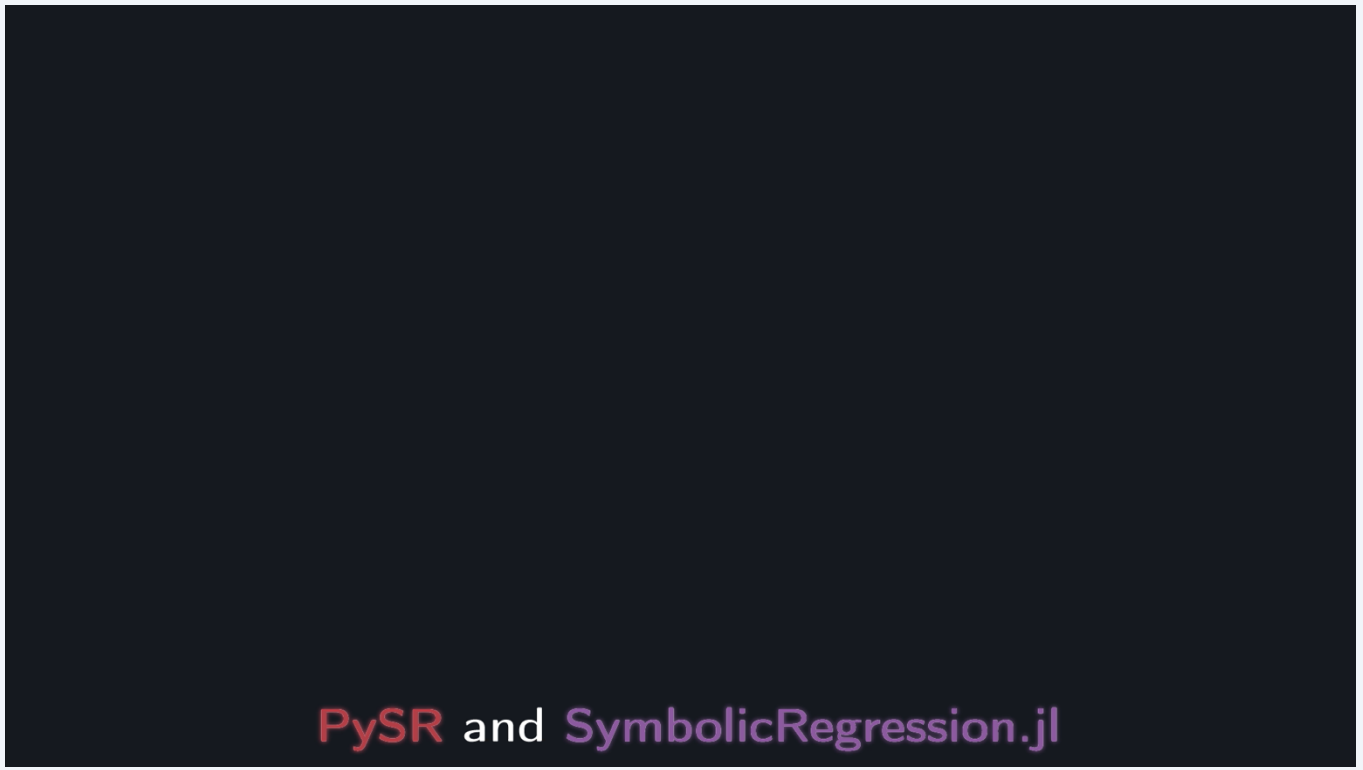
Symbolic regression — searching the space of equations

Given data (x_i, y_i) *find a closed-form expression* ***g(x)*** *that approximates the mapping while staying compact.*

- Equations represented as expression trees
- Evolutionary / Pareto search over operators (+, −, ×, ÷, sin, √…)
- Selects candidates trading off accuracy against complexity
- Implemented with PySR [1] (evolutionary algorithm based)

Outcome: not just black-box predictions but equations a human can read.

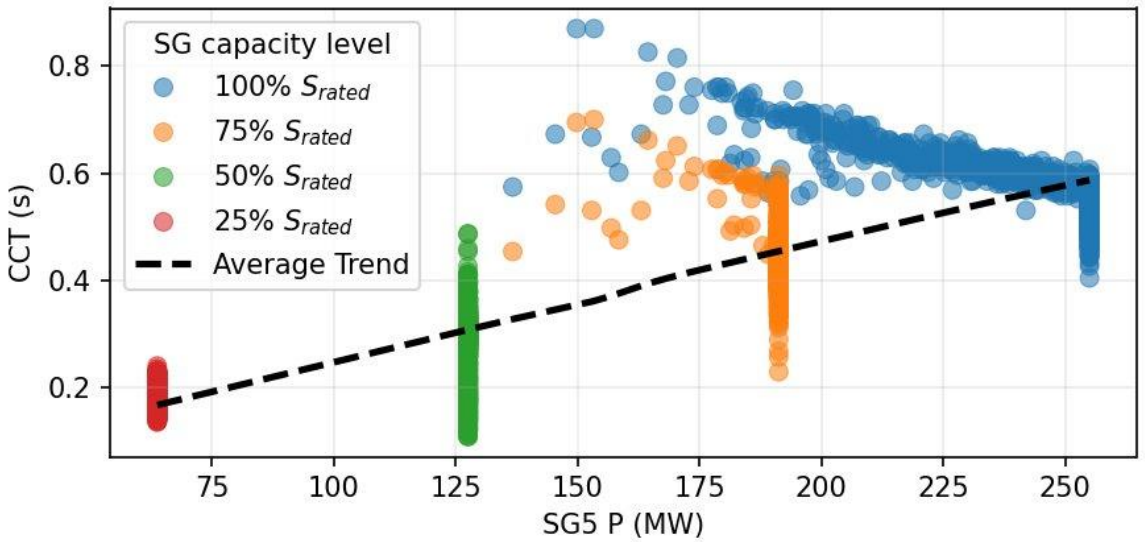
Illustration of the symbolic-regression search process



[1] M. Cranmer, “Interpretable machine learning for science with PySR and SymbolicRegression.jl,” arXiv:2305.01582, 2023. (and https://github.com/MilesCranmer/PySR)

Power-system dynamic behaviour can be piecewise

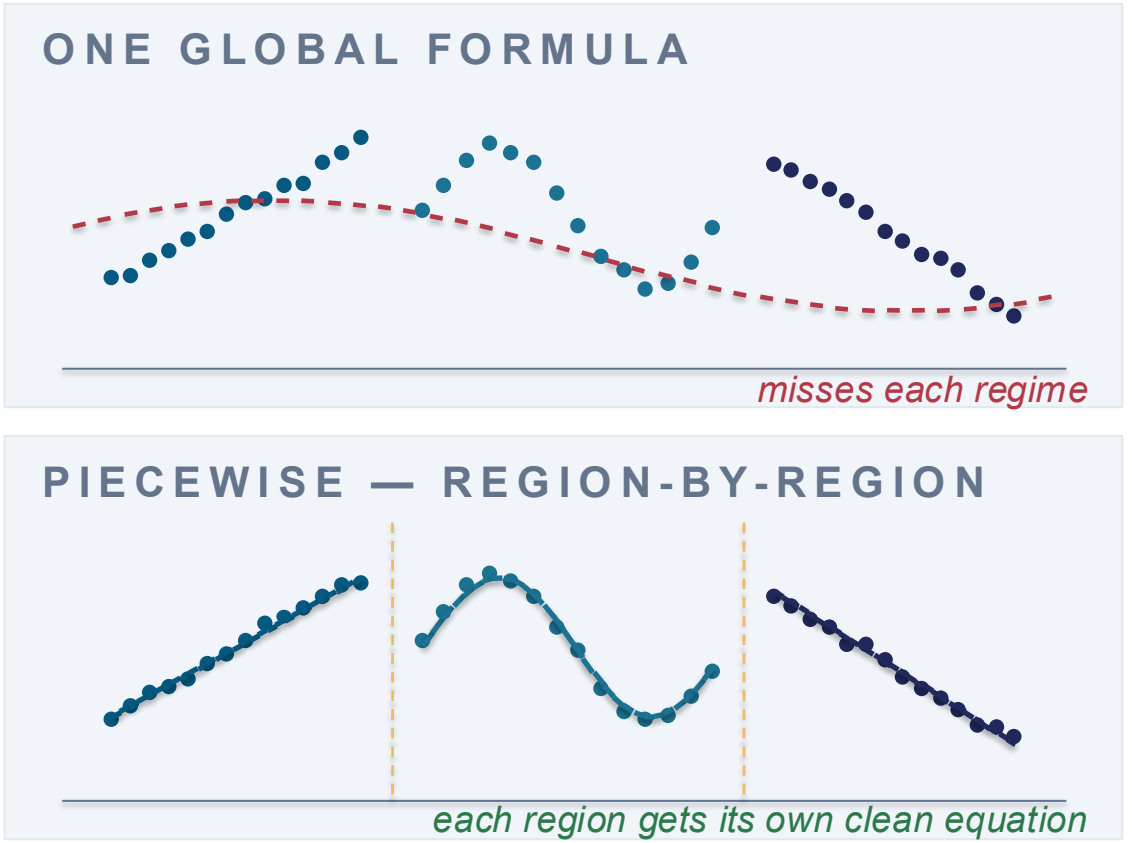
Stability mechanisms shift abruptly with control modes and SG capacity. A single global formula either becomes dense and unreliable, or hides the regime change altogether.



The piecewise view

- Partition the operating space into homogeneous regions
- Fit a separate, compact symbolic model in each region
- Recover both the partition rules and the local equations
- Equations stay simple and faithful to local physics

Global SR vs. Piecewise SR



MANCHESTER 1824
The University of Manchester

What we want to learn

Approximate a potentially non-smooth mapping $f: \mathbb{R}^d \to \mathbb{R}$ (CCT) by a piecewise SR model (PcSR):

$$\widehat{f}_\alpha(x) = \sum_{\ell=1}^{L(\alpha)} \mathbf{1}\{x \in \mathcal{R}_\ell^\alpha\}\, g_\ell^\alpha(x), \qquad \mathbf{1}\{x \in \mathcal{R}_\ell^\alpha\} = \begin{cases} 1, & x \in \mathcal{R}_\ell^\alpha \\ 0, & x \notin \mathcal{R}_\ell^\alpha \end{cases}$$

- $x \in \mathbb{R}^d$: input feature vector (dispatch, inertia, angles, …)
- $\widehat{f}_\alpha$: the overall piecewise predictor at pruning level a
- a : cost-complexity pruning level of the regression tree
- $L(\alpha)$: number of leaves (regions) at pruning level a
- $\mathcal{R}_\ell^\alpha$: the ℓ-th region from that pruned tree
- g_ℓ^α: regional symbolic model trained on samples in $\mathcal{R}_\ell^\alpha$

Pc-SR jointly chooses (i) the partitions and (ii) the local symbolic models to balance accuracy with interpretability.

The University of Manchester

Region-specific symbolic regression for stability assessment

01 Task-aware partitioning

Cost-complexity-pruned Regression Decision Tree

A family of trees with rule-based regions along the CCP path.

02 Local symbolic models

Compact PySR regressors per leaf

Constrained operator set + tight complexity cap.

03 Selection

Multi-objective score

Pick the simplest partition with acceptable accuracy.

04 Merging (optional)

Post-hoc consolidation

Merge regions whose equations are equivalent.

Partition and equations are co-optimized yielding stable, physically-meaningful regimes with simple local equations.

Cost-complexity pruning (CCP) of regression trees

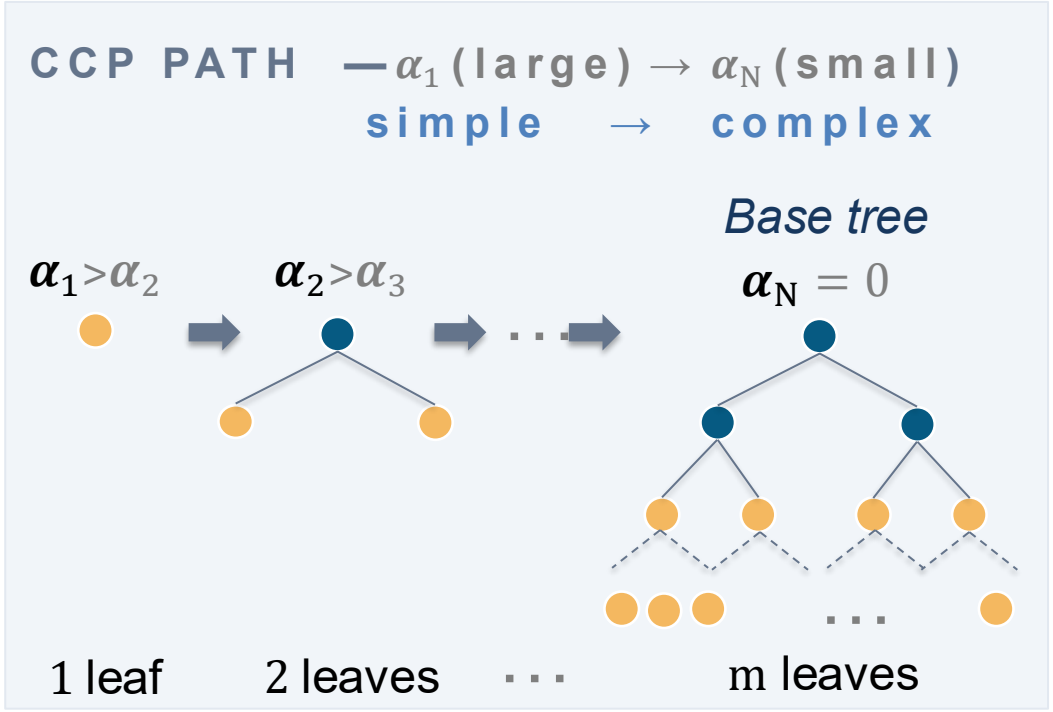
The University of Manchester

Why a regression tree?

- **Task/target-aware splits** guided by stability metric (CCT), unlike unsupervised clustering
- **If–then rules** on individual feature thresholds; human-readable
- **CCP path** spans simple → complex partitions in a principled way

Procedure

1. Fit a *base tree* (specific depth and samples / leaf)
2. Compute and **sort** the effective alphas $\{\alpha_k\}$ via CCP
3. Cap candidate set (e.g. 10–50) to bound runtime
4. Each α_k → a candidate partition, larger α → simpler tree



Algorithmic efficiency:
Many leaves are unchanged across successive prunings: their SR equations are cached and reused, drastically reducing the number of expensive SR runs.

A constrained symbolic search per region

Search space is constrained and tailored to anticipated functional forms to enhance interpretability and reduce runtime.

Per-leaf PySR configuration

Binary operators	+ , − , × , ÷
Unary operators	log, sin, cos, √· (problem-specific)
Forbidden nesting	log(log(·)) and similar pathologies
Iteration budget	50 (synthetic) / 300 (39-bus)
Max complexity	10–20 nodes
Goal	low-complexity formulas, not deep ones

Each fitted model g_ℓ is summarized by:

- the recovered equation string
- its symbolic complexity (number of nodes)
- MSE_ℓ + the full distribution of squared errors e_i^2

EXAMPLE: leaf 1 with 500 samples

$$g_1 = 2 \cdot x_2 + \sin(x_3)$$

Complexity : 6 nodes

MSE_1 : 2.5×10^{-3} and $[e_1^2, \ldots, e_{500}^2]$

MANCHESTER 1824
The University of Manchester

Three complementary metrics for partition selection

MSE

Global accuracy

$$MSE(\alpha) = \sum_{\ell=1}^{L(\alpha)} w_\ell \; MSE_\ell, \quad w_\ell = \frac{n_\ell}{n}$$

with

$$MSE_\ell = \frac{1}{n_\ell} \sum_{i \in \mathcal{R}_\ell^\alpha} (g_\ell^\alpha(x_i) - y_i)^2$$

Sample-weighted mean of per-leaf errors. Reflects overall fit while preventing big leaves from dominating.

$CVaR_\beta$

Tail-risk safeguard

$$CVaR_\beta(\alpha) = \mathbb{E}\left[e_i^2 \mid e_i^2 \geq VaR_\beta\right],$$

$$e_i^2 = \left(\widehat{f_\alpha}(x_i) - y_i\right)^2$$

$\beta = 0.95$

Penalises the worst 5 % of squared errors — catches large failures hidden by the global MSE.

$LeafLp$

Leaf-level imbalance

$$LeafLp(\alpha) = \left(\sum_{\ell=1}^{L(\alpha)} w_\ell \; MSE_\ell^p \right)^{1/p},$$

p = 4

p > 1 makes ill-fit leaves dominate. As p→∞ it tends to the worst leaf — enforces a “no bad leaves” stance.

Each metric protects a different failure mode — together they avoid both over-fitting and over-segmentation.

MANCHESTER 1824
The University of Manchester

Composite score and the simplicity-bias rule

COMPOSITE SCORE

(5)

$$S_w(\alpha) = w_{\text{MSE}}\,\widetilde{MSE}(\alpha) + w_{\text{CVaR}}\,\widetilde{CVaR}_\beta(\alpha) + w_{\text{Lp}}\,\widetilde{LeafLp}(\alpha), \quad \sum_i w_i = 1,\ w_i \geq 0,$$

$\tilde{\bullet}$: *metrics normalised by their own minimum across α* · $w = (0.5, 0.3, 0.2)$ *baseline weighting*

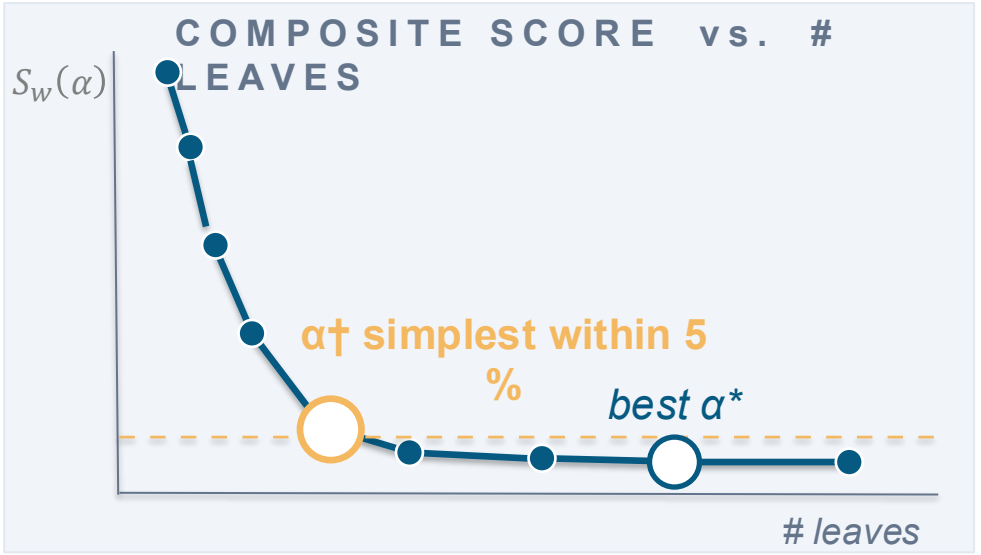
OPTIMIZATION

For A: set of candidate α, optimal $\alpha^\star$:

$$\alpha^\star = \arg\min_{\alpha \in A} S_w(\alpha)$$

Simplicity-bias rule

If multiple α candidates lie within 5 % of the best score, keep the simplest tree (fewest leaves) — denoted α†.



Post-hoc merging of equivalent regions

1. Canonicalisation	2. Structural check	3. Embedding similarity	4. Numerical validation	5. Iterate
Reduces false mismatches before similarity comparison (SymPy-based)	Compare variable & operator sets; incompatible pairs rejected	all-MiniLM-L6-v2 sentence transformer → cosine similarity	Cross-evaluate on both regions for borderline pairs	Reassign samples, retrain, repeat until convergence

Canonicalisation

Step 1

- **Parsing:** convert strings to SymPy expressions (sympify)
- **Normalise divisions:** rewrite $x / 0.5 \rightarrow x \cdot 2.0$
- **Round floats:** enforce consistent precision (e.g., $1.99999 \approx 2.0$)
- **Explicit multiplicative form:** variables as $r \cdot x$ (avoids x vs. $1.0 \cdot x$ mismatches)
- **Simplification:** apply algebraic rules (e.g. $x + x \rightarrow 2x$) via SymPy.simplify
- **Variable renaming:** rename to $x_1, x_2, \ldots$; avoids name-convention mismatches (e.g., DP25 vs. δ_5)

Cosine-similarity decision rule

Step 3

sim ≥ 0.90 → merge regions directly

0.80 ≤ sim < 0.90 → numerical validation

sim < 0.80 → keep regions separate

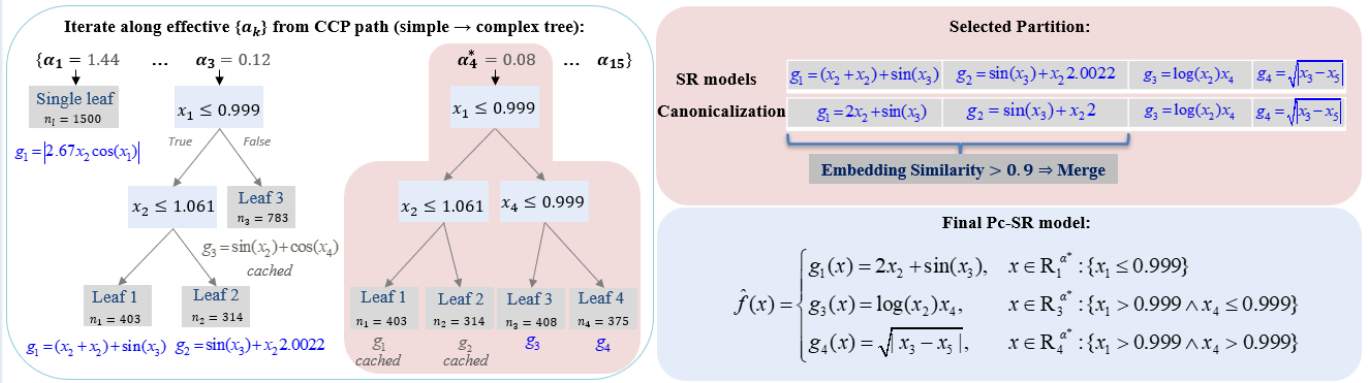
Synthetic 3-regime benchmark — known regions & equations

Ground truth

N = 1500 x=[0, 2) · added noise on y with σ = 0.05

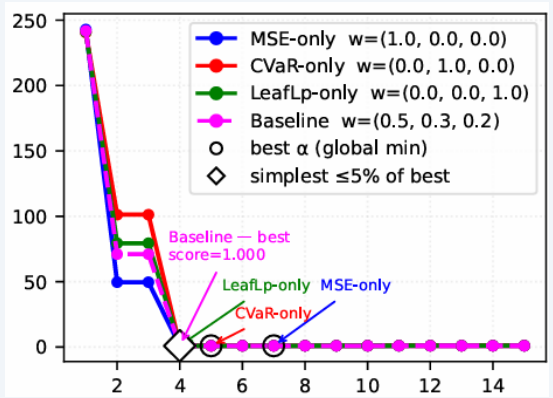
Region	Condition	Expression y	Samples
0	$x_1 < 1$	$2 \cdot x_2 + \sin(x_3)$	717
1	$x_1 \geq 1 \ \wedge \ x_4 < 1$	$\log(x_2 + 0.001) \cdot x_4$	408
2	$x_1 \geq 1 \ \wedge \ x_4 \geq 1$	$\sqrt{\lvert x_3 - x_5 \rvert}$	375

Pc-SR walk-through



Zoom

Composite score $S_w(\alpha)$ vs. # leaves



The University of Manchester

Recovering analytical CCT on SMIB system

SMIB systems with closed-form CCT → used as ground truth to validate Pc-SR.

Classical SG ↔ infinite bus

$$CCT_{SG} = \sqrt{\frac{4H(\delta_{cr} - \delta_0)}{2\pi f P_m}} = 0.1128\sqrt{\frac{H(\delta_{cr} - \delta_0)}{P_m}}$$

Sampled inputs: $S_{\text{rated}} \in [100, 500]$ MVA · $P_m \in [0.2, 0.95]$ pu · $H \in [1, 10]$ s. N=500 LHS.

Droop-controlled grid-forming converter ↔ infinite bus

$$CCT_{Conv} = \frac{\delta_{cr} - \delta_0}{m_p 2\pi f p^*} = 0.00318\frac{\delta_{cr} - \delta_0}{m_p p^*}$$

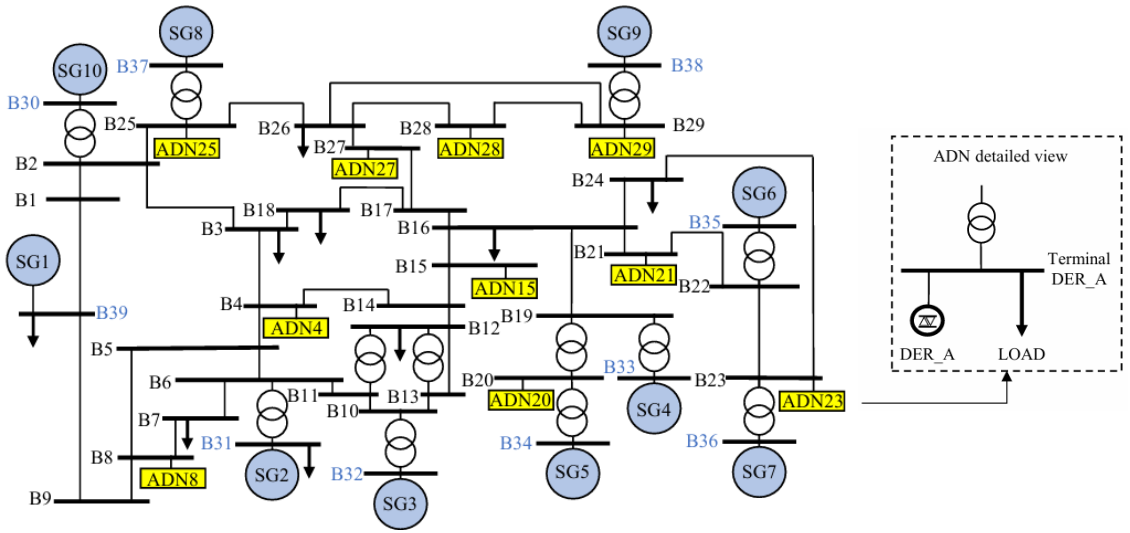
Sampled inputs: $S_{\text{rated}} \in [500, 1000]$ MVA · $p^* \in [0.2, 0.95]$ pu · $m_p \in [1, 10]$ %. N=500 LHS.

Balanced 3ϕ self-clearing fault applied at the point of common coupling. CCT is determined from RMS simulations.

Exact analytical structures rediscovered — Pc-SR collapsed to a single global expression in both cases.

Constants within **0.3 % (SG)** and **0.06 % (GFM)** of the analytical values.

Modified 39-bus system with active distribution networks



50-dimensional feature space

δ_i	initial rotor angles (10)
$SGKEi$	SG kinetic energy = $S_{\text{rated},i} \cdot H_i$ (10)
$SGPi$	SG active-power dispatch (10)
DPk	DER active-power dispatch (10)
LPk	load active power (10)

$i = 1, \ldots, 10$ SGs

k: indexing buses hosting ADNs

Simulation setup

- 10 ADNs (der_a model with dynamic voltage support)
- SGs as 6th-order machines with AVR + PSS (except SG1)
- 4 SG capacity levels: 100, 75, 50, 25 % of S_{rated}
- Sobol sampling of loads & DER injections; AC OPF in MATPOWER
- 3ϕ self-clearing fault at Bus 29; CCT via RMS binary search in PowerFactory
- 3754 converged scenarios from 4352 OPF problems

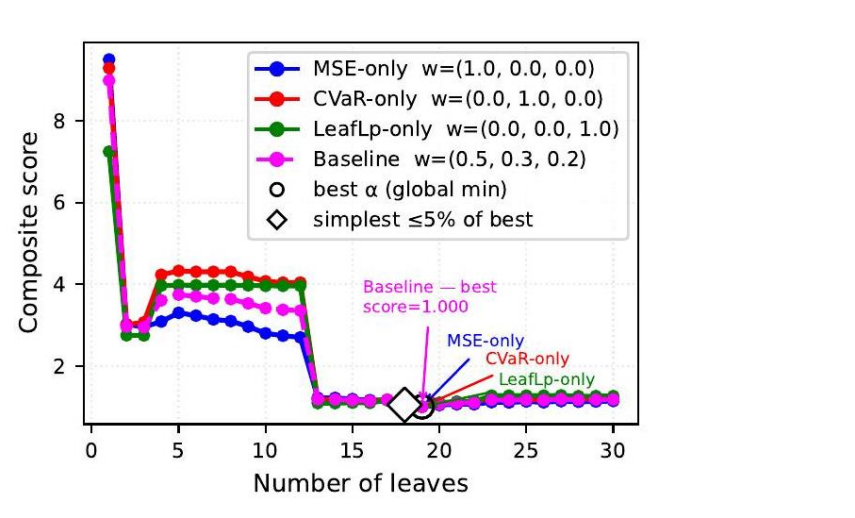
Values adapted for the 39-bus network

Min samples / leaf	50 → 100	*higher dimensionality*
Iteration budget	50 → 300	*harder search space*
Unary operators	+ sin, cos, √·	*physically motivated*
α-candidate cap	up to 50	*richer CCP path*

Pc-SR matches black-box accuracy with full transparency

R^2 across all 18 leaves	MSE (s^2)	RMSE (s)	$CVaR_{95}$ (s^2)
0.999	**$2.3 \cdot 10^{-5}$**	**0.0047**	**$1.6 \cdot 10^{-4}$**

Composite score $S_w(\alpha)$ vs. # leaves (CCT dataset)



Selected partition (α †)

18 leaves · depth 5

Root + first split on **SGKE features**
Further splits: **SG dispatch & rotor angles**

Matches black-box performance

XGBoost on the same dataset:
MSE $1.99 \cdot 10^{-5}$ · RMSE 0.00447 s · R^2 0.9992

End-to-end runtime ≈ 70 min

A regime-conditioned CCT equation

Representative leaf — region $\mathcal{R}_{18}^{\alpha^\dagger}$ (samples at 100 % Srated)

$$\mathcal{R}_{18}^{\alpha^\dagger} = \{ x \in \mathbb{R}^{\mathbb{d}} : \{SGKE5 > 8.12, SGKE8 > 21.26, SGP6 > 3.79, SGP7 > 2.88\}$$

$$g_{18}^{\alpha^\dagger}(x) = 0.012167\,(\delta_5 \cdot DP25) + \frac{\delta_1 + 1.3266}{\delta_9 + SGP9}$$

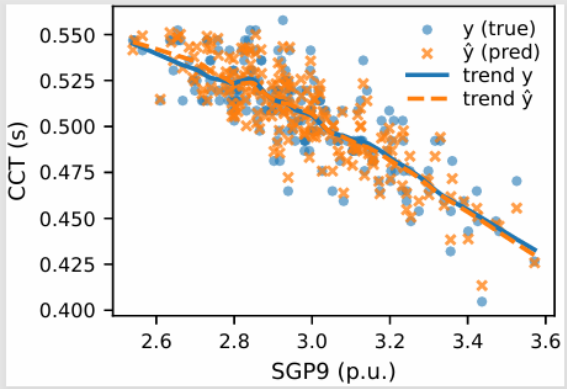
$R^2 = 0.96$ · $RMSE = 0.0057$ s

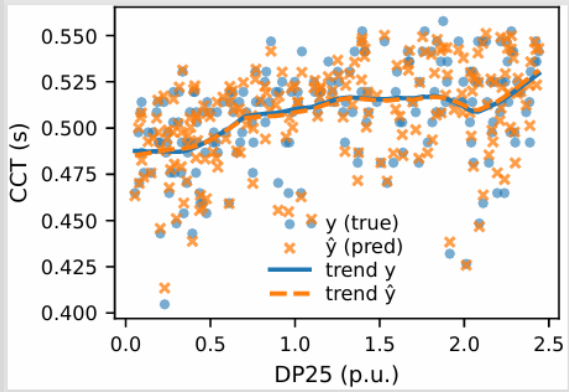
(δ_1 + 1.33) / (δ_9 + *SGP9*)

Rational term — higher pre-fault dispatch *SGP9* and larger initial angles δ_9 shrink CCT → consistent with equal-area considerations: more pre-fault loading leaves less decelerating margin.

0.012 (δ_5 · *DP25*)

Bilinear interaction — DER injection at bus 25 combined with favourable angular conditions at *SG5* enhances local synchronising torque → CCT increases.





Normalised sensitivities → $S_{\widetilde{SGP9}}/S_{\widetilde{DP25}} \approx 27$: SGP9 is the dominant CCT driver in this regime

MANCHESTER 1824
The University of Manchester

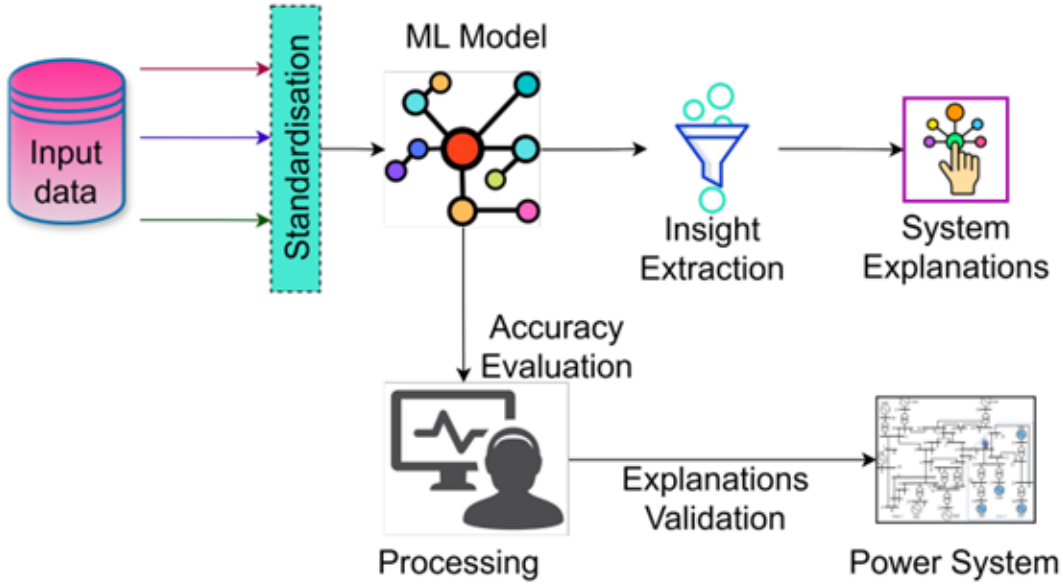
FUNCTIONAL APPROXIMATIONS FROM ML MODELS

Kolmogorov Arnold Networks

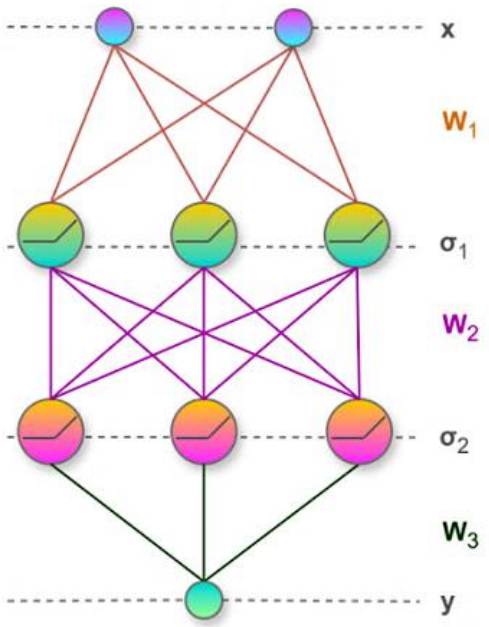
A. B. Kilembe, P. N. Papadopoulos, “Interpretable Power System Stability Assessment Using Kolmogorov-Arnold Networks”, PSCC 2026, Cyprus.

End-to-End Interpretable ML with KANs

- **Generate scenario datasets**
  - Use AC Optimal Power Flow (OPF)
- **Simulate system dynamics**
  - RMS time-domain simulations (TDS)
  - Build a power system stability dataset
- **Train KAN models**
  - Learn the system dynamics behaviour
- **Extract symbolic expressions**
  - Convert learned models into analytical/functional forms
- **Advantages**
  - Derive analytical relationships that are difficult to obtain traditionally
  - Replace black-box ML with interpretable models
- **Result**
  - Improved accuracy + interpretability
  - Enhanced trust in system dynamics analysis



Multilayer Perceptron (MLP) neural networks



Class of feed-forward artificial neural networks

Contains at least three layers, i.e., the input layer, the hidden layer with σ_v neurons each, and the output layer.

For an MLP model with V hidden layers, feature vector $\mathbf{x}$, weights matrix $\mathbf{W}_v$ of size (σ_v, σ_{v+1}) and bias vector $\mathbf{b}_v$ of size $(\sigma_v, 1)$, the output $\mathbf{z}$ of layer v is expressed as:

$$\mathbf{z}_1 = \mathbf{W}_1^\top \mathbf{x} + \mathbf{b}_1$$
$$\hat{\mathbf{z}}_1 = \Theta_1(\mathbf{z}_1)$$
$$\mathbf{z}_{v+1} = \mathbf{W}_{v+1}^\top \hat{\mathbf{z}}_v + \mathbf{b}_{v+1}$$

$$\hat{y} = \mathbf{W}_{V+1}^\top \hat{\mathbf{z}}_V + \mathbf{b}_{V+1}$$

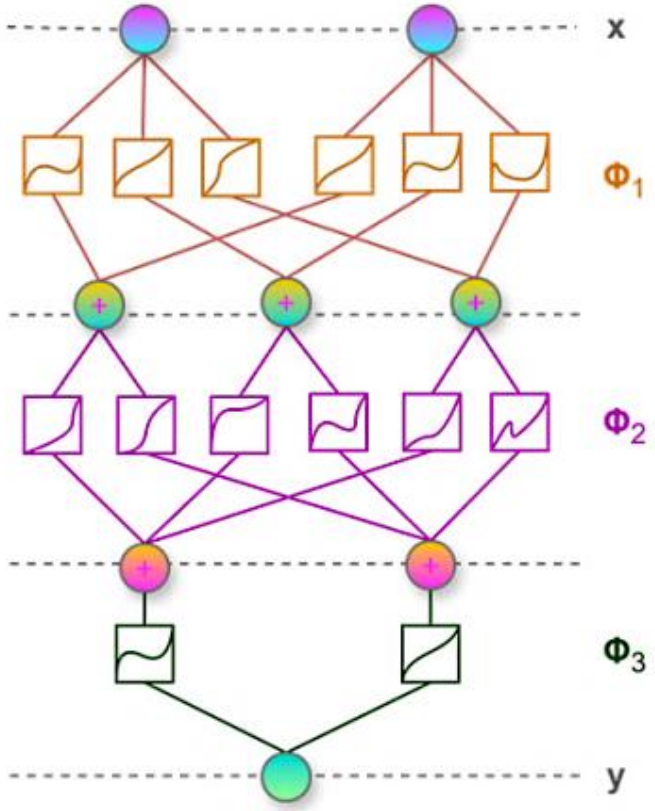
- Nodes apply an non-linear activation function, e.g. Rectified Linear Unit (ReLU) $\mathbf{z}_v = \max(0, \hat{\mathbf{z}}_v)$

Kolmogorov Arnold Neural Networks

- KANs are inspired by the Kolmogorov-Arnold representation theorem
- The theorem states that any multivariate continuous function on a bounded domain can be expressed as a finite composition of continuous univariate functions:

$$f(\mathbf{x}) = f(x_1, ..., x_n) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right)$$

- Similar to MLPs, KANs' compositional structure includes fully connected layers
- But instead of fixed weights on the edges, they are learnable univariate activation functions, ϕ, parametrised by splines.
- Each node (or neuron) is simply a summation operator on the incoming univariate functions



$$x_{v+1,q} = \sum_{p=1}^{n_v} \phi_{v,q,p}(x_{v,p}), \quad q = [1, \cdots, n_{v+1}]$$

MANCHESTER 1824
The University of Manchester

Kolmogorov Arnold Neural Networks

- Given a supervised problem and a KAN model with V layers and θ set of ϕ activation functions
- Training KANs aims to discover an optimal set of activation functions:

$$\theta = \{\mathbf{\Phi}_{V-1}, \mathbf{\Phi}_{V-2}, ..., \mathbf{\Phi}_0\}$$

where:

$$\mathbf{x}_{v+1} = \underbrace{\begin{pmatrix} \phi_{v,1,1}(\cdot) & \phi_{v,1,2}(\cdot) & \cdots & \phi_{v,1,n_v}(\cdot) \\ \phi_{v,2,1}(\cdot) & \phi_{v,2,2}(\cdot) & \cdots & \phi_{v,2,n_v}(\cdot) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{v,n_{v+1},1}(\cdot) & \phi_{v,n_{v+1},2}(\cdot) & \cdots & \phi_{v,n_{v+1},n_v}(\cdot) \end{pmatrix}}_{\mathbf{\Phi}_v} \mathbf{x}_v$$

Final output:

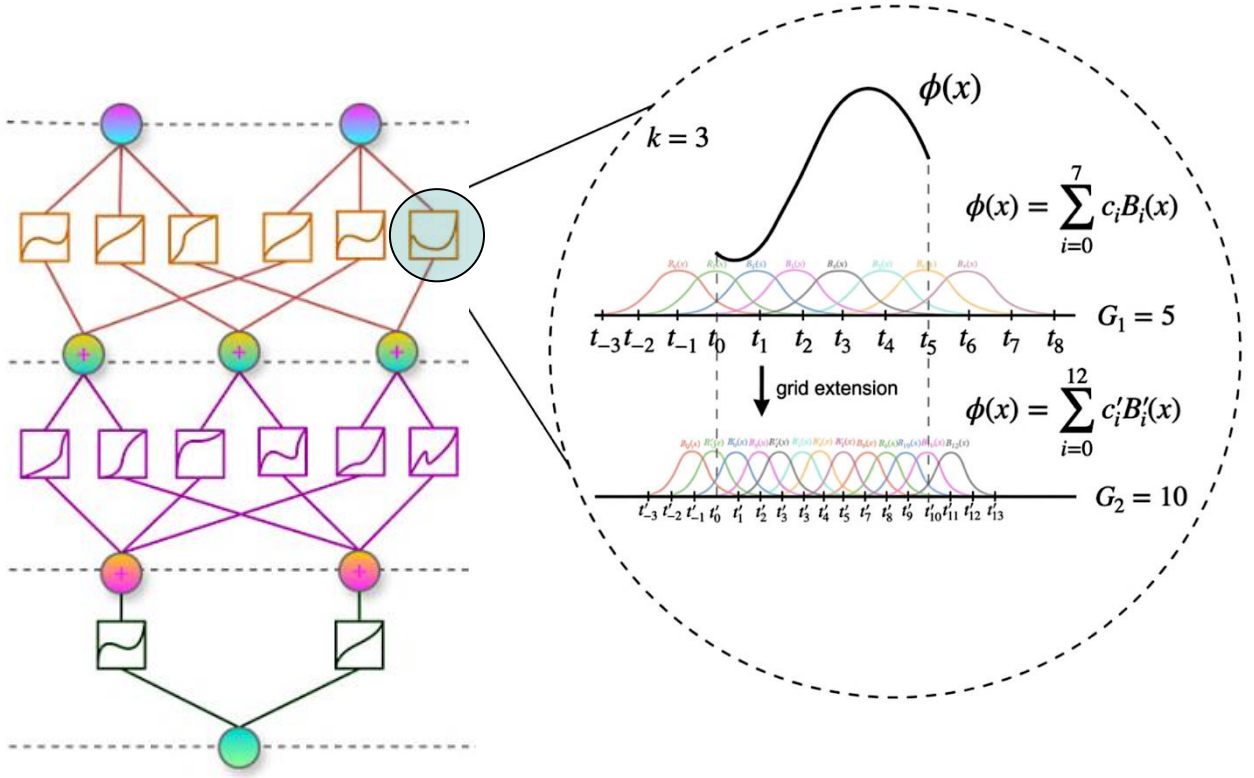
$$f_{\text{KAN}}^{\theta}(\mathbf{x}) = (\mathbf{\Phi}_{V-1} \circ \mathbf{\Phi}_{V-2} \circ \cdots \circ \mathbf{\Phi}_1 \circ \mathbf{\Phi}_0)\mathbf{x}$$

KANs activation functions

- Instead of fixed weights on edges, KANs have learnable univariate activation functions, ϕ, parametrised by splines.

$$\phi(x) = w_b \left(\frac{x}{1 + e^{-x}} \right) + w_s \sum_i c_i B_i(x)$$

Parameters c_i, w_b, and w_s are trainable. Controls the overall magnitude of the activation function

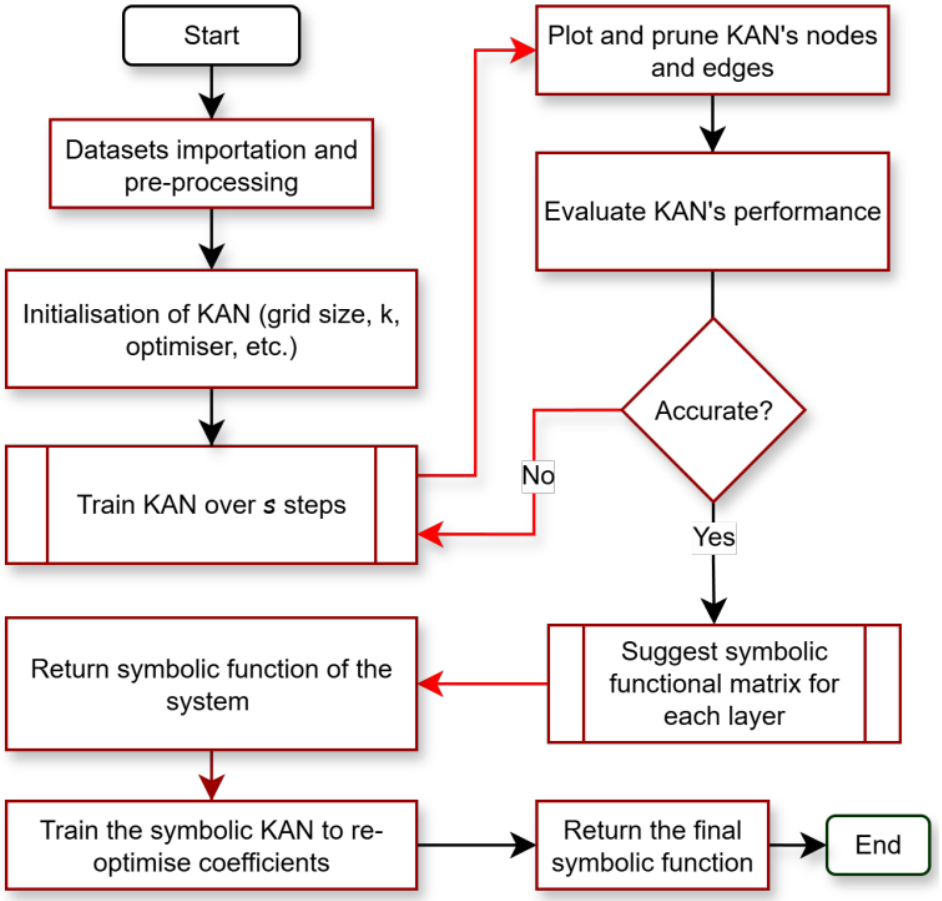


- Sigmoid as basis function
- Linear combination of B-splines
- Granularity of B-splines is determined by parameter *G*
- The order of B-spline polynomials is determined by parameter *k*

[1] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Solja˘ciÅLc, T. Y. Hou, and M. Tegmark, “Kan: Kolmogorov-Arnold networks,” 2024. https://arxiv.org/abs/2404.19756

MANCHESTER 1824
The University of Manchester

KANs for stability assessment



- **Load datasets and initialise the KANs – one for each location**
  - Architecture: $[n_0, n_1, .., n_V]$ and set parameters G, k,etc.
- **Apply regularisation**
  - L1 regularisation (λ), Entropy regularisation (λ_{ent})
  - All encourage sparsity (interpretability vs accuracy trade-off)
- **Train the model to predict the stability metric, then prune the graph**
  - Removes redundant activations
  - Use node importance scores with respect to a set threshold
- **Obtain simplified KAN**
  - Reduced complexity, improved interpretability
- **Symbolic regression (auto-discovery)**
  - Replace B-splines with mathematical functions
  - Library: $x^2, \sqrt{x}, e^x, \log, \sin, \tan$, etc.
- **Final output**
  - A function mapping system variables > to predicted metric

MANCHESTER 1824
The University of Manchester

Symbolic expressions discovery

- Auto-discovery of a symbolic KAN is based on selecting mathematical operations that minimise the MSE between the operators and the replaced B-spline function.
- Each learned spline $\phi(x)$is replaced by a symbolic function $g(x)$chosen from a library:

$$\mathcal{G} = \{x^2, \sqrt{x}, e^x, \log x, \sin x, \tan x, \arcsin x, \dots\}$$

- To find the best symbolic approximation:

$$g^* = \arg\min_{g \in \mathcal{G}, \theta_g} \mathbb{E}\left[\left(\phi(x) - g(x; \theta_g)\right)^2\right]$$

- Auto-discovery performed sequentially, starting from the first layer and first node
- Replace each spline: $\phi(x)_{v,q,p} \rightarrow g(x)_{v,q,p}$
- For example, the splines can directly be replaced with the following math operators:

$$\phi_1(x) \approx x^2; \ \ \phi_2(x) \approx \sin(x)$$

- After all substitutions, the model becomes an explicit symbolic formula.

MANCHESTER 1824
The University of Manchester

Case studies

- To validate the proposed approach in predicting bus-level RoCoFs, we use:
  - Modified IEEE 9-bus system
  - Modified IEEE 39-bus system
  - Texas 2000-bus system

- AC OPF is run in MATPOWER (solver: MATLAB Interior Point Solver (MIPS))

- RMS-TDS are performed on the AC-OPF solutions in DIgSILENT PowerFactory

- CIGs are Type IV fully rated converter models of the Western Electricity Coordinating Council (WECC)

- Renewables penetration ranges between 40% to 90%

- Our disturbance is the *N* - 1 SG outage of the largest generation

MANCHESTER 1824
The University of Manchester

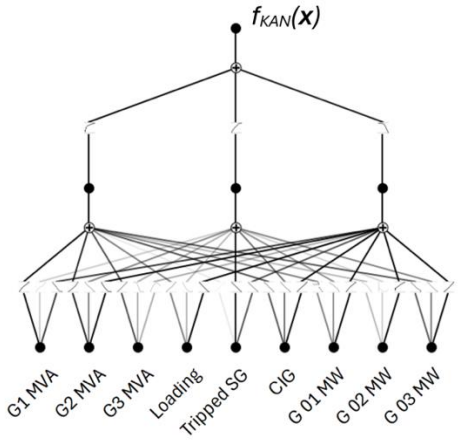
Predictions accuracy

- In the modified IEEE 9-bus system, KANs outperform MLPs (e.g. more than two orders of magnitude in many cases) in terms of the mean MSE
- A similar trend is observed in both the modified IEEE 39-bus system and the Texas 2000-bus system
- KANs require more training time on average at the expense of superior expressiveness

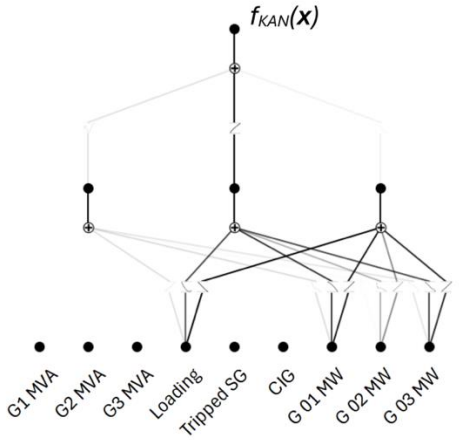
TABLE I: Model Performance in Predicting RoCoF in Three Test Systems

Test System	ML Model	Max MSE	Mean MSE	Min MSE	Training (s)
IEEE 9-bus	MLP	0.0020	0.0017	0.0015	0.91
	KAN	0.0004	8.52×10^{-5}	1.51×10^{-6}	61.14
IEEE 39-bus	MLP	0.0038	0.0018	0.0004	2.27
	KAN	0.0004	0.0001	1.32×10^{-5}	149.77
Texas 2000-bus	MLP	0.0045	0.0017	0.0006	3.88
	KAN	4.86×10^{-5}	1.64×10^{-5}	6.75×10^{-6}	301.23

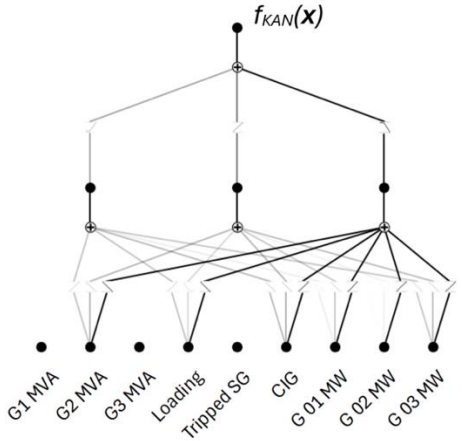
Feature importance visualisation in KANs



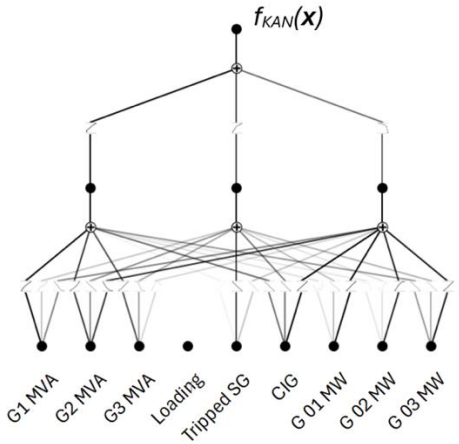
(a) KAN's activation functions over the training dataset



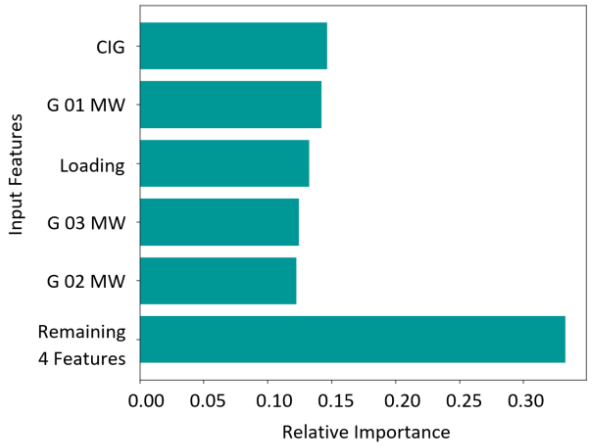
(b) Under no CIG penetration



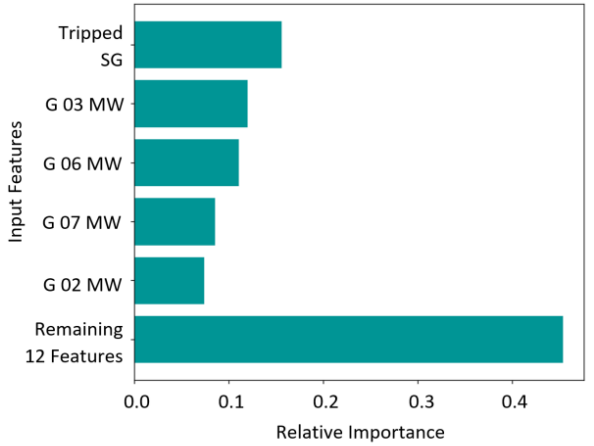
(c) Under high CIG penetration



(d) Under low system loading conditions



(a) Modified IEEE 9-bus system at Bus 1



(b) Modified IEEE 39-bus system at Bus 30

- Identifying important features in power systems reveals insights into complex dynamics
- KANs provide a mechanism to uncover these and understand stability risks
- Feature importance = the behaviour of activation functions over the input domain/sample (equivalent to global)

$$F_{x_p} = \sum_q A^I_{0,q,p} A^O_{1,q}$$

MANCHESTER 1824
The University of Manchester

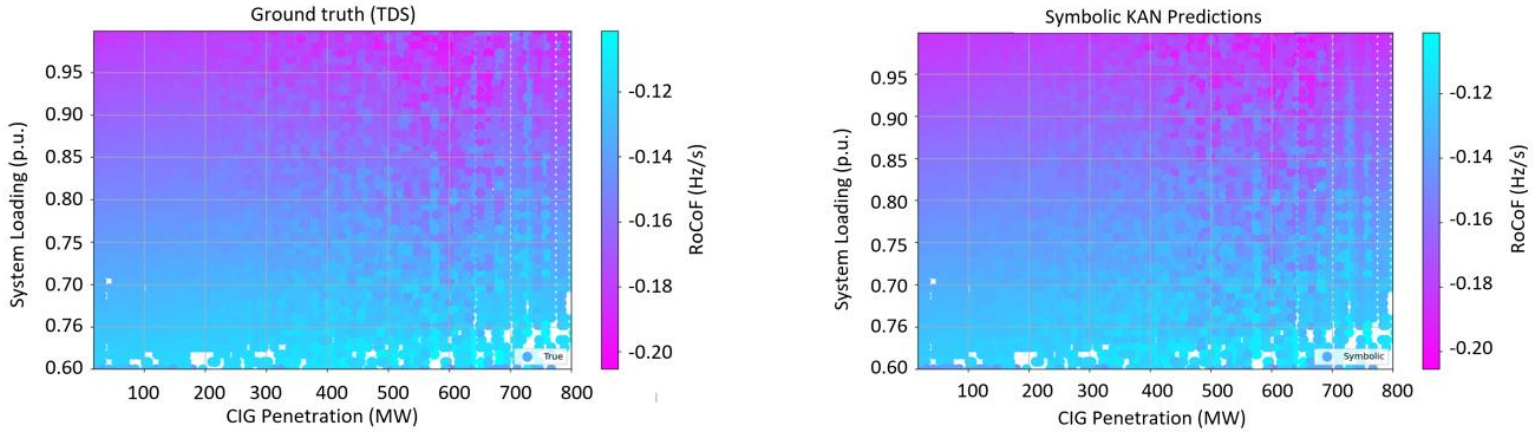
Sample of derived functional expressions

$$\begin{aligned}
\mathrm{RoCoF}_{\mathrm{Bus}\,30} &= a_1 \sin\Big(\mathbf{b}_1^\top \mathbf{x} + \mathbf{c}_1^\top \cos(\mathbf{d}_1 \mathbf{x} + \mathbf{e}_1) + r_1\Big) \\
&\quad + a_2 \cos\Big(\mathbf{b}_2^\top \mathbf{x} + \mathbf{c}_2^\top \sin(\mathbf{d}_2 \mathbf{x} + \mathbf{e}_2) + r_2\Big) + a_3 \\
\mathrm{RoCoF}_{\mathrm{Bus}\,31} &= a_1 - a_2 \sin\Big(\mathbf{b}_1^\top \mathbf{x} + \mathbf{c}_1^\top \exp(\mathbf{d}_1 \mathbf{x}) + \mathbf{e}_1^\top \sin(\mathbf{f}_1 \mathbf{x} + \mathbf{g}_1) \\
&\quad + \mathbf{h}_1^\top \cos(\mathbf{i}_1 \mathbf{x} + \mathbf{j}_1) + r_1\Big) \\
\mathrm{RoCoF}_{\mathrm{Bus}\,32} &= \mathbf{b}_1^\top \mathbf{x} + \mathbf{c}_1^\top (\mathbf{x} - \mathbf{m}_1)^2 + \mathbf{d}_1^\top \cos(\mathbf{e}_1 \mathbf{x} + \mathbf{f}_1) \\
&\quad + \mathbf{g}_1^\top \tan(\mathbf{h}_1 \mathbf{x} + \mathbf{i}_1) + a_2 - \frac{a_3}{(x_3 - r_1)^2}
\end{aligned}$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_{17}]^\top$ is the input feature vector; $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{d}_1$, $\mathbf{d}_2$, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{f}_1$, $\mathbf{g}_1$, $\mathbf{h}_1$, $\mathbf{i}_1$, $\mathbf{j}_1$, and $\mathbf{m}_1$ are coefficient vectors; and a_1, a_2, a_3, r_1, r_2 are scalar constants.

Performance degradation of symbolic eqs.

- There is a strong similarity between the two, highlighting KAN's ability to preserve accuracy
- The maximum MSE by the symbolic expression is **9.7×10^{-3}** (full KAN's: **4.0×10^{-4}**)



(a) Ground truth samples (b) Symbolic KAN's samples

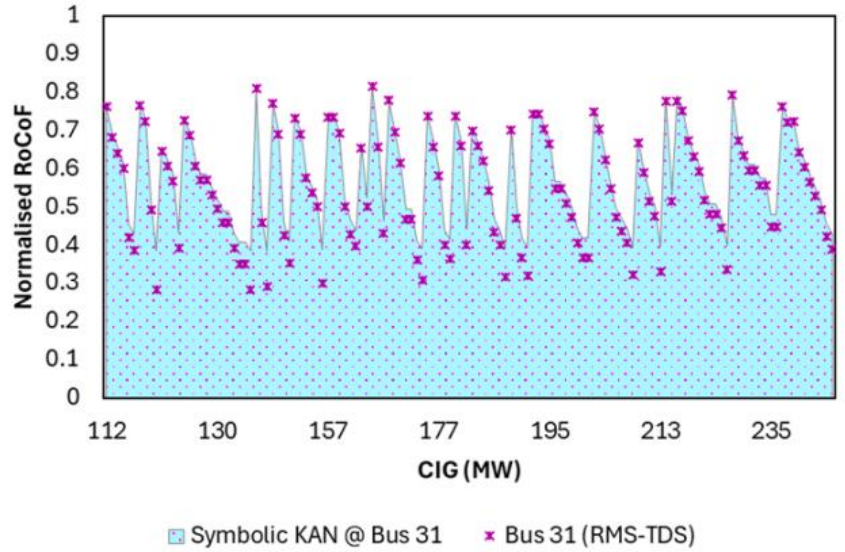
Fig. 6: Visual comparison of RoCoF in the modified IEEE 39-bus system at Bus 30

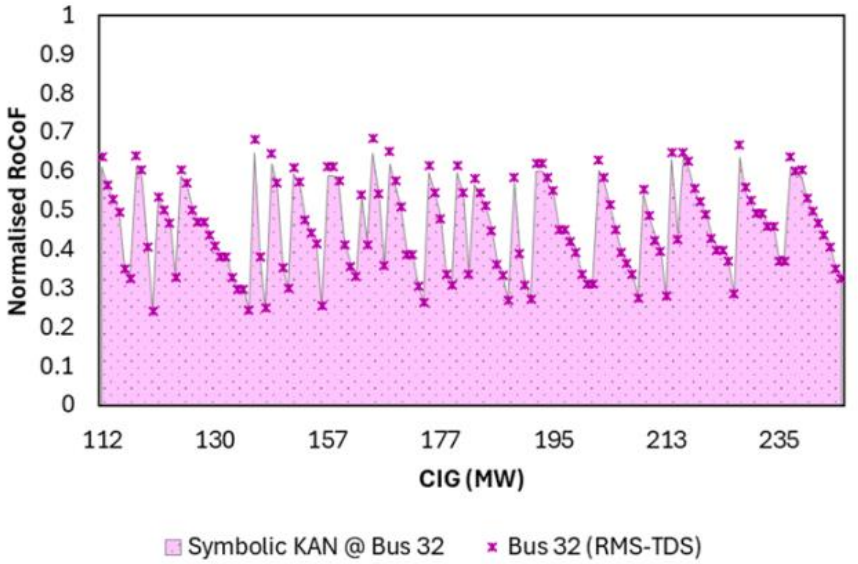
TABLE II: Accuracy of the Symbolic Expressions in the modified IEEE 39-bus System

Model	Max MSE	Mean MSE	Min MSE
Symbolic KAN	0.0097	0.0056	0.0016

Stability boundary calculation with KANs

- **Setup for analysis**
  - SG dispatch determined via AC-OPF optimisation
  - Vary CIG penetration: 110–250 MW
  - Minimise the number of changing variables for clarity
- **Key result**
  - KAN symbolic model accurately tracks RoCoF trends
  - Closely matches time domain results
- **Insight**
  - Symbolic KAN enables clear visual stability boundary analysis
  - Supports the interpretation of how renewable penetration affects system frequency stability





Local sensitivity analysis

- A local sensitivity analysis quantifies the influence of small perturbations in individual input features.

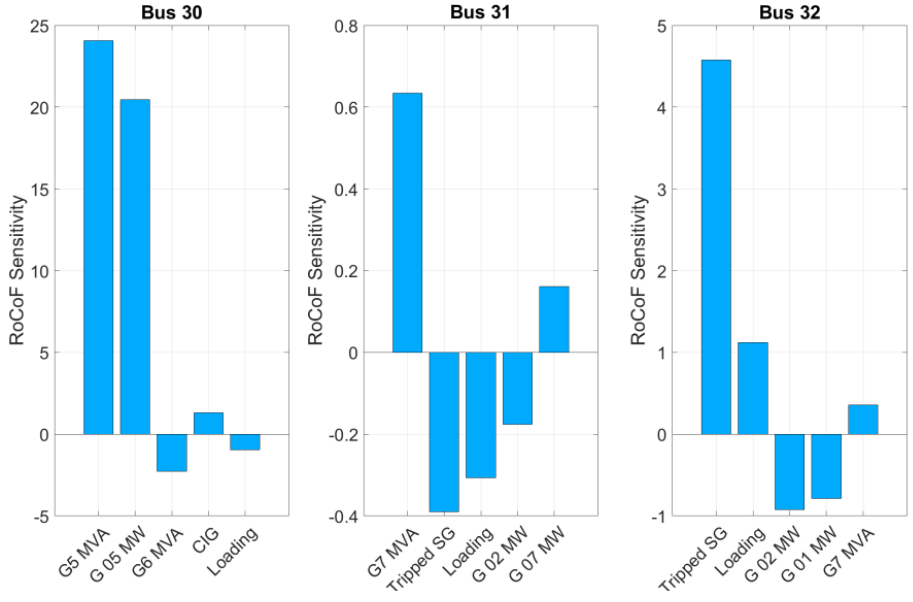
- The sensitivity of RoCoF to the *ith* feature is defined as:

$$S_{\mathrm{RoCoF}}^{(i)} = \frac{\partial f(\mathbf{x})}{\partial x_i}$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_n]$, and $f(\mathbf{x})$ is the KAN's RoCoF function.

- This can reveal useful insights

- **Bus 30**
  - Strong positive influence from SG 05
  - Smaller generator allows de-loading of larger generators (reduces N−1 impact)
- **Bus 31**
  - SG 07 rating is important (inertia-related effect)
- **Bus 32**
  - Most sensitive to tripped generator (fault location/type effect)



MANCHESTER 1824
The University of Manchester

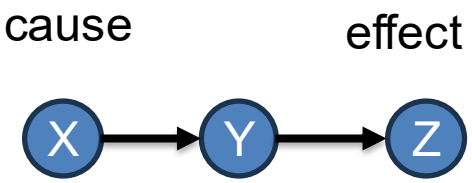
STRUCTURAL CAUSAL MODELLING

Identifying causal drivers for stability assessment

Faheem Ul Haq, Panagiotis N. Papadopolous, Robin Preece, “Identifying Causal Drivers of Transient Stability Boundary using Structural Causal Modeling”, to be submitted.

Correlation vs causation – causal theory

- ML models learn statistical associations which are not necessarily causal
  - They might learn from correlations in the data
- Correlations measure the statistical association between the variables without implying a cause-effect relation.
- Causal theory asserts that the variables within a process inherently reflect the cause-and-effect relationships dictated by the underlying physical principles governing the system or the process.
- An effect arises only as a consequence of its cause and cannot occur in its absence.
- Causal relations are governed by the underlying physical principles/ laws

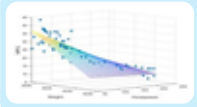


Structural Causal Modelling for stability assessment

- **Causal Discovery**: Identification of underlying cause–effect relationships among variables, resulting in a Directed Acyclic Graph (DAG) representing the process structure.
- **Refutation**: The learned DAG is subjected to statistical refutation tests based on the faithfulness assumption to ensure consistency with the observed data.
- **Prediction modelling**: The refuted DAG is used to identify the Markovian parents of the target, forming a causally relevant feature set to train and test an ML model for predicting the stability boundary.
- **Causal Effect Estimation**: The causal effect of each parent variable is quantified to identify the most significant drivers of system stability.
- **Intervention**: An active intervention is performed on the VSS to enhance the stability margins of the system.
- **Counterfactual**: A counterfactual reasoning is provided to why the causal model suggested a particular intervention by proving that alternative interventional would have a negative impact on system stability.


Causal discovery


Refutation process


Prediction modelling


Causal Estimation

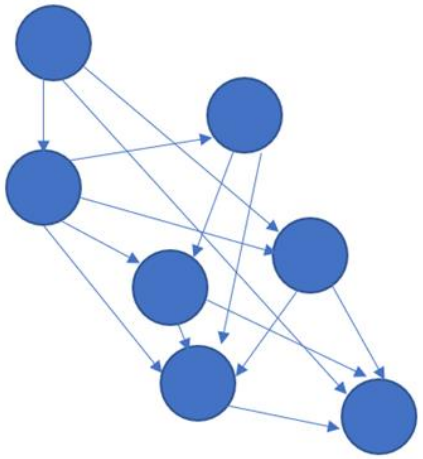

Intervention


Counterfactual

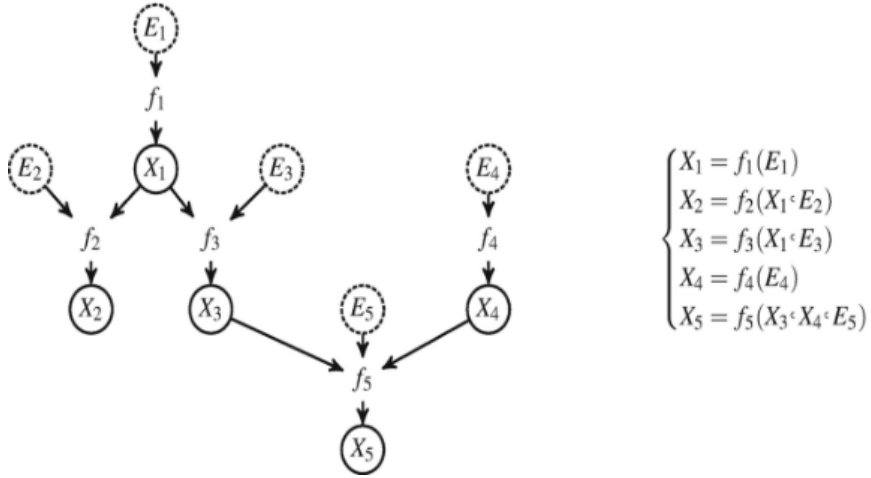
MANCHESTER 1824
The University of Manchester

Causal Model

- A causal model M, is defined as a set $M =< G, \Theta_G >$, where G is a causal structure and a set of parameters θ_G compatible with G, such that Θ_G assigns a function $X_i = f(Pa_{X_i}, U_i)$.



Causal Graph



Functional Causal Model

Causal Discovery using Directed Acyclic Graph-Graph Neural Network (DAG-GNN)

- DAG-GNN is a deep learning-based continuous optimisation method used to infer causal structures from observational data.

- This framework is built upon a simplified representation of causal relationships, known as the structural equation model, expressed as:

$$X = A^T X + Z$$

- *X* Variables in observational data
- *A* Adjacency matrix/causal graph
- *Z* Noise matrix.

$$X = A^T X + Z$$

Generative model for observational data

$$X = (1 - A^T)^{-1} Z$$

Generalised non-linear form

$$X = g_1(1 - A^T)^{-1} g_2 Z$$

Recovery model for Noise

$$Z = (1 - A^T) X$$

Recovery model for Noise

$$Z = g_3(1 - A^T) g_4 X$$

- Causal structure *A* can be identified if *X* and *Z* are known. However, only *X* is known while *Z* remains entangled in *X*.
- *Z* can be disentangled if *X* and *A* are known. However, we don't know the causal structure.
- Variational inference says *Z* can be disentangled by using a family of known probability distributions as proxy to encode its true posterior.

Yu, Yue et al. "DAG-GNN: DAG Structure Learning with Graph Neural Networks." International Conference on Machine Learning (2019).

Causal Discovery using Directed Acyclic Graph-Graph Neural Network (DAG-GNN)

- Accurate disentanglement of latent noise *Z* from observed data *X* is essential to validate a candidate adjacency matrix *A*, assessed through the faithful reconstruction of *X* from *Z*. Therefore, two complementary objectives are monitored:

- **Approximation quality of the noise posterior:** A measure of how well the family of proxy distributions $q(Z|X)$ approximates the true prior $P(Z)$ via a KL-divergence.

$$D_{KL}(q(Z|X)||P(Z))$$

- **Reconstruction Fidelity:** Given samples $Z \sim q(Z|X)$ we evaluate how well the generative model $P(X|Z)$ explains the observed data in expectation:

$$E_{q(Z|X)}[logP(X|Z)]$$

Combining both objectives yields Evidence Lower Bound (ELBO)

$$ELBO = E_{q(Z|X)}[logP(X|Z)] - D_{KL}(q(Z|X)||P(Z))$$

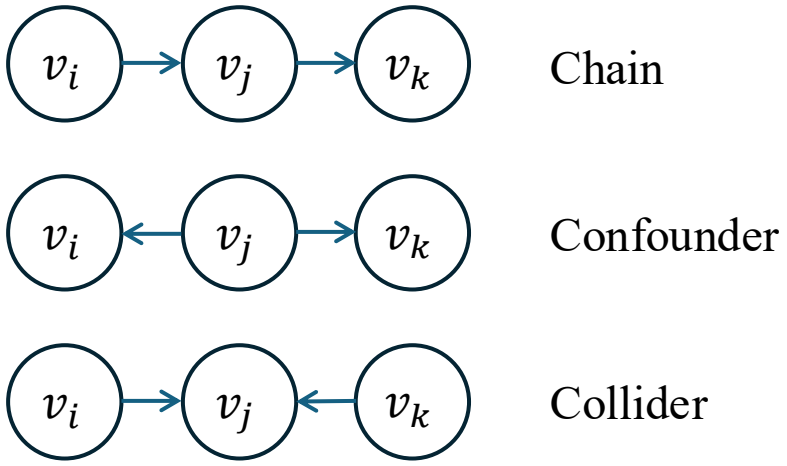
The training process of DAG-GNN is the maximization of ELBO

$$\min_{A,\phi} -ELBO$$

$$s.t.\ \ trace\left(\left(I - \gamma A \odot A\right)^{m}\right) - m = 0$$

Which yields A^* that represents the causal structure of X

Refutation of the causal graph



- To validate the accuracy of the adjacency matrix A, we subject the corresponding DAG to statistical refutation tests.
- Central to this validation is the faithfulness assumption, which provides a link between the graphical structure and the statistical properties of observed data.
- Faithfulness: A causal graph A over a set of variables with a joint probability distribution P, is deemed faithful if A accurately encapsulates the conditional independencies present in P.

Independence test for chains and confounders:

$$D_{KL}(v_i, v_k|v_j) = \sum_{n=1}^{\infty} P(v_i, v_k|v_j) log\left(\frac{P(v_i, v_k|v_j)}{P(v_i|v_j)P(v_k|v_j)}\right)$$

Independence test for Colliders:

$$D_{KL}(v_i, v_j) = \sum_{n=1}^{\infty} P(v_i, v_k) log\left(\frac{P(v_i, v_k)}{P(v_i)P(v_k)}\right)$$

D-separation criterion:

For chains and confounders:
Terminal nodes are conditionally independent on mediation node.

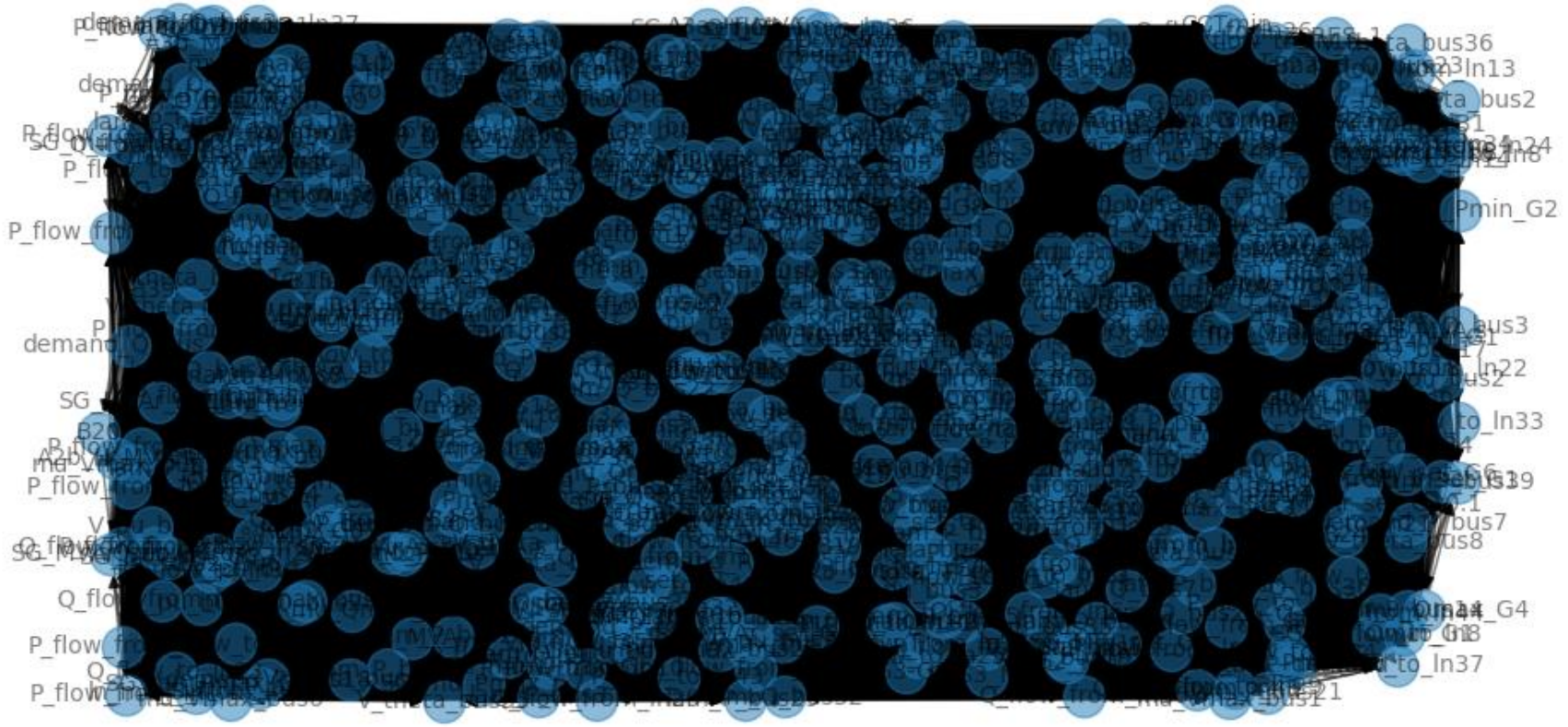
$$v_i \perp v_k|v_j$$

For colliders:
Terminal nodes are marginally independent.

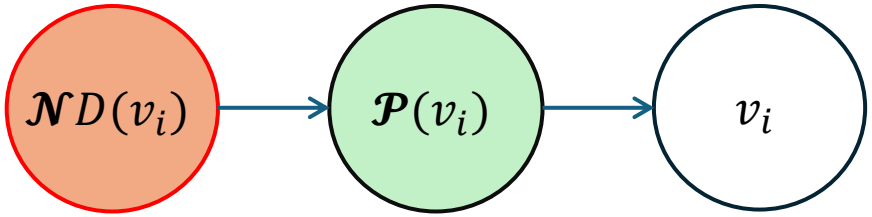
$$v_i \perp v_k$$

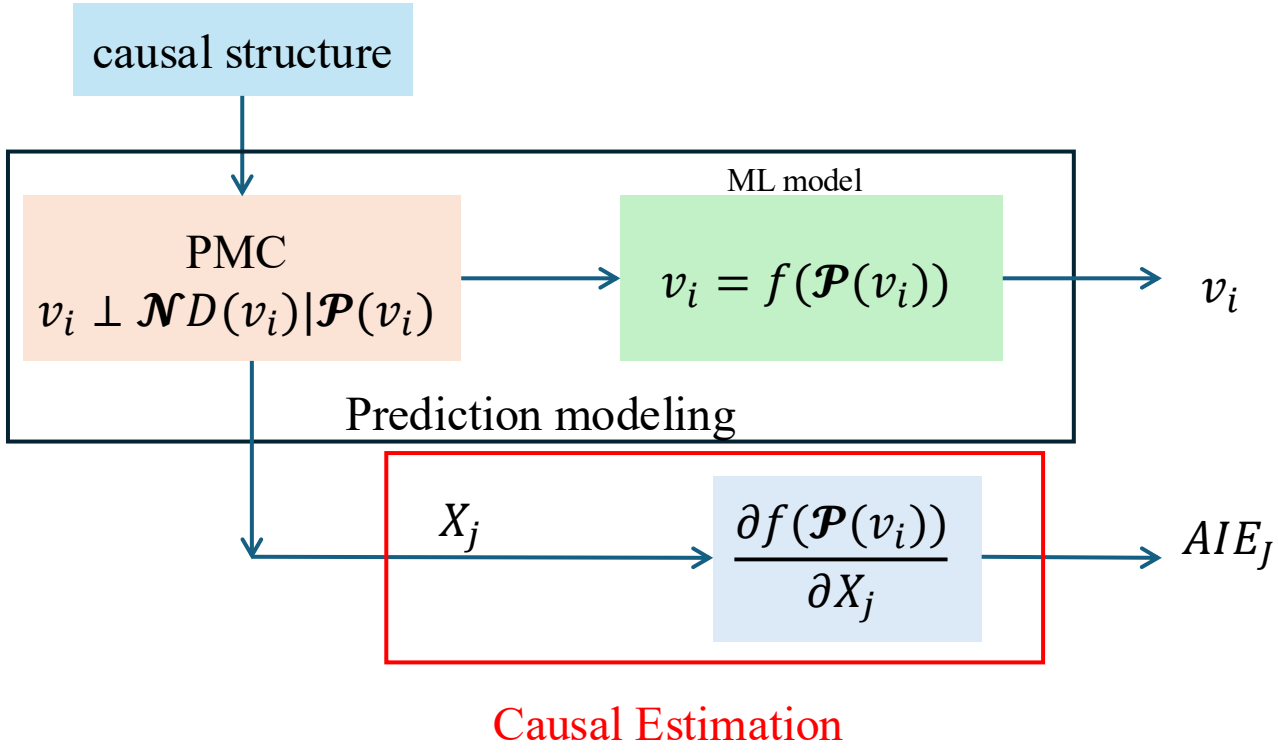
MANCHESTER 1824
The University of Manchester

Full causal graph can be messy



Prediction Modelling and Causal Estimation

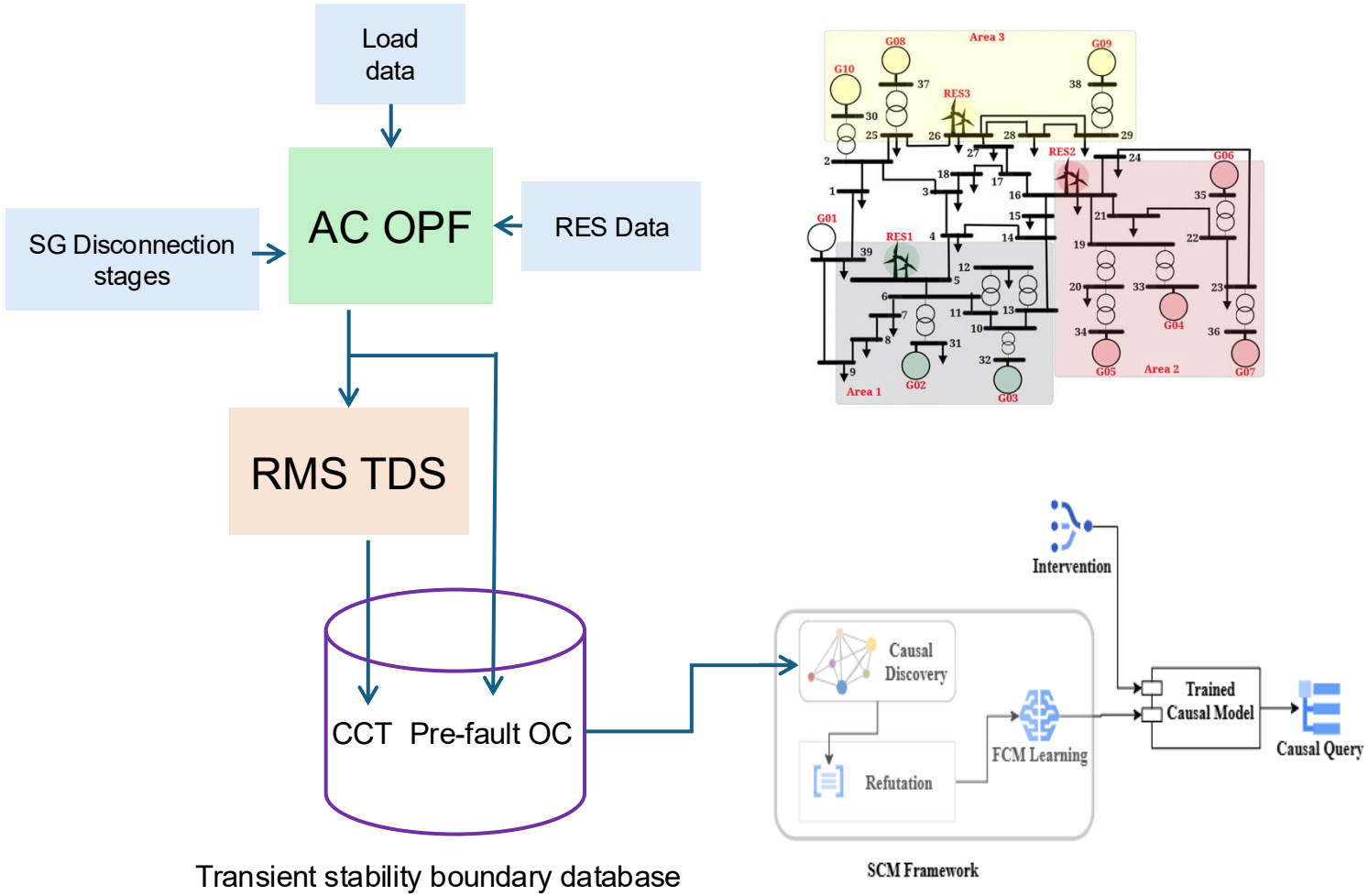




Causal Estimation

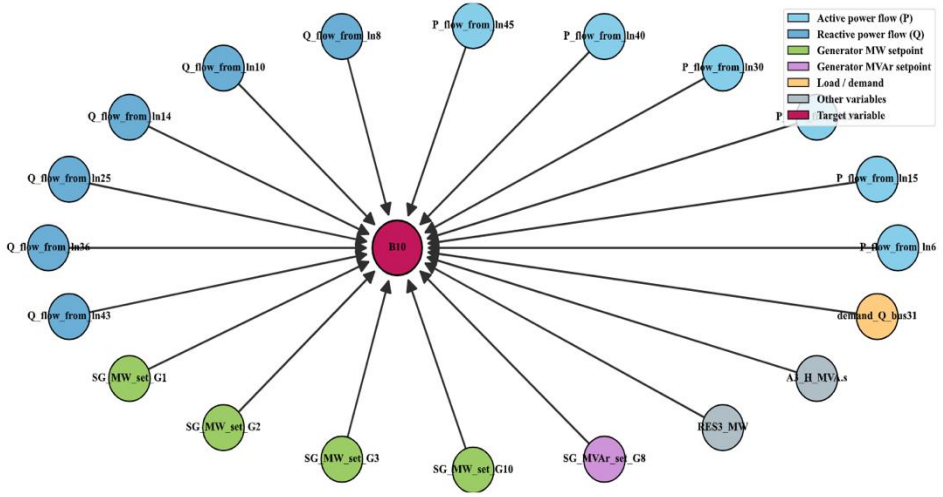
- Parental Markov condition: If the direct causes/parental nodes $\mathcal{P}(v_i)$ of a target are observed, every other non-descendant $\mathcal{N}D(v_i)$ becomes irrelevant.
- This reduces the number of features for a target used to train ML models for predictive purposes.
- Causal Effect of each parental node can be estimated using Average Intervention Effect (AIE)
- Larger magnitudes of AIE represent the variables of significance for stability.

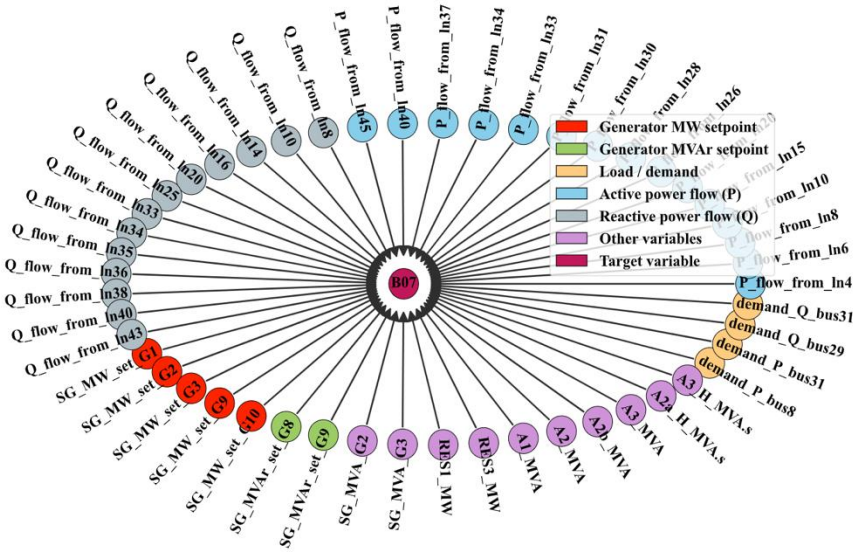
Case study and implementation

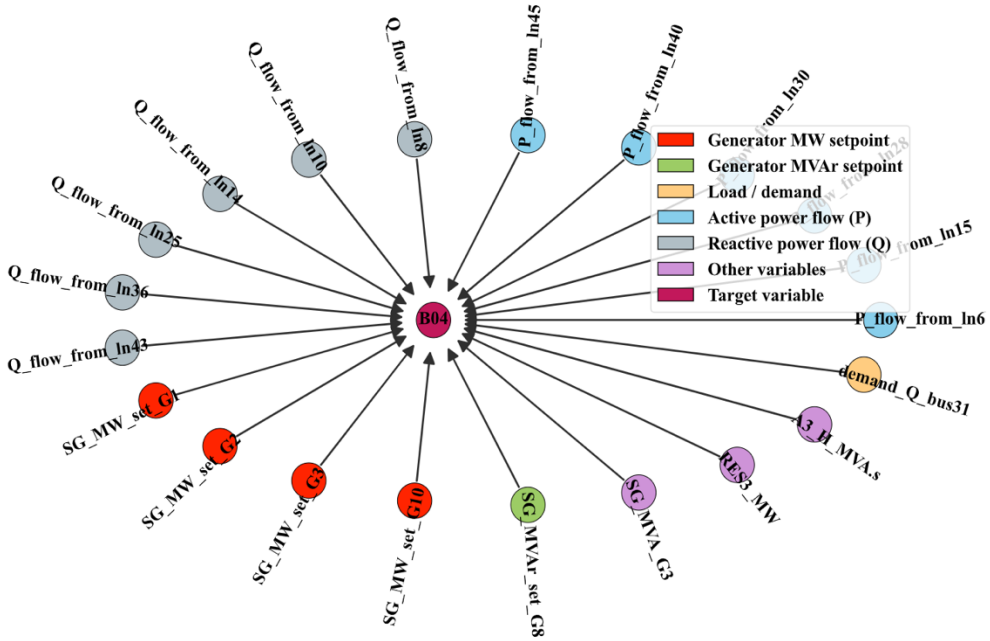


- Implementation to IEEE 39 bus test system with renewables
- The pre-fault conditions along with the CCTs, defines transient stability boundary with pre-fault system conditions as features and CCTs as target stability boundaries.
- This database is subjected to the SCM framework to determine the causal structure of the transient stability of the system.

Causal discovery results







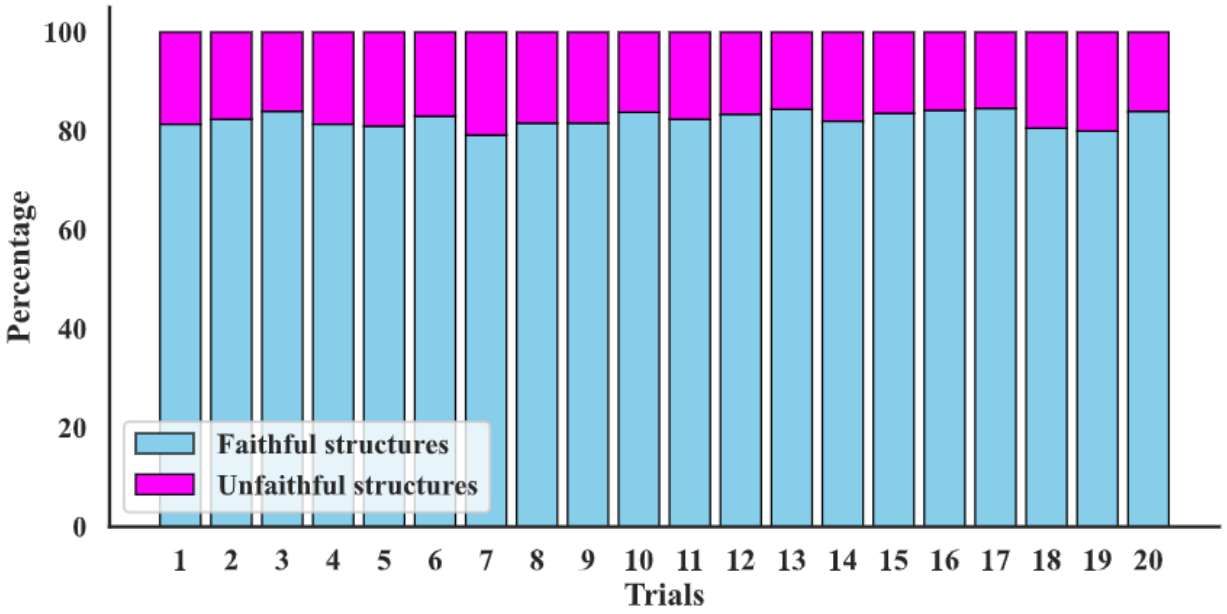
- Each DAG represents a subgraph of the global causal structure retaining only the first-order Markov parents of the respective stability boundary targets.
- The causal drivers identified: generator dispatches, network power flows, system loading, etc.
- A small fraction of the input feature space is identified as direct causal drivers of the transient stability boundary, while most operating parameters exhibiting either mediated or no direct influence.

Causal graphs for local transient stability boundaries corresponding to buses 10, 04 and 07 respectively

Refutation results

- Causal discovery results are ideally verified against ground truth, which represents the gold standard for validation.
- However, the availability of ground truth is itself a fundamental challenge in power system studies.
- Statistical testing of the faithfulness assumption offers one such alternative, but the number of testable substructures grows in a combinatorial manner with the number of nodes, rendering exhaustive verification computationally intractable.
- Exhaustive refutation is challenging.
- Random sampling of collider structures is therefore adopted, providing a tractable Monte Carlo approximation of the overall refutation behaviour.

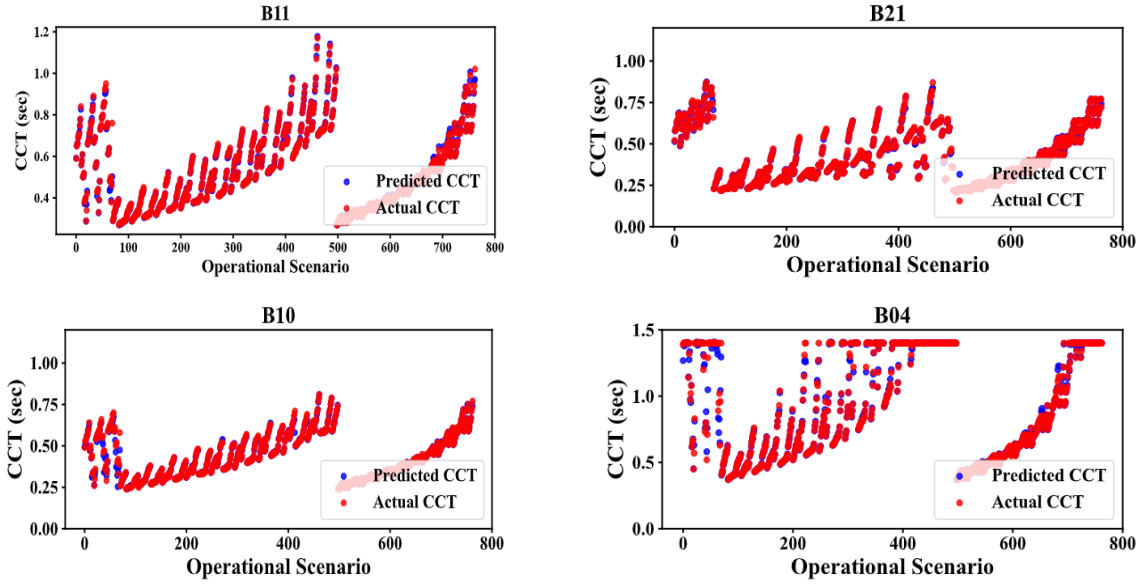
Refutation of collider structures



- On average 84% of the structures are faithful across all trials, used as a quantitative measure for the performance of the causal discovery approach.
- The remaining 16% of unfaithful structures cannot be statistically verified and require domain knowledge for further validation.
- Unfaithful structures enables the identification and pruning of false edges, yielding a more reliable causal graph.

Predictive modelling results

- ML models trained form refuted causal graphs exclusively from first order Markovian parents.
- This significantly reduces the input feature space.
- The causal predictive model is compared with conventional ML models with all explanatory variables as input feature space and SHAP based dimensionally reduced feature space.
- The results suggests that causal model retains the same explanatory power as that of the conventional ML model and SHAP based dimensionally reduced feature space.



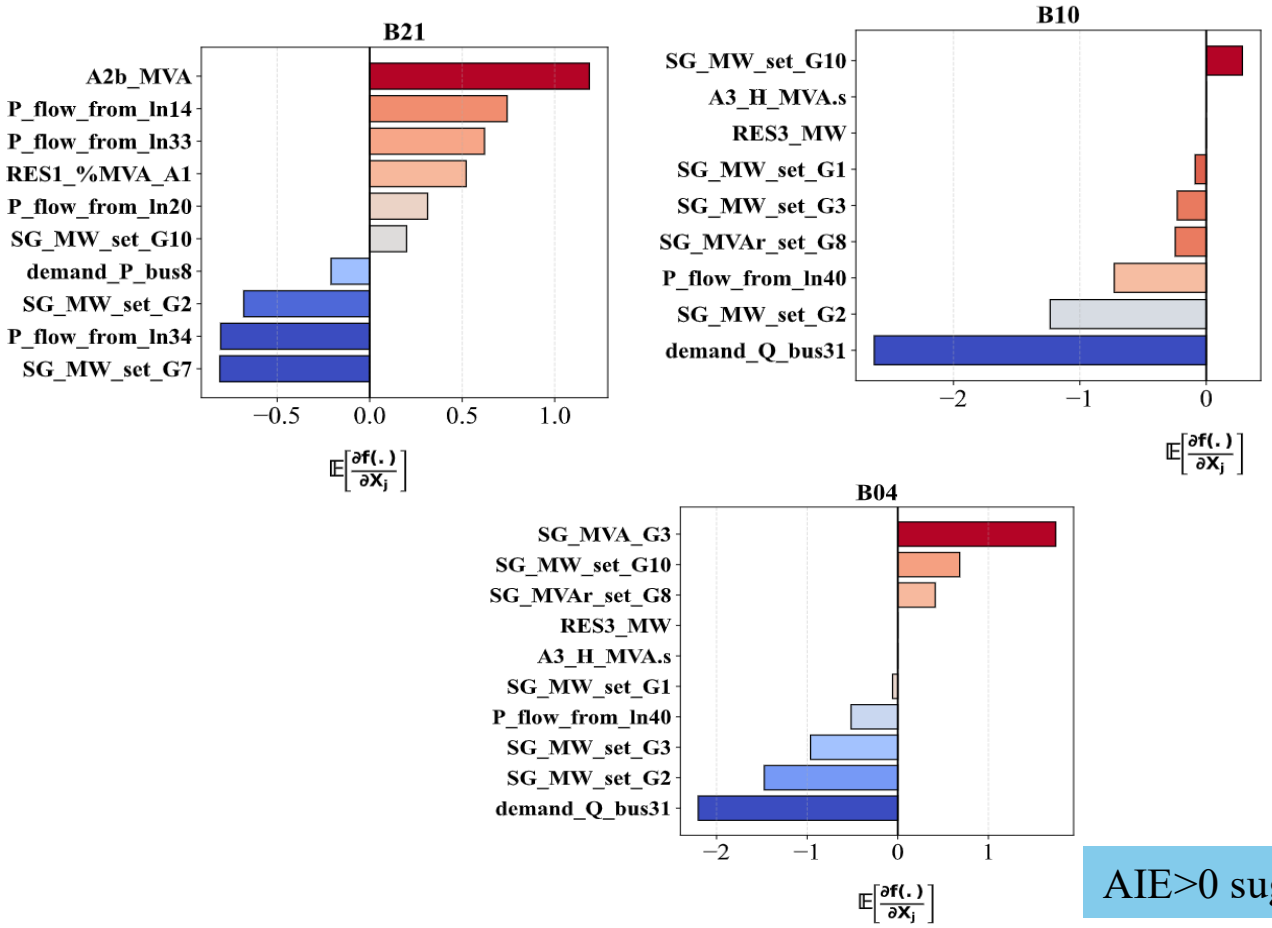
Comparison between true and predicted boundary using proposed causal model.

Metric	Full-Feature Model			SHAP-Based Reduced-Order Model			Causal Model		
	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
R^2	0.986	0.952 (B19)	0.996 (B23)	0.976	0.925 (B03)	0.996 (B12)	0.979	0.903 (B23)	0.999 (B22)
MSE	0.00053	0.00000491 (B12)	0.00206 (B27)	0.00114	0.00000419 (B12)	0.00872 (B03)	0.00066	0.00000624 (B12)	0.00242 (B03)
MOE	0.230	0.012 (B20)	0.827 (B03)	0.238	0.015 (B20)	0.918 (B03)	0.292	0.016 (B22)	0.868 (B27)
|MUE|	0.301	0.018 (B12)	0.870 (B11)	0.361	0.035 (B12)	0.869 (B11)	0.238	0.036 (B22)	0.695 (B11)

Predictive Performance Comparison of Full-Feature, SHAP-Based, and Causal Models.

Causal Estimation results

Average intervention effect



- Estimating the causal effect provides a quantitative measure of the degree to which each identified causal variable influences the transient stability boundary.
- This helps the identification of variables of significance for stability.
- Average Intervention Effect (AIE) estimation guides the interventional strategies to modify the stability boundary and enhance the stability margins.

AIE>0 suggests increasing the feature value increases the corresponding CCT.

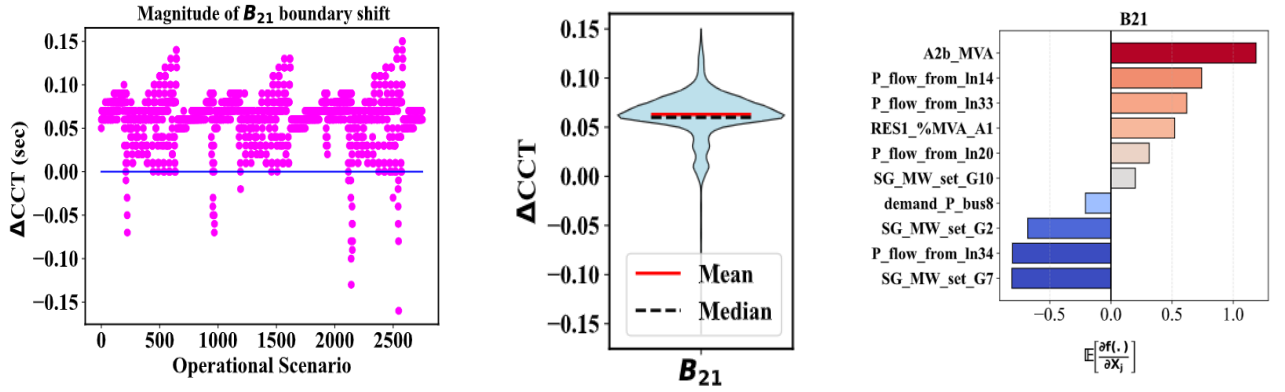
AIE<0 suggests decreasing the feature value increases the corresponding CCT.

Interventions

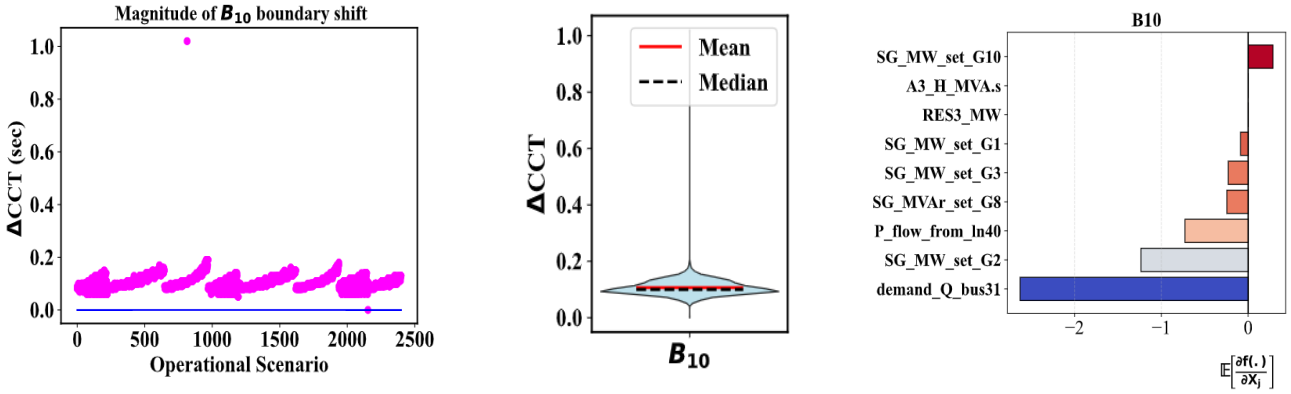
- An intervention, denoted $do(Z=z)$, is an external operation that forcibly sets a variable Z to a specific value z while leaving other structural equations unchanged.
- The causal effect of an intervention on Z for a target variable Y is quantified as the difference in the expected value of Y between its intervened and natural states, given by:

$$\Delta Y = \mathbb{E}[Y \mid do(Z=z)] - \mathbb{E}[Y]$$

- Guided by the AIE, interventions were performed enhance the stability boundary of bus 10 and 21.
- Guided by the negative sign of the AIE, the dispatch of $G7$ was reduced by 20%, resulting in an improvement in 97% of operating scenarios with an average increase of 63 ms in the CCT of $B21$.
- Similarly for bus 10, a 20% reduction on G2 dispatch resulted in improvement of 99% operating scenarios with an average improvement of 100 ms in the CCT of $B10$.



Impact on B21 stability boundary from intervention on SG_MW_set_G7 (Reduced by 20%)



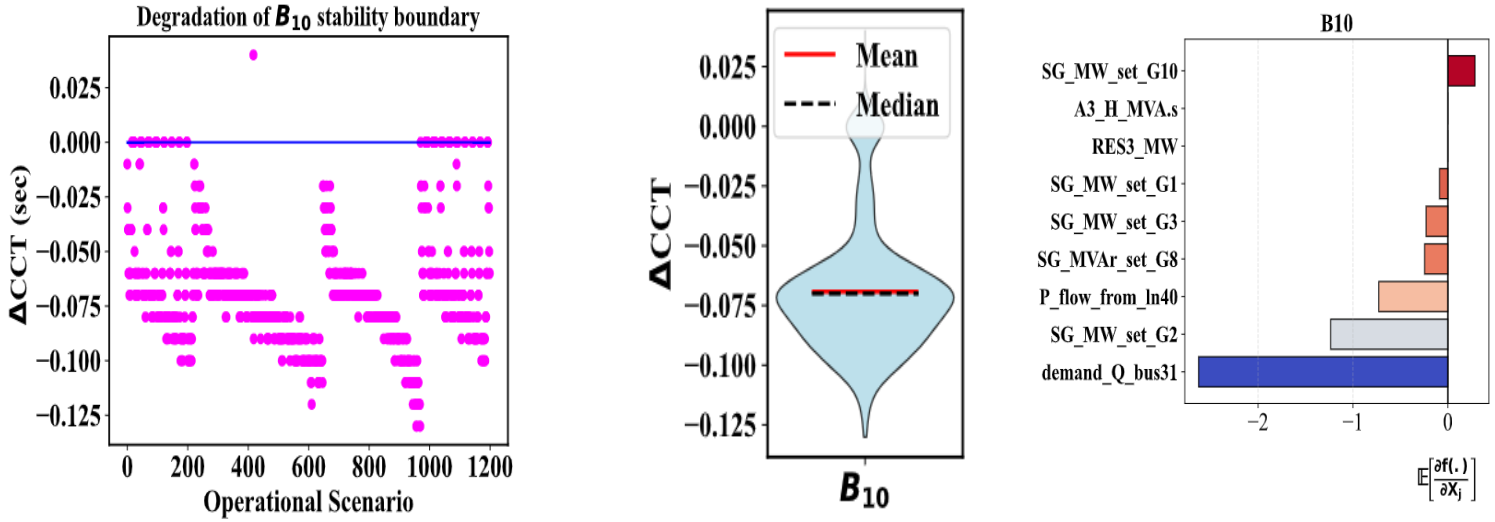
Impact on B10 stability boundary from intervention on SG_MW_set_G2 (Reduced by 20%)

Counterfactual intervention

- Counterfactual refers to a “what-if” scenario that is contrary to the observed fact, describing what the outcome would have been if an intervention had been applied differently.

- For instance, the AIE indicates that reducing the active power dispatch of $G2$ increases the CCT of $B10$, which is confirmed during the intervention process. A counterfactual analysis then examines the hypothetical scenario: what would have happened to the CCT of $B10$ had the dispatch of $G2$ been increased instead.

- Counterfactual intervention to $G2$ by increasing its active power dispatch by 20%, resulted in a degradation of the stability boundary by an average of 75 ms.



Impact on B10 stability boundary from counterfactual intervention on SG_MW_set_G2 (increased by 20%)

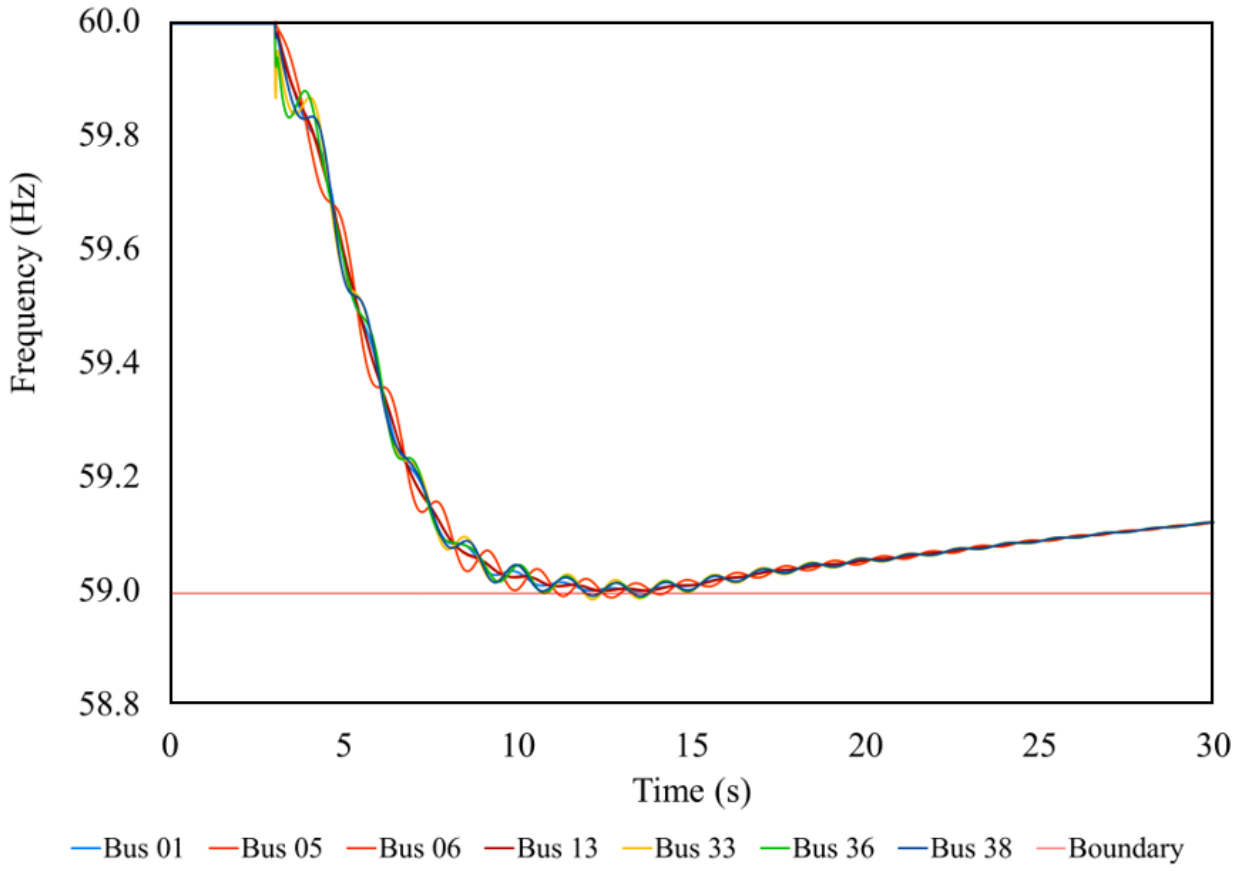
MANCHESTER 1824
The University of Manchester

INTRODUCING DETAILED DYNAMICS IN OPTIMISATION

Dynamics-aware optimisation: incorporating constraints from NNs in optimisation – an application to frequency stability

Machine Learning for Locational/Regional Frequency Stability

- In low inertia systems with a lot of converters, frequency dynamics are increasingly exhibiting locational/regional characteristics
- Center of Inertia (COI) approach (frequency as global phenomenon) can lead to:
  - A false sense of security leading to unforeseen locational relay activation.
  - Higher operational costs by securing frequency control resources more than needed.
- Analytical derivation of locational frequency dynamics is challenging (complex nonlinear dynamics)
- Solving DAEs is detailed but computationally intensive – hard to implement in optimization frameworks



A. Kilembe, P. N. Papadopoulos, W. Bukhsh, " *Neural Network-Constrained Optimal Power Flow for Locational Frequency Stability*," under 1st round of revision.

AC-OPF with locational frequency stability constraints

Objective Function

- The AC-OPF is a well-studied non-linear optimisation problem in power systems.
- Dispatches generators while minimising generation cost:

$$\min \sum_{i \in \mathcal{G}} C_i(P_{G,i})$$

subject to

$$g_0(x_0, u_0) = 0$$
$$h_0(x_0, u_0) \leq 0$$

Frequency Stability Constraints

- Embedded within the OPF to satisfy stability requirements:

$$h_w(x_0, u_0) \leq 0 \leq f_{\max}$$

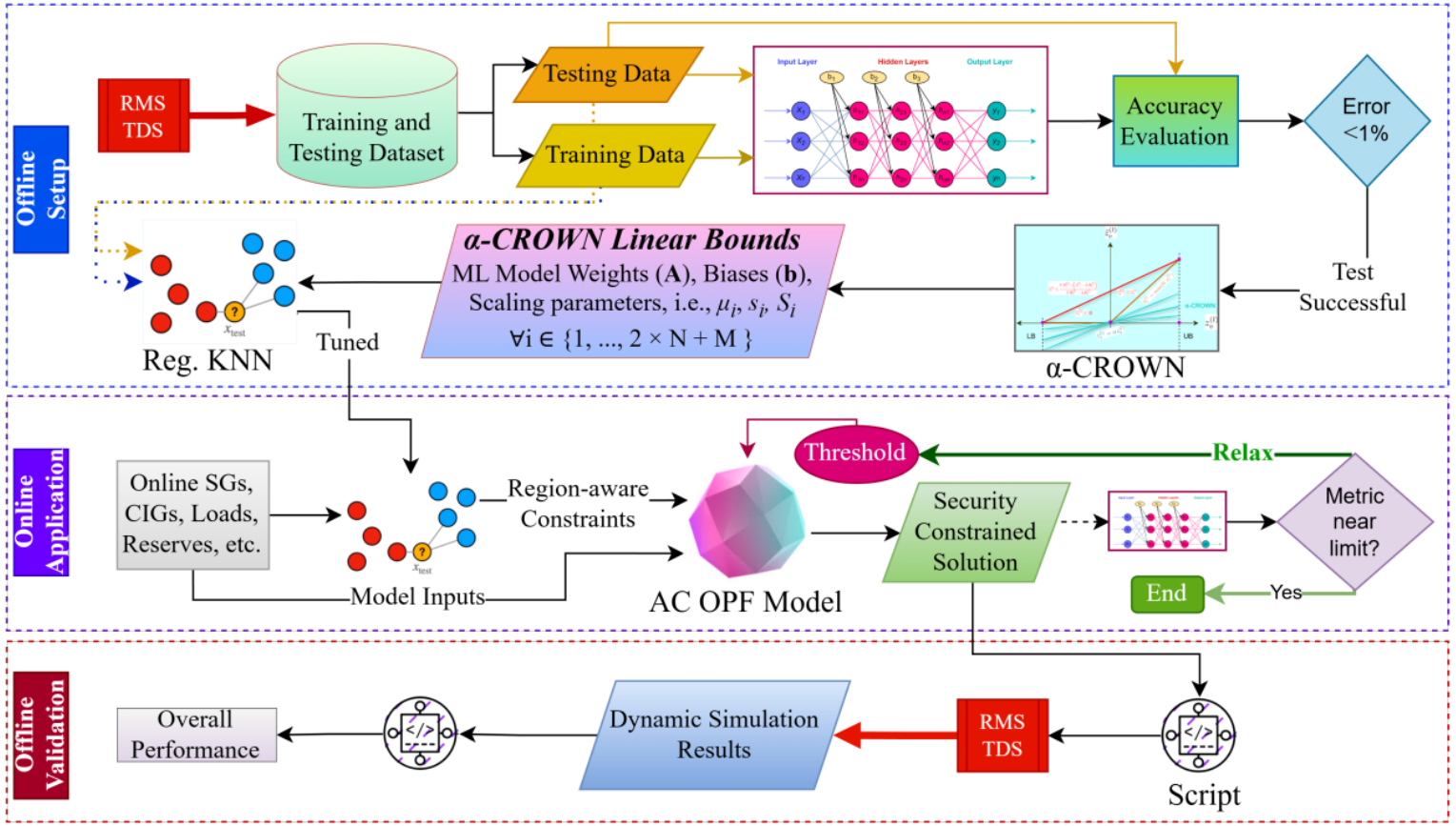
- The Centre of Inertia (COI) is often utilised to balance computational efficiency and modelling complexity:

$$RoCoF_{COI} = \frac{-\Delta P + \Delta P^{FR}}{2H}$$

Final optimisation vector contains standard system variables: $\hat{\mathbf{x}}^\top = [\Theta_m \quad V_m \quad P_g \quad Q_g]$

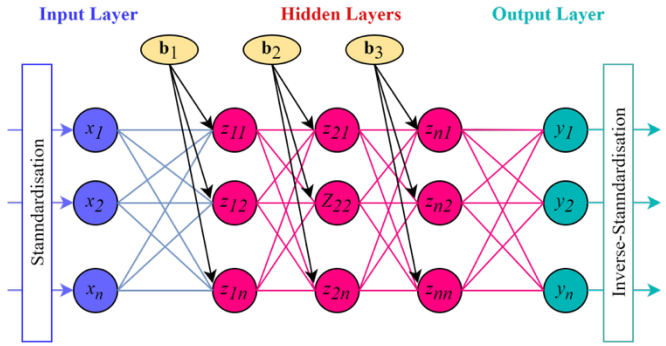
Methodology (3 stages)

- Train a Neural Network to capture regional frequency stability characterisitcs (e.g. nadir and RoCoF at various locations)
  - Using time domain simulations so capturing detail
- Linearise the trained neural network and formulate constraints to implement in optimization
  - Simpler approach to obtain linear constraints
- Maintain detail and avoid over-securing



A. Kilembe, P. N.. Papadopoulos, W. Bukhsh "Robust Neural Network-Constrained OPF with Verification-Assisted Region-Aware Constraints for Locational Frequency Stability," to be submitted.

Linear bounds for ReLU function



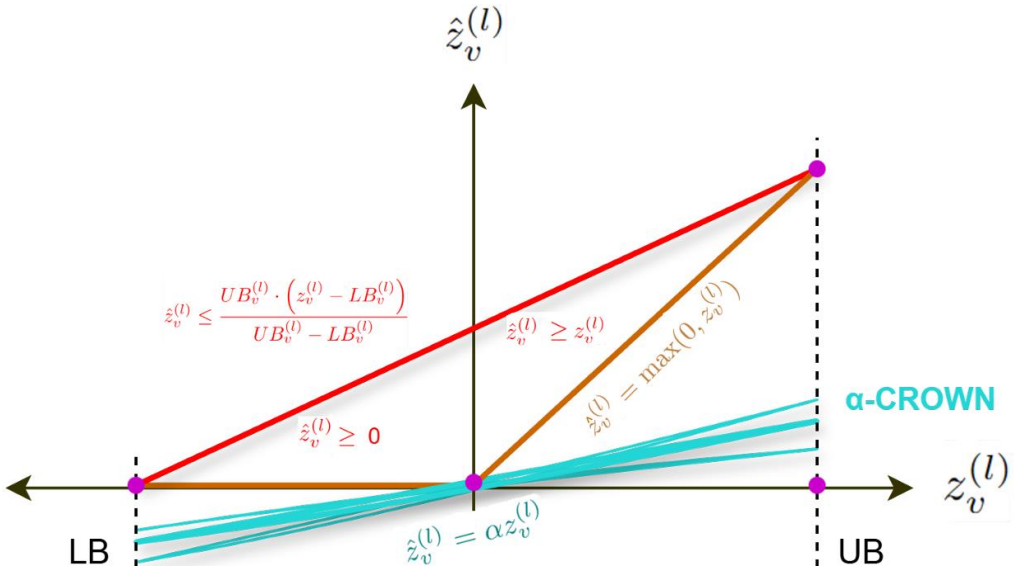
$$\hat{z}_v^{(l)}(x) = \mathrm{ReLU}(z_v(x))$$

The bounds on the neuron's output can be stated as:

$$\underline{a}_v^{(l)} z_v^{(l)}(x) + \underline{b}_v^{(l)} \leq \hat{z}_v^{(l)}(x) \leq \bar{a}_v^{(l)} z_v^{(l)}(x) + \bar{b}_v^{(l)}$$

Where:

$$\underline{a}_v^{(l)} = \begin{cases} \bar{a}_v^{(l)} = 0, & \text{(always inactive neurons)} \\ \bar{a}_v^{(l)} = 1, & \text{(always active neurons)} \\ \alpha_v^{(l)}, \quad 0 < \alpha_v^{(l)} < 1, & \text{(unstable neurons)} \end{cases}$$



Bounds on the neural network prediction are given as:

$$\mathbf{W}_{V+1}^{\top} \underline{\mathbf{D}}_V \mathbf{z}_V(\mathbf{x}) + \underline{\mathbf{b}}_{V+1} \leq f(\mathbf{x}) \leq \mathbf{W}_{V+1}^{\top} \bar{\mathbf{D}}_V \mathbf{z}_V(\mathbf{x}) + \bar{\mathbf{b}}_V$$

K. Xu, H. Zhang, S. Wang, Y. Wang, S. Jana, X. Lin, and C.-J. Hsieh, "Fast and complete: Enabling complete neural network verification with rapid and massively parallel incomplete verifiers," ICLR 2021.

Methodology cont'd

- The generated affine bounds fulfil the following conditions:

$$\min_{\mathbf{x}\in\mathcal{X}(\mathbf{x}_0)} \left(\underline{\mathbf{A}}_\alpha^\top \mathbf{x} + \underline{\mathbf{b}}_\alpha\right) \leq \min_{\mathbf{x}\in\mathcal{X}(\mathbf{x}_0)} f_{\text{True}}^{\text{LB}}(\mathbf{x}),$$

$$\min_{\mathbf{x}\in\mathcal{X}(\mathbf{x}_0)} \left(\bar{\mathbf{A}}_\alpha^\top \mathbf{x} + \bar{\mathbf{b}}_\alpha\right) \geq \max_{\mathbf{x}\in\mathcal{X}(\mathbf{x}_0)} f_{\text{True}}^{\text{LB}}(\mathbf{x}).$$

- $\mathbf{A}_\alpha$ and $\mathbf{b}_\alpha$ are candidate parameters (weights and biases, respectively) embedded within the optimisation framework as security constraints
- Note that two KNN models (for upper and lower bound) are fitted to predict these parameters $\mathbf{A}_\alpha$ and $\mathbf{b}_\alpha$
- Final optimisation vector with the neural networks-based security constraints becomes:

$$\hat{\mathbf{x}}_{\text{freq}}^\top = \begin{bmatrix} \Theta_m & V_m & P_g & Q_g & \mathcal{F}_i & \theta_j \end{bmatrix}$$

F_i: represents additional variables not in the standard optimisation vector, but useful for training the neural network

ϑ_j: is the set of model biases.

Case studies

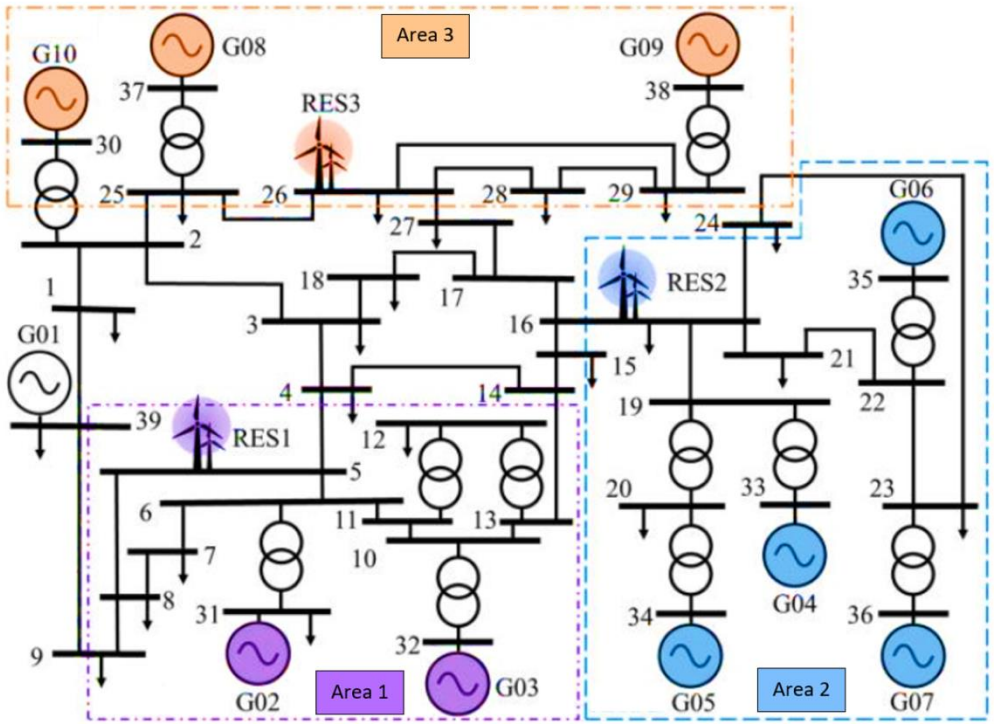
Case 1: *Fixed (N-1) Disturbance Location With Single Area CIG Penetration*

Case 2: *(N-1) Disturbance location changes based on the largest generating SG unit*

Model: IEEE 39-bus system (3 areas)
Number of SGs: 10
Number of CIGs: 3

CIG model: Type 4 wind generation Western Electricity Coordinating Council (WECC) model

The dataset is split 70%-30% for model training and testing, respectively



Verification of upper and lower bounds

Lower bound:

$$f_{pred}(x) - f_{True}(x) \leq 0$$

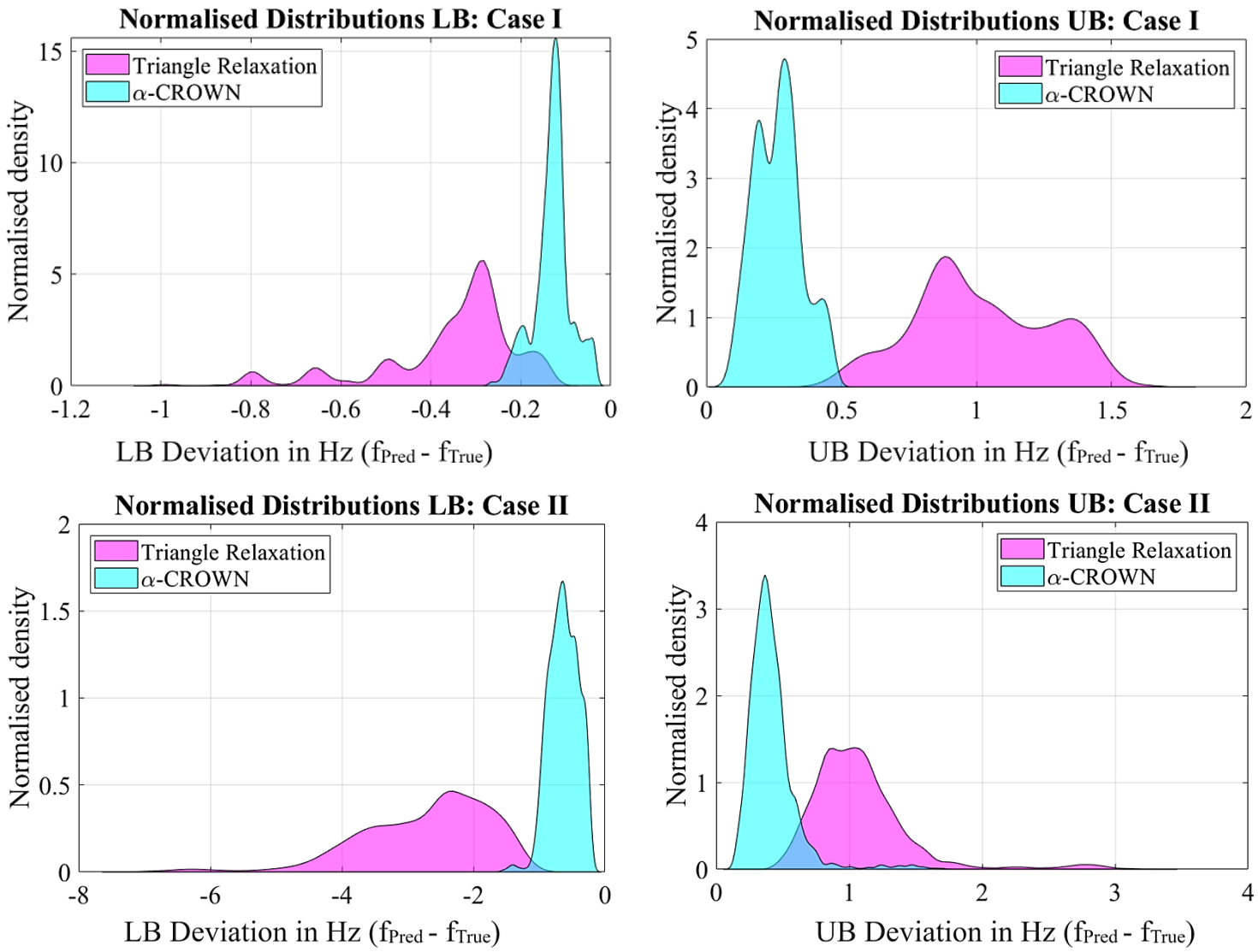
&

Upper bound:

$$f_{pred}(x) - f_{True}(x) \geq 0$$

This condition is always true for all samples $\mathbf{x}(\mathbf{x_0})$ around the sample $\mathbf{x_0}$: $\{\mathbf{x}| \ ||\mathbf{x} - \mathbf{x_0}||_\infty < \epsilon\}$

We apply an L_∞ ball of 5%



MANCHESTER 1824
The University of Manchester

Frequency dynamics prediction results

- In both cases, the neural networks learn to predict the locational frequency dynamics with high accuracy
- Since such predictions are computed via matrix multiplication, i.e.,

$$\boldsymbol{ML_{Pred}} = \boldsymbol{input} \times \boldsymbol{Weights} + \boldsymbol{bias}$$

  they are super fast
- This avoids solving differential equations, significantly reducing computational cost

TABLE I
ACCURACY OF THE NEURAL NETWORK MODEL FOR LOCATIONAL FREQUENCY STABILITY METRICS PREDICTION FOR CASE I

Metric	*Max. RMSE*	*Mean RMSE*	*Max. MAE*
RoCoF (Hz/s)	0.0021	0.0020	0.0195
Nadir (Hz)	0.0095	0.0093	0.0489

TABLE IV
ACCURACY OF THE NN MODEL FOR LOCATIONAL FREQUENCY STABILITY METRICS PREDICTION FOR CASE II

Metric	*Area*	*Max. RMSE*	*Mean RMSE*	*Max. MAE*
RoCoF (Hz/s)	Area 1	0.0224	0.0047	0.0168
	Area 2	0.0108	0.0057	0.0069
	Area 3	**0.0036**	**0.0033**	**0.0025**
Nadir (Hz)	Area 1	**0.0222**	**0.0205**	0.0159
	Area 2	0.0224	**0.0205**	**0.0156**
	Area 3	0.0225	0.0208	**0.0156**

Security-constrained results

Case I

- While all models converged 100% of the time, some caused violations
- **OPF** (with no security constraints) and **Numerical** lead to frequency violations

Case II

- Not all models converged 100% of the time
- Not all models secured the system
- Only the proposed *α-CROWN-Reg-KNN* converged 100% and led to 0 violations

TABLE IV
MODEL OPTIMISATION RESULTS FOR LOCATIONAL FREQUENCY STABILITY (NADIR (HZ)) FOR CASE I

s.t.	*Unstable OCs (%)*	*Convergence Rate (%)*
OPF	82.98	-
TDS-Opt	0.00	100
Numerical Estimation	19.35	100
Δ-OPF	**0.00**	100
α-CROWN-KNN-1	41.67	100
α-CROWN-Reg-KNN	**0.00**	100

TABLE V
MODEL OPTIMISATION RESULTS FOR LOCATIONAL FREQUENCY STABILITY (NADIR (HZ)) FOR CASE II

s.t.	*Unstable OCs (%)*			*Convergence Rate (%)*
	Area 1	*Area 2*	*Area 3*	
OPF	100	91.67	98.55	-
TDS-Opt	**0.00**	**0.00**	**0.00**	**100**
Numerical Estimation	1.39	76.39	97.10	**100**
Δ-OPF	**0.00**	**0.00**	**0.00**	63.89
α-CROWN-KNN-1	50.00	15.28	**0.00**	**100**
α-CROWN-Reg-KNN	**0.00**	**0.00**	**0.00**	**100**

TABLE VI
TIME REQUIRED BY DIFFERENT OPTIMISATION MODELS

s.t.	*Mean Optimisation Seconds*	
	Case I	*Case II*
OPF	0.05	4.84
TDS-Opt	57.39	154.33
Numerical Estimation	0.05	5.79
Δ-OPF	0.06	5.44
α-CROWN-KNN-1	1.08	21.80
α-CROWN-Reg-KNN	1.10	16.18

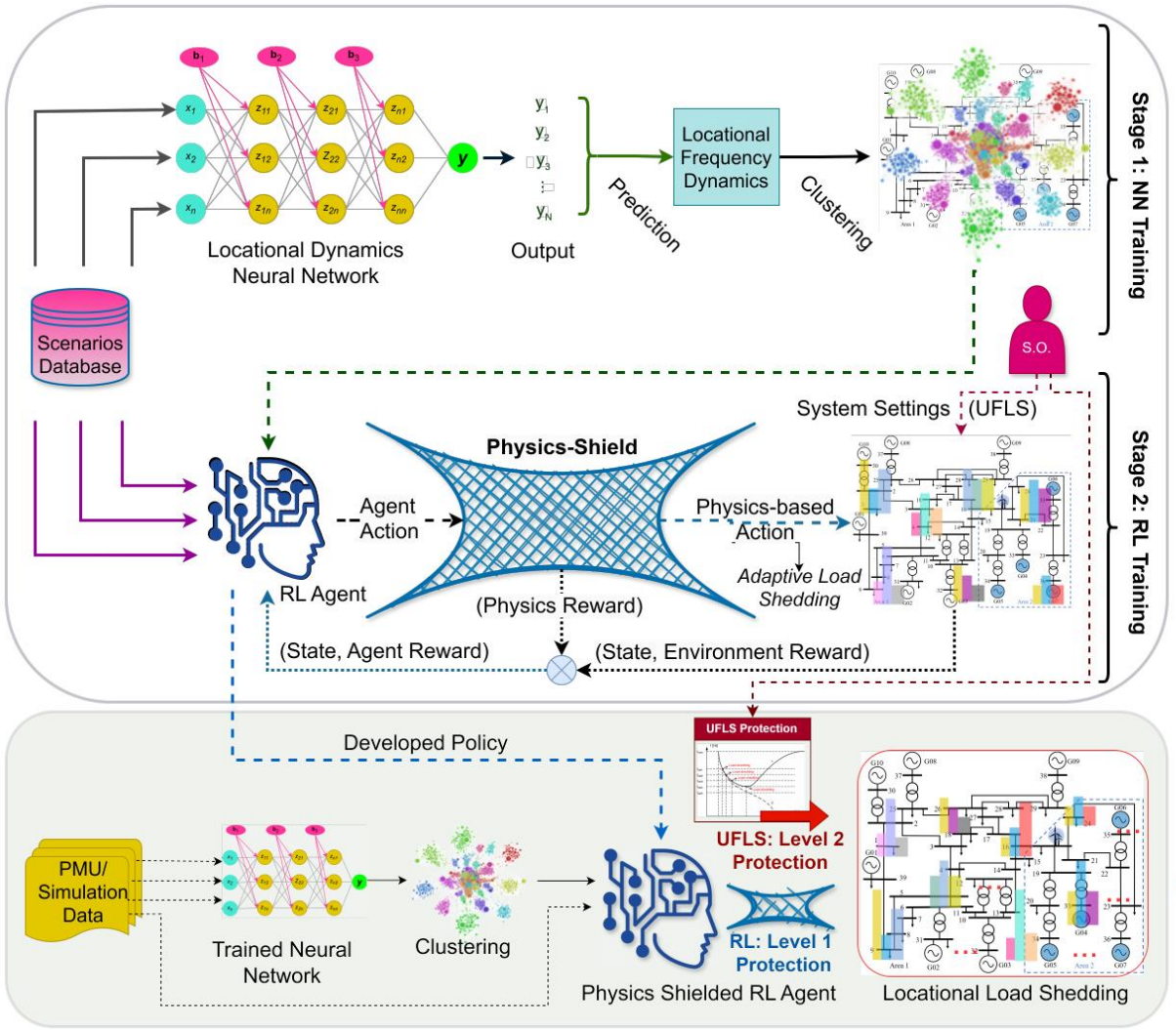
MANCHESTER 1824
The University of Manchester

HOW FAR CAN WE GO? - CONTROL AND AUTOMATION

Reinforcement Learning for emergency frequency control

Adaptive Load Shedding through Physics-Informed Deep Reinforcement Learning

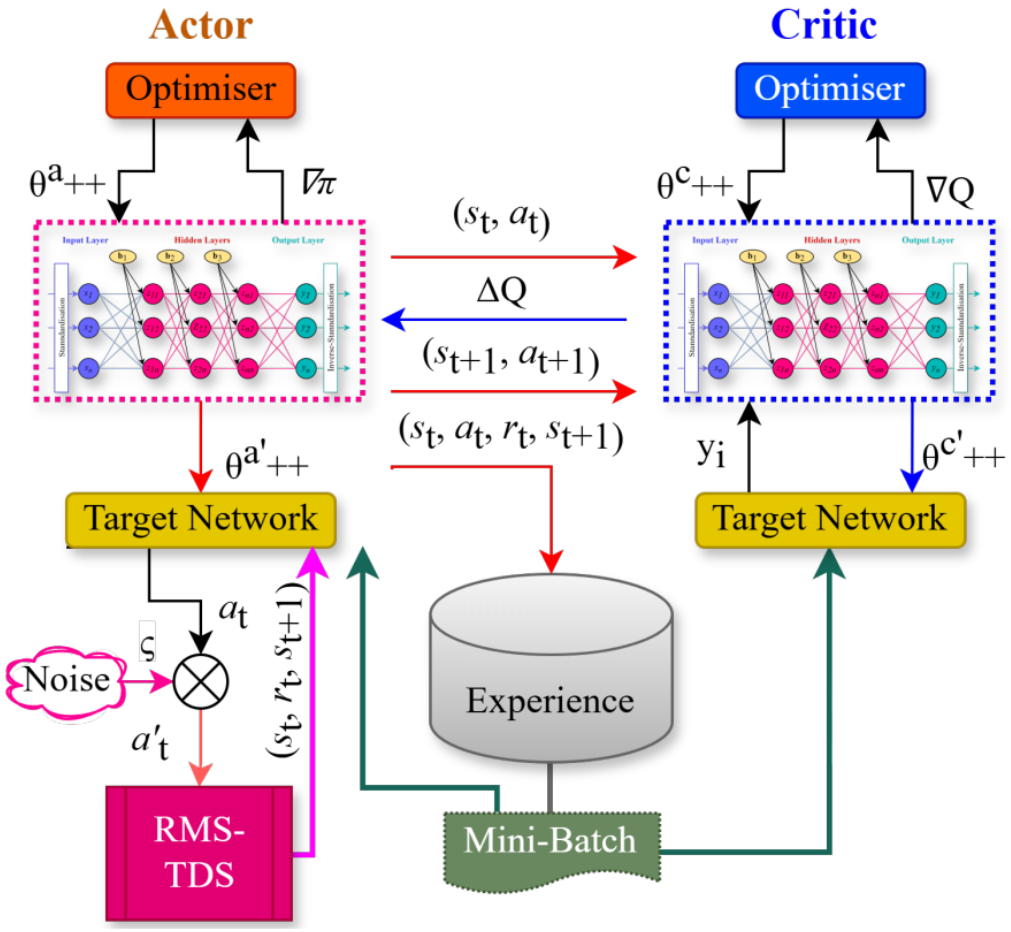
- Moving further from preventive securing of the system to real time emergency control
- Instead of pre-defined under-frequency load shedding settings, decide when, where and how much load to shed adaptively.
  - Utilising information from current operating condition (measurement)
  - Taking into consideration locational/regional frequency dynamics
- Physics Informed Reinforcement Learning (RL) **<u>to address scalability</u>**
  - Physics-Shield (swing equation based) and Neural Networks for coherent areas detection
  - Neural Networks and clustering for dimensionality reduction
  - Improves training and performance
- **<u>Complementary</u>** to Under-Frequency-Load-Shedding
  - UFLS can still be in operation as last resort



A. Kilembe, P. N. Papadopoulos, "*Adaptive Load Shedding using Physics-Informed Deep Reinforcement Learning,*" to be submitted.

DDPG algorithm

- Deep Deterministic Policy Gradient (DDPG) is an off-policy, model-free actor–critic reinforcement learning algorithm designed for continuous, multi-dimensional action spaces.
  - ***actor-critic:*** *uses two networks for action selection and value prediction*
  - ***off-policy:*** *learns from a replay buffer of past experiences*
  - ***continuous:*** *outputs continuous variables rather than discrete (e.g., MW, MVA, speed, etc.)*
  - ***multi-dimensional:*** *not limited to a single output*



MANCHESTER 1824
The University of Manchester

Physics shielded RL

- Embed the governing physics into the RL training process through a physics shield

$$\psi_{\max}^{\Omega_{\mathrm{PI}}} = |(\psi_{\max} - \Delta P_l)t^*| = \frac{2\sum_{i=1}^{N} s_i H_i}{f_n} \times \frac{\Delta f_{\mathrm{COI}}^{\min}}{\varphi_{\max}}$$

$$\psi_{\min}^{\Omega_{\mathrm{PI}}} = |(\psi_{\min} - \Delta P_l)t^*| = \frac{2\sum_{i=1}^{N} s_i H_i}{f_n} \times \frac{\Delta f_{\mathrm{COI}}^{\max}}{\varphi_{\min}}$$

- The physics shield activates whenever the agent's actions exceed swing equation limits

$$\psi_{\mathrm{RL}} \leftarrow \begin{cases} \psi_{\mathrm{RL}}, & \text{if } \psi_{\min}^{\Omega_{\mathrm{PI}}} \leq \psi_{\mathrm{RL}} \leq \psi_{\max}^{\Omega_{\mathrm{PI}}} \\ \psi_{\mathrm{RL}} \left(\min\left(1, \frac{\psi_{\max}^{\Omega_{\mathrm{PI}}}}{\psi_{\mathrm{RL}}}\right) \times \max\left(1, \frac{\psi_{\min}^{\Omega_{\mathrm{PI}}}}{\psi_{\mathrm{RL}}}\right) \right), & \text{otherwise} \end{cases}$$

- The physics shield reward is given by:

$$r_{\Omega_{\mathrm{PI}}} = \begin{cases} \phi_{\Omega\mathrm{PI}}(\psi_{\max}^{\Omega_{\mathrm{PI}}} - \psi_{\mathrm{RL}}), & \text{if } \psi_{\mathrm{RL}} \geq \psi_{\max,\Omega_{\mathrm{PI}}} \\ \phi_{\Omega\mathrm{PI}}(\psi_{\mathrm{RL}} - \psi_{\min}^{\Omega_{\mathrm{PI}}}), & \text{if } \psi_{\mathrm{RL}} \leq \psi_{\min,\Omega_{\mathrm{PI}}} \\ 0, & \text{otherwise} \end{cases}$$

Algorithm 1 NN-Physics-Shielded RL for AUFLS Emergency Frequency Control

Inputs: System operating condition: bus clustering labels (ρ), generator ratings ($\mathbf{s}_g$), generator active powers ($\mathcal{P}_g$), total system load scaling ($\mathcal{L}$)

Output: Load shedding amount (for specific loads in different locations) (ξ_i), and activation time (after the disturbance) of the AUFLS RL (t_{Act})

Stage 1: NN Training:

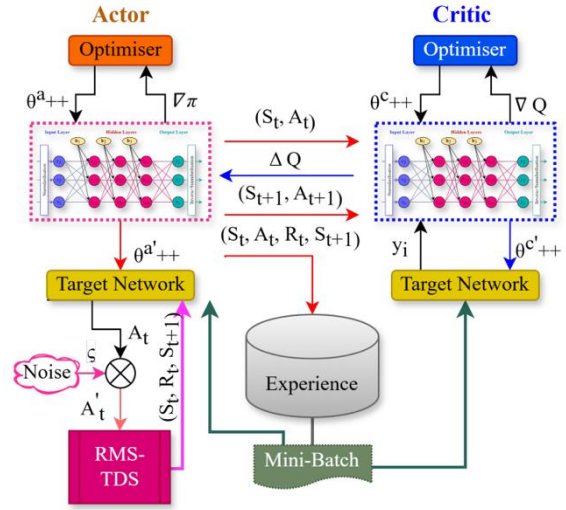
train neural network (NN) to predict bus-level RoCoF from pre-disturbance features
evaluate NN accuracy (RMSE and R^2)
if NN accuracy test passed, continue; **else** retune and retrain
initialise K-Means with $K \leftarrow 2$
while true:
  apply K-Means on predicted RoCoF; compute maximum intra-cluster inertia
  if intra-cluster inertia $>$ 5% of RoCoF threshold: $K \leftarrow K + 1$
  else: store K as the cluster number of the power system, **break**

Stage 2: RL Training:

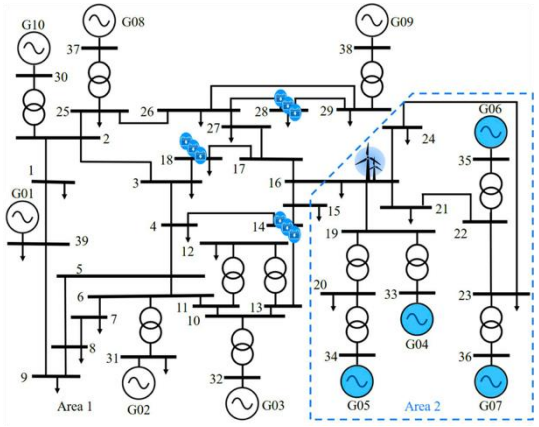
for state space, s, in training dataset **do**
  trained NN(s) predicts bus level RoCoF across PQ load buses of the system
  K-Means clusters the predicted RoCoFs into K coherent groups
  update $s \leftarrow$ bus indices labelled according to the coherence clusters
    for $j_{\mathrm{episode}} \in [1, \ldots, \mathcal{E}_0]$ episodes **do**
      if sampled exploration factor $\eta > \epsilon_{\mathrm{threshold}}$
        exploit agent's learned actions
      select action a based on the updated state s
      clip action $a +$ decaying Gaussian noise within [0,1]
      pass action a through Physics Shield Ω_{PI}
      get Physics Shield-based reward, $r_{\Omega_{\mathrm{PI}}}$, update a
      initialise RMS-TDS and apply the disturbance
      update total agent reward $r_{\mathrm{RL}} \leftarrow r_{\Omega_{\mathrm{PI}}} + r_f$
      store system transition in memory buffer, $\mathcal{G}$
      if frequency nadir within safe bounds, i.e., $s_{\mathrm{Metric}} < \mathcal{H}_{\mathrm{Hz}}^{\mathrm{limit}}$
        break (exit current episode loop)
    sample mini-batch $\mathcal{B}$ from $\mathcal{G}$, update $\theta^a, \theta^{a'}, \theta^c, \theta^{c'}$
    if s, a remain unchanged over Ψ interactions
      store scheme $[\xi_i, t_{\mathrm{Act}}], \forall i \in \mathcal{D}_{\mathrm{flexible\text{-}loads}}$
      break (exit full training loop)

Case Studies

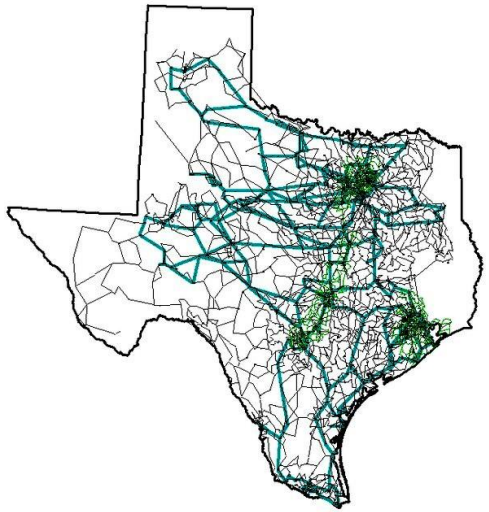
- The CIG penetration is up to **40%** of the system generation
- Steady state operating conditions are generated using the AC-OPF
- Time domain simulations in the RMS framework are performed for “worst case” N-2 contingency in PowerFactory
- RL parameters are determined through randomised sklearn-GridSearch
- Frequency nadir used as stability metric with a boundary at 57 Hz.



NN AND RL ALGORITHMS ARCHITECTURE AND HYPERPARAMETER SETTINGS



Modified IEEE 39 Bus Network



Synthetic Texas 2000 Bus Network

NN Model Parameter	NN-1 (IEEE 39-bus)	NN-2 (Texas 2000-bus)
Learning Rate	0.001	0.001
Batch Size, $\mathcal{B}$	64	64
Number of Layers	3	3
Number of Neurons	64	64
Activation, Θ	ReLU	Tanh
RL Model Parameter	**RL-1 (IEEE 39-bus)**	**RL-2 (Texas 2000-bus)**
Learning Rate	$1e^{-4}$	$1e^{-3}$
Batch Size, $\mathcal{B}$	128	128
Experience Replay	$1e^{5}$	$1e^{5}$
Number of Layers	3	3
Number of Neurons	100:128:64	100:256:128
Activation, Θ	ReLU & Sigmoid	ReLU & Sigmoid
Episodes, $\mathcal{E}$	200	250
Initial Exploration	0.5	0.5
Exploration Decay	0.9	0.9
Min. Exploration	0.01	0.01
Epsilon decay, ϵ_{decay}	1000	500
ϵ_{start} - ϵ_{end}	0.9 - 0.05	0.9 - 0.05

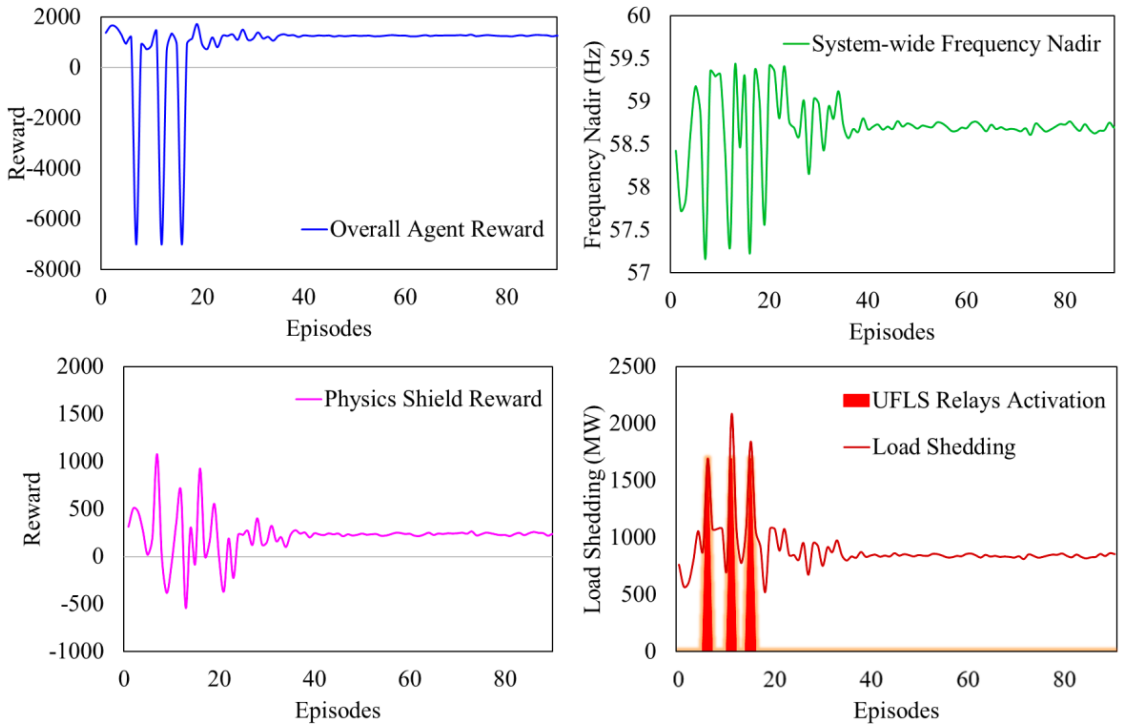
Results: Stage 1 and Training Performance

PERFORMANCE EVALUATION OF ML MODEL FOR ROCOF PREDICTION IN THE IEEE 39-BUS NETWORK

RMSE (Min)	RMSE (Mean)	RMSE (Max)	R^2 (Training / Validation)
0.0022	0.0053	0.0089	0.9976 / 0.9661
Train Time (Secs.): 45.52			

PERFORMANCE EVALUATION OF ML MODEL FOR ROCOF PREDICTION IN THE TEXAS 2000-BUS NETWORK

RMSE (Min)	RMSE (Mean)	RMSE (Max)	R^2 (Training / Validation)
0.0045	0.0063	0.0082	0.9960 / 0.9941
Train Time (Secs.): 150.79			



AUFLS training progress of the RL physics-shielded agent in the Modified IEEE 39-bus network

Results: Stage 2 and Shedding Performance

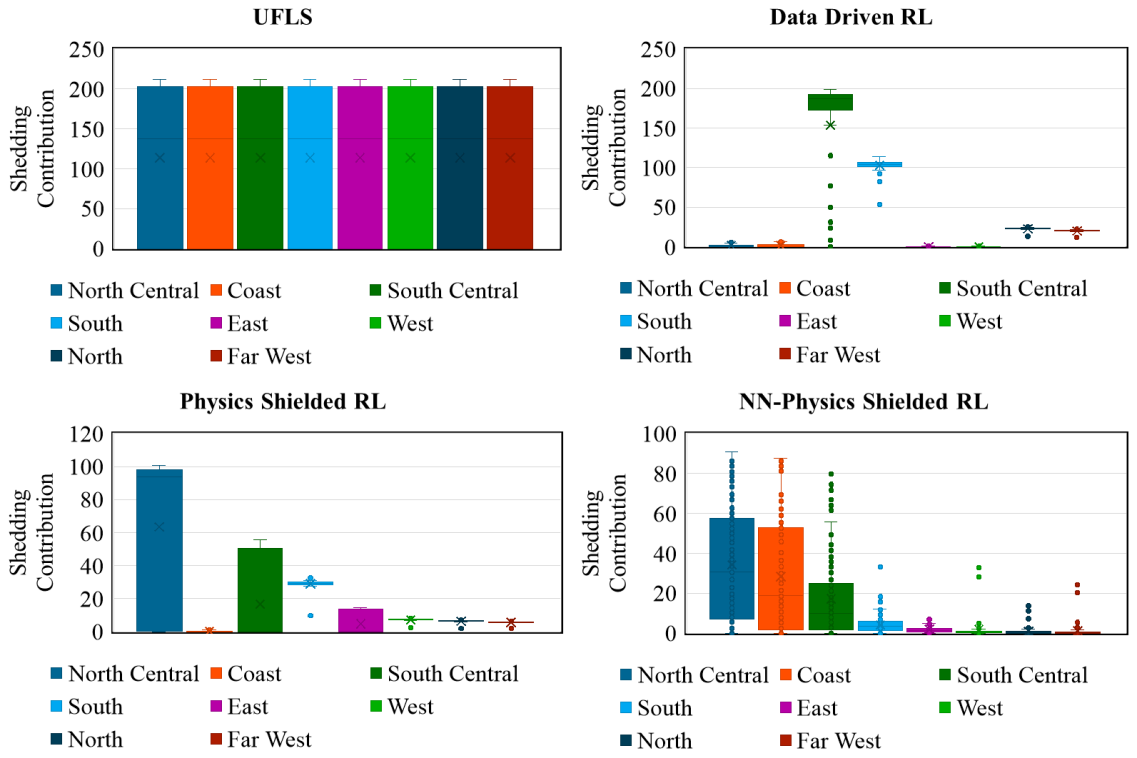
MEAN LOAD SHEDDING (MW) PERFORMANCE BY DIFFERENT SCHEMES: IEEE 39 BUS NETWORK

Control Model	Mean Nadir (Hz)	Mean Load (MW)	Train time
UFLS Scheme	58.28	971.40	-
Data-Driven RL	58.67	692.47	190 mins.
Physics-Shielded RL	58.43	627.63	129 mins.
NN-Physics-Shielded RL	58.25	588.63	125 mins.

MEAN LOAD SHEDDING (MW) PERFORMANCE BY DIFFERENT SCHEMES: TEXAS 2000 BUS NETWORK

Control Model	Mean Nadir (Hz)	Mean Load (MW)	Train time
UFLS Scheme	59.53	598.22	-
Data-Driven RL	59.63	299.09	700 mins.
Physics-Shielded RL	59.62	127.94	540 mins.
NN-Physics-Shielded RL	59.63	90.578	455 mins.

- Reduced load shedding in all cases
- Scalability for large-scale systems



Shedding Performance in the Texas 2000 Bus Network

MANCHESTER 1824
The University of Manchester

TOPOLOGY AND PROPAGATION OF DYNAMICS

Graph based methods considering dynamics to provide insights and deal with topology changes

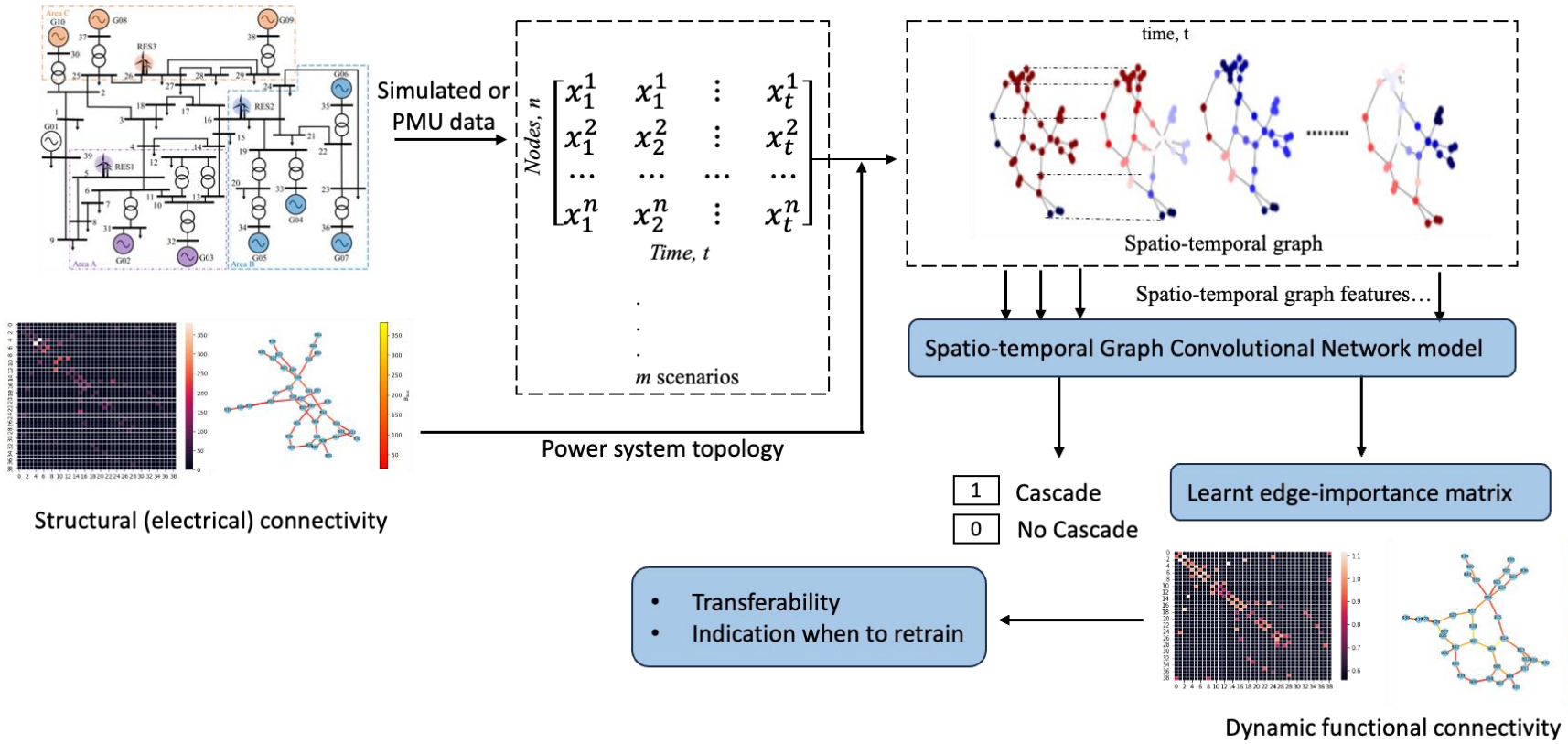
MANCHESTER 1824
The University of Manchester

Graph-based approaches and machine learning

- Power systems can be naturally represented by graphs
- Planned/unplanned topology changes can occur
  - A challenging problem when training machine learning models due to its combinatorial nature
- Graph based methods can help
  - Improve performance
  - Improve generalization capability and transferability
  - Provide insights into vulnerable regions
- Further to that, we can also provide indications of when re-training might be needed under topological changes
- An important aspect relates to taking into account aspects related to how dynamic phenomena propagate in graph-based approaches

Dynamic Functional Connectivity Graph – considering dynamics in graph-based approaches

- Graphs representation usually based on static connectivity
- For complex phenomena (e.g. cascading failures), dynamics play an important role
- A spatio-temporal approach: We can re-weight a graph by learning connections related to dynamics
- Gain useful insights on how dynamic phenomena propagate or influence the task at hand (e.g. cascade or instability prediction)
- Speed is a key advantage of overall framework, assessment (cascade happening or not) in <1ms



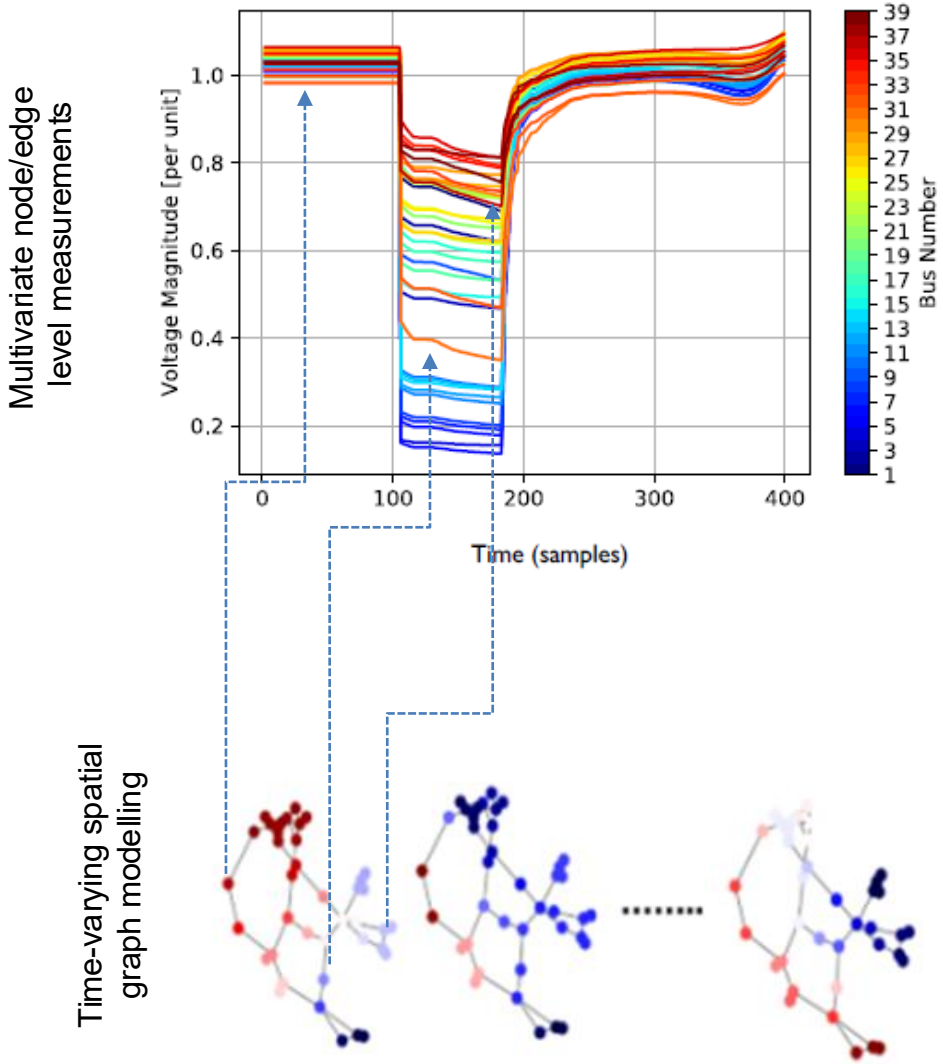
[1] T. Ahmad, Y. Zhu, P. Papadopoulos, “Predicting Cascading Failures in Power Systems using Graph Convolutional Networks,” NeurIPS 2021 Workshop on Tackling Climate Change with Machine Learning, https://www.climatechange.ai/papers/neurips2021/76.

[2] T. Ahmad, P. N. Papadopoulos, " Dynamic functional connectivity graph for assessing cascading events in power system,”, PSCC 2024, Electric Power Systems Research, Vol. 235, 2024, 110724, https://doi.org/10.1016/j.epsr.2024.110724.

Spatio-temporal Power System Graphs

- Conventional power system graph --- weighted, undirected, spatial graph $\mathcal{G}\ (\mathcal{V}, \mathcal{E})$, $\mathcal{V}$- nodes (buses including generators & shunt elements), $\mathcal{E}$ - edges (transmission lines/transformers),
- Spatio-temporal (ST) graphs constructed using time-domain trajectories of power system variables, at spatial locations indexed by spatial graph.
  - Spatial edges: $E_S = \{v_{ti}v_{tj} | (i,j) \in \mathbb{N}\}$
  - Temporal edges: $E_T = \{v_{ti}v_{(t+1)i}\}$

Spatio-temporal power system graph



MANCHESTER 1824
The University of Manchester

Spatio-temporal Graph Convolutional Networks for cascading event prediction

- **Step 1:** Creation of spatio-temporal (ST) graphs using time-varying voltage magnitudes, $V^{mag}(t) = v_{t,i}^{mag} | t = 1,2,\dots,T; i = 1,\dots N.$

- **Step 2:** Creation of ST neighbourhood:

$$\Lambda(v_{ti}) = \{v_{q,j} | e(v_{tj}, v_{ti}) \leq S\}, |q - t| \leq \lfloor \Gamma/2 \rfloor$$

- **Step 3**: ST-GC operation on node, v_{ti} with respect to convolutional kernel, $w(.)$ normalisation factor, $\mathcal{N}_{ti}$

$$f_{out}(v_{ti}) = \frac{1}{\mathcal{N}_{ti}} \sum_{v_{qj} \in \Lambda(v_{ti})} f_{in}(v_{qj}).w(v_{qj})$$

Spatial kernel - $W_{SG} \in R^{C\ X\ P}$ Temporal kernel - $W_{TG} \in R^{M\ X\ T}$

- **Step 4:** Spatial graph convolution at time t, w.r.t to symmetrically normalised Laplacian:

$$f_t' = D^{-\frac{1}{2}} \tilde{A} D^{-\frac{1}{2}} f_t W_{SG}$$

$f_t \in R^{N\ X\ C}$ denotes C types of input features for N nodes at the t^{th} time – frame

$f_{t'} \in R^{N\ X\ P}$ - P output features

This matrix is shared across all $st-GCN$ layers by replacing $(\tilde{A})$ by $(\tilde{A})$x M (element-wise multiplication)
M – edge-importance matrix and trained as a parameter.

- **Step 5:** Temporal convolution (1-D): $f_t' \odot W_{TG} \in R^{T\ X\ T}$
- **Step 6**: Global average pooling and output vector transformed to class probabilities by a fully connected SoftMax layer with a sigmoid activation.

Dynamic Functional Connectivity graph

- *Edge-importance matrix, M* takes into account both topological information and dynamic behaviour of power systems.
- Improves performance for classifying scenarios leading to cascades
- M-matrix can also be projected as a dynamic functional connectivity (DFC) specific to cascading events.
- Constrained to be positively weighted and symmetric.
- The fully connected graph (corresponding to M) used to derive a sparse graph (similar to topological power system graph but re-weighted) using the *κ-neighbourhood sparsification* scheme.

$$DFC = \begin{cases} m_{ij} & m_{ij} < \kappa \\ m_{ij} = 0 & m_{ij} \geq \kappa \end{cases}$$

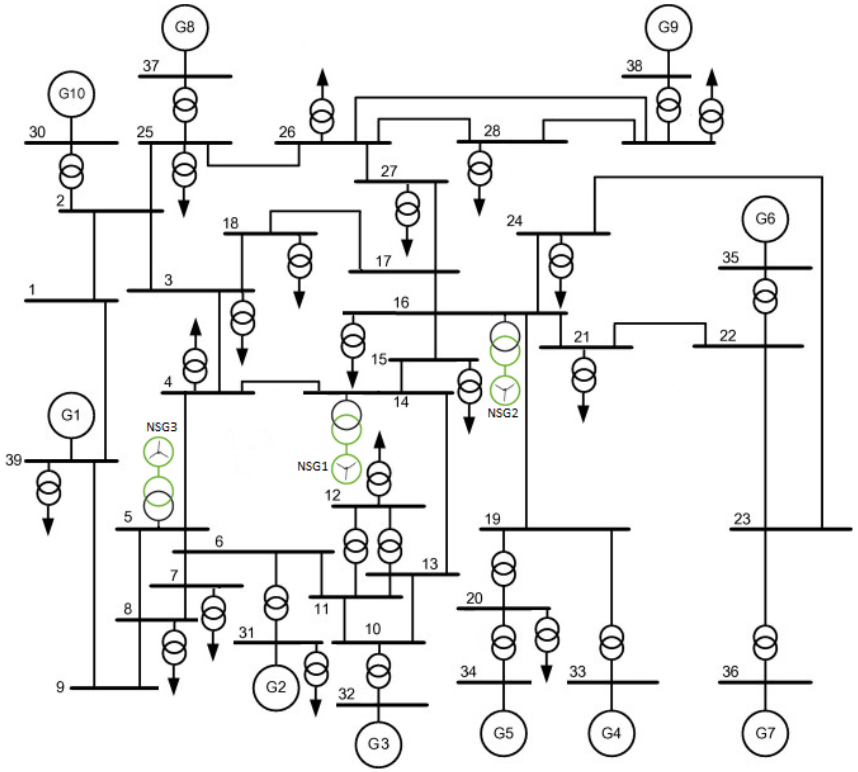
Complex Network theory-based centrality measures

Centrality	Calculation
Degree	$C_d(v) = \frac{\|\mathcal{L}(v,j)\|}{n-1}$
Betweenness	$C_b(v) = \frac{\sum_{i \neq v \neq j \in V} \sigma_{ij}(v)/\sigma_{ij}}{(n-1)(n-2)/2}$
Closeness	$C_c(v) = \frac{\sum_{j \in V \setminus v} d(v,j)}{n-1}$, $C_{c'}(v) = \frac{n-1}{\sum_{j \in V \setminus v} d(v,j)}$
Eigen-vector	$C_e(v) = \frac{1}{\lambda_{max}} \sum_{j=1}^{n} A(v,j) x_j$

- Centrality indices commonly used to understand and identify structural conditions that favours an edge or node to affect the behaviour of other elements.
- Node-centrality measures, calculated for both DFC and Y-bus based graphs, respectively and used to derive importance rank of various buses.

Case studies - Modelling of cascading events

Modified IEEE 39-bus



- Type 4 wind generation
- Protection devices
  - Synchronous generators
    - Over/under-speed
    - Pole-slip
    - Under-voltage
    - Over-excitation limiter
  - Non synchronous generation
    - Over/Under-voltage
    - Over/Under-frequency
  - Loads
    - UFLS (4 stages)
- Other relevant devices
  - AVRs, PSS, Governors, wind with FRT, (secondary f control)
  - Load tap changers

Example rules

If

- V<V1pu for t>t1ms

or

- V<V2pu for t>t2ms

or

- f<f1pu for t>t3ms

trip

[1] G. A. Nakas, P. N. Papadopoulos, “Investigation of the Impact of Load Tap Changers and Automatic Generation Control on Cascading Events”, PowerTech 2021, Madrid, Spain (online), June 28 – July 2, 2021.

[2] G. A. Nakas, P. N. Papadopoulos, “Investigation of Cascading Events in Power Systems with Renewable Generation”, ISGT Europe 2020, Delft, Netherlands (online), 26-28 October 2020.

Case studies – an example of a cascading event

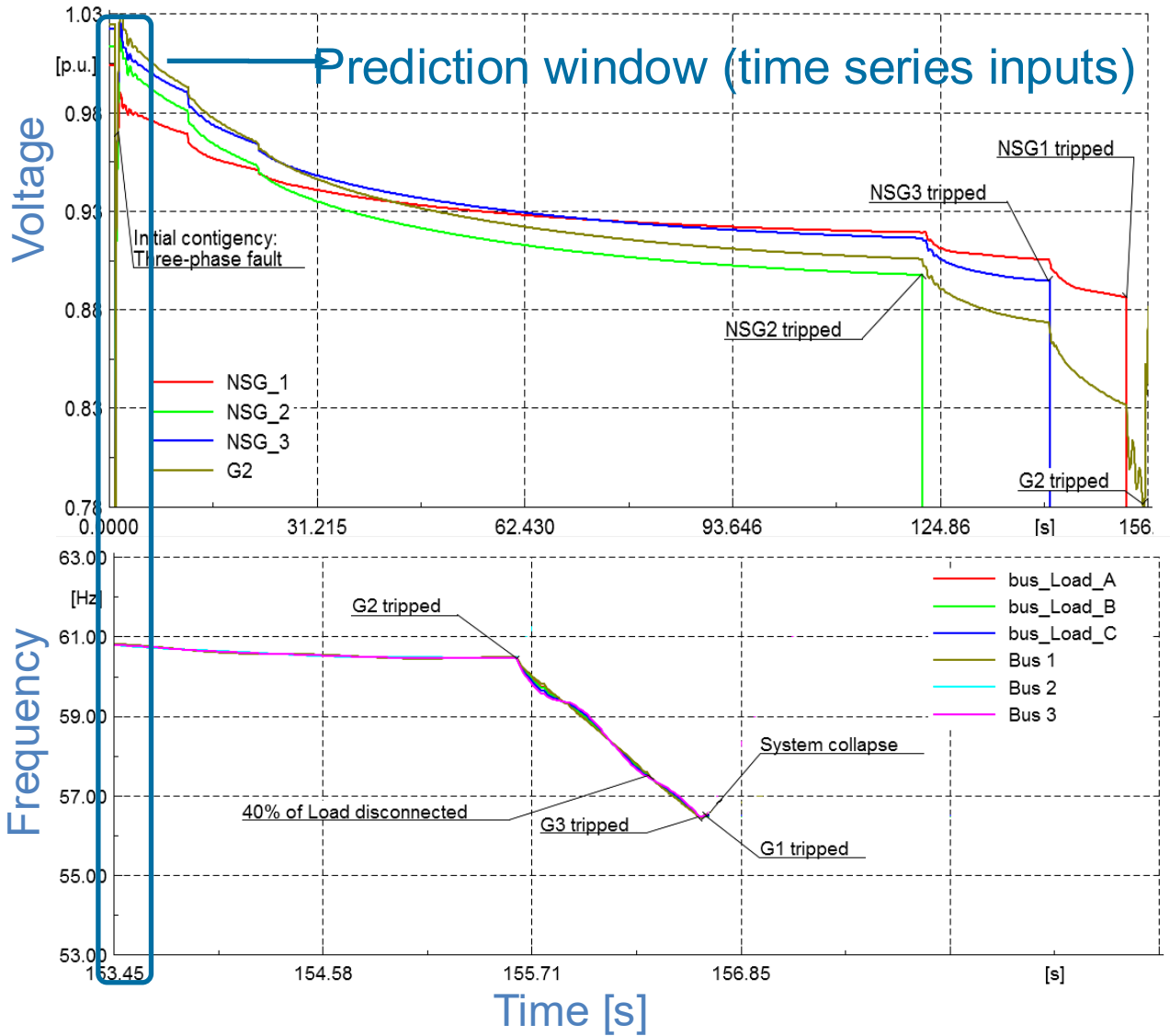
The University of Manchester

1st event 2nd event 3rd event…

[(NSG2-UV), (NSG3-UV), (NSG1-UV), (G2-UV), (Load_C-UF), (Load_A-UF), (Load_B-UF), (G3-US), (G1-US)]

Component that trips

Reason for tripping



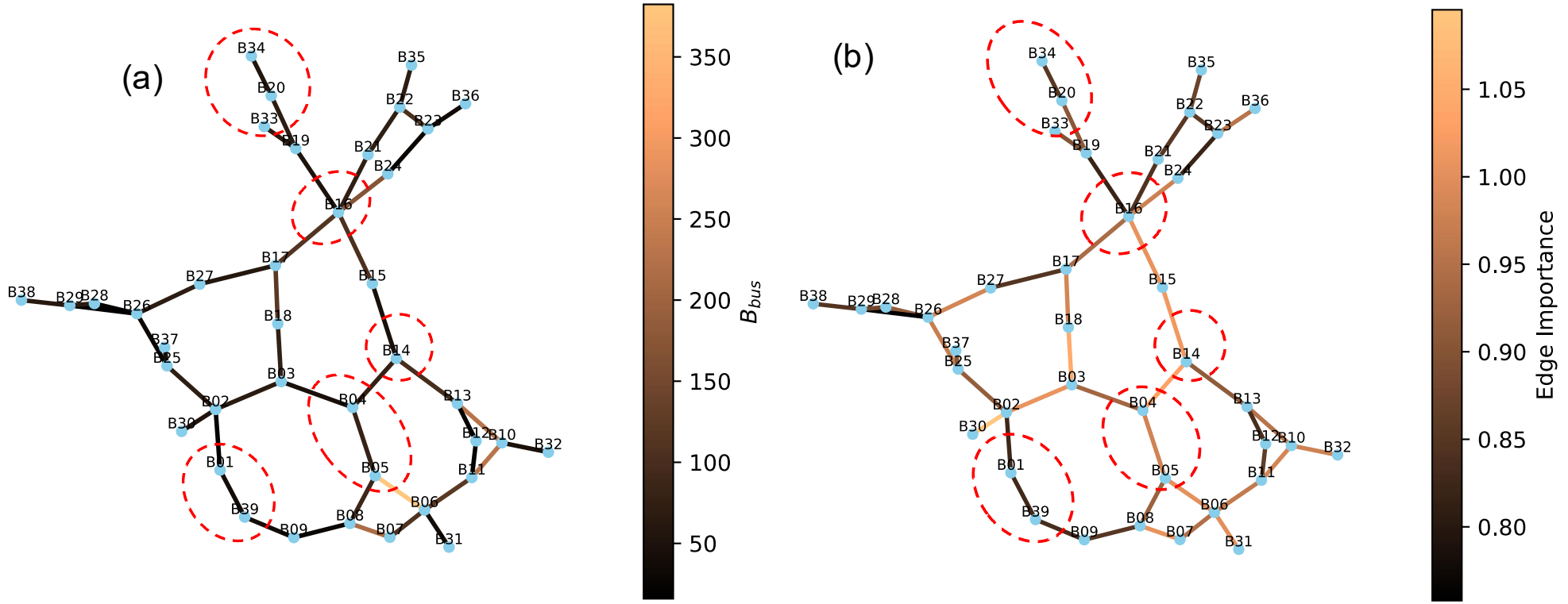
Time of Event (s)	Event Description
1	Three – Phase Fault on 90% of Line 5-7
1.07	Line 5-7 trips by Line Distance Protection Relays
122.26	NSG2 disconnected by UV Relay
141.30	NSG3 disconnected by UV Relay
152.80	NSG1 disconnected by UV Relay
155.61	G2 disconnected by UV Relay
156.34	40% of Load C disconnected by UF Relay
156.34	40% of Load A disconnected by UF Relay
156.34	40% of Load B disconnected by UF Relay
156.62	G3 disconnected by Under – Speed Relay
156.65	G1 disconnected by Under – Speed Relay

[1] G. Nakas et al., “Online Identification of Cascading Events in Power Systems With Renewable Generation Using Measurement Data and Machine Learning," IEEE Access, vol. 11, pp. 72343-72356, 2023, doi: 10.1109/ACCESS.2023.3294472. https://ieeexplore.ieee.org/document/10179902.

[2] G. Nakas, P.N. Papadopoulos, “Online Identification of Cascading Event sequences in Power Systems using Deep Learning” , under 1st round of revision.

Admittance-based vs Dynamic Functional Connectivity Graph

(a) Ybus based graph (b) DFC graph



- Dashed circles: buses/lines with most frequently tripped components.
- In DFC strong links surround those buses where highest number of cascading events take place – important information in terms of cascade propagation and dynamics
- *E.g., B16* (connected to *NSG2*) and *B30* (connected to *G1*), are strongly connected with higher edge-weight paths in the DFC graph

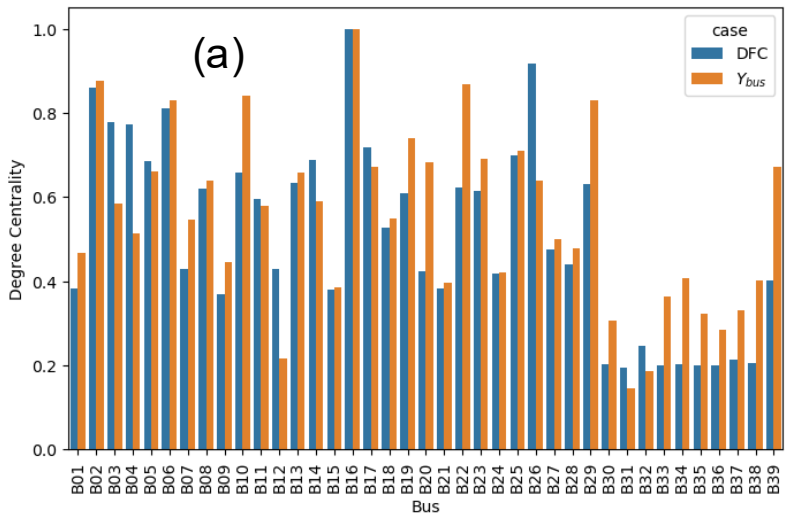
Event sequence	#pattern appeared
[(“NSG2”*, “OverVoltage”)]	2219
[(“NSG2”, “OverVoltage”), (“G1”**, “Out of step”)]	336
[(“NSG2”, “OverVoltage”), (“NSG1”, “OverVoltage”), (“G1”, “Out of step”)]	279
[(“NSG2”, “OverVoltage”), (“G1”, “Under-Speed”), (“NSG3”, “UnderVoltage”)]	243

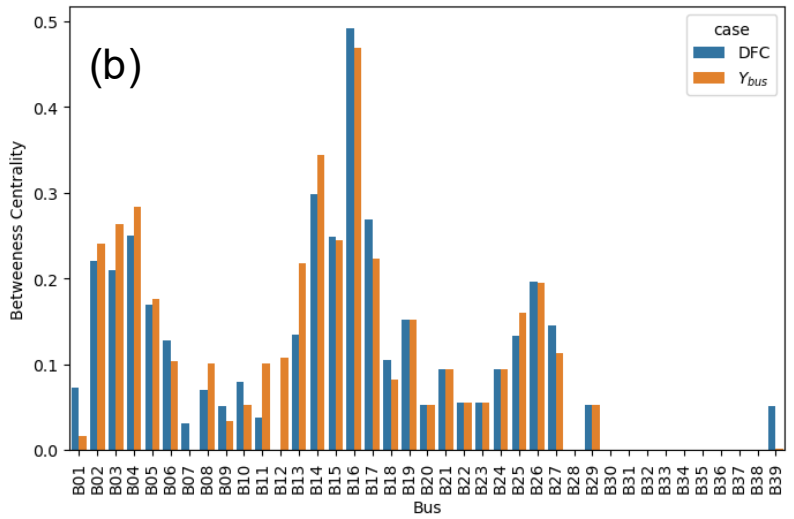
* NSG1, NSG2, NSG3 represent wind generators, ** G1, G4, G5 represent synchronous generators.

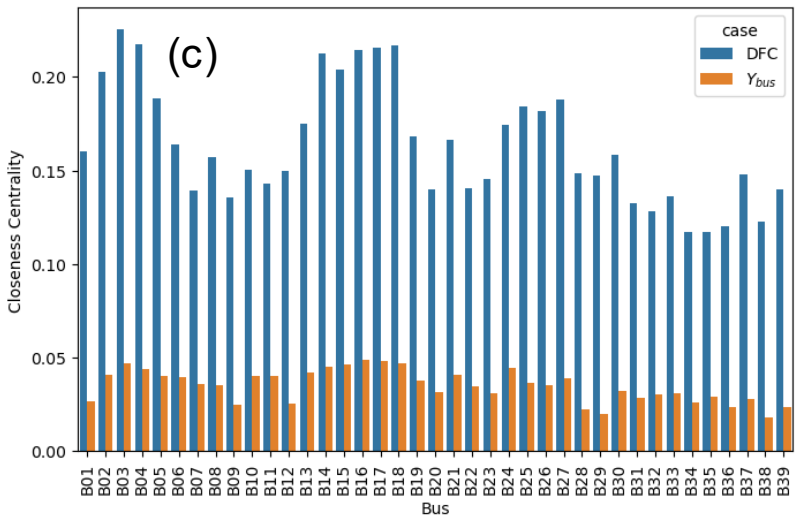
T. Ahmad, P. N. Papadopoulos, " Dynamic functional connectivity graph for assessing cascading events in power system,” Electric Power Systems Research, Vol. 235, 2024, 110724, https://doi.org/10.1016/j.epsr.2024.110724.

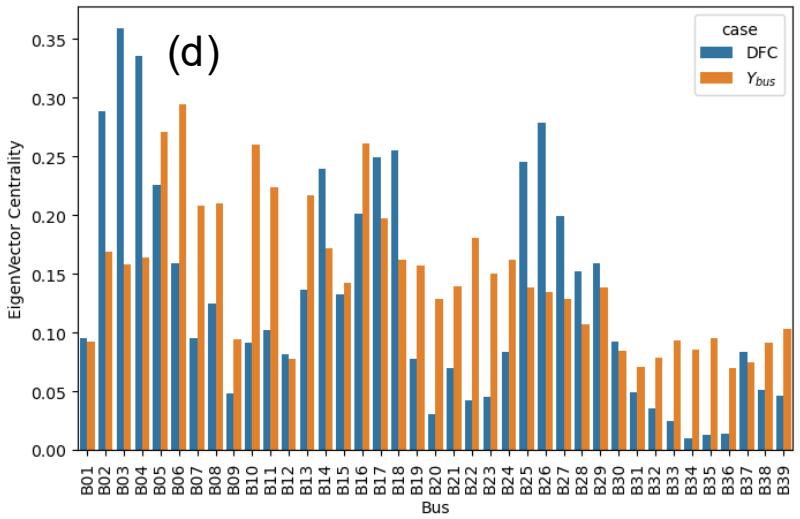
Centrality indices

Centrality measures for DFC and Ybus based graph (T to B) (a) degree centrality (b) Betweenness centrality (c) Closeness centrality (d) Eigen-vector centrality







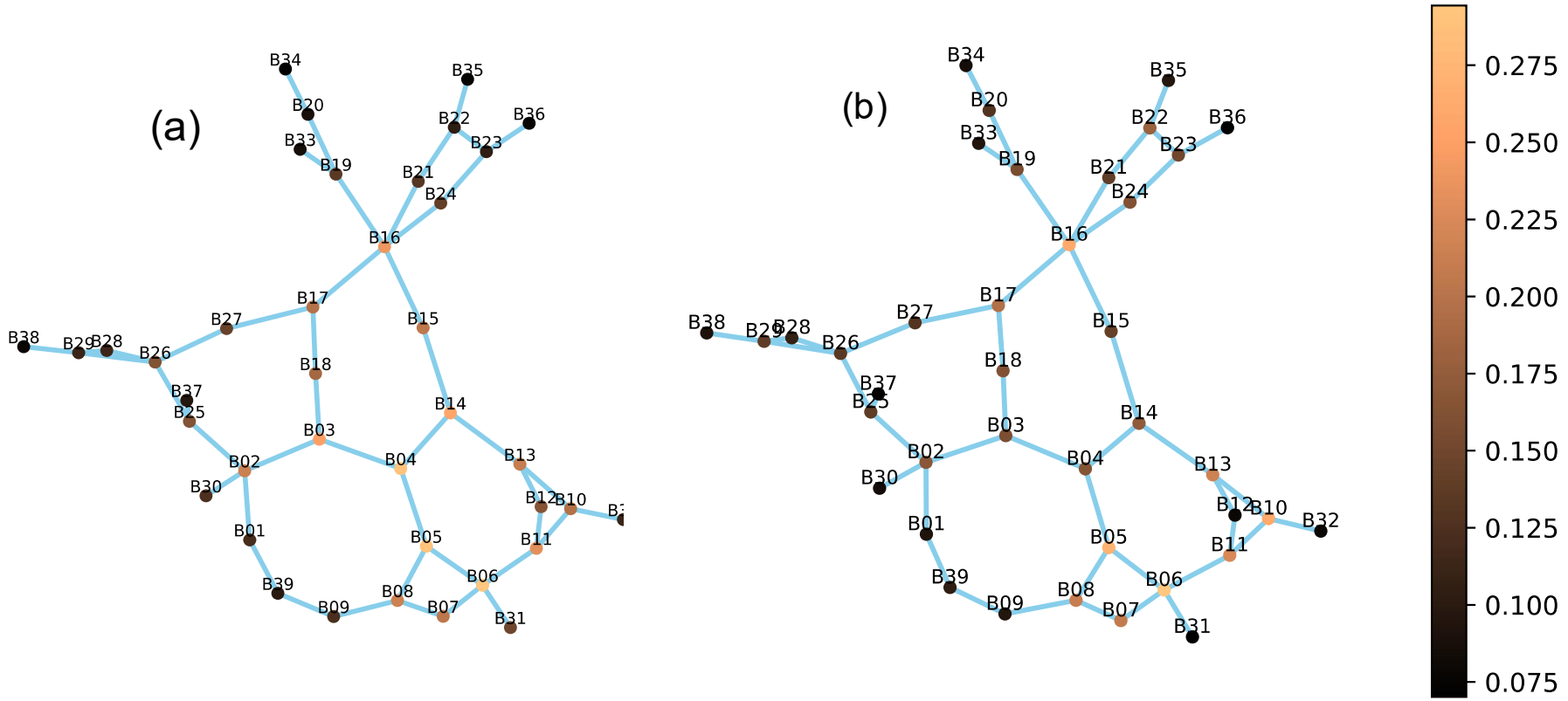


- For DFC graph a large amount of degree centrality can be shifted into a small number of nodes in the system.
- The eigen-vector centrality of the DFC graph shows a distribution similar to degree centrality, but a sparser one.

T. Ahmad, P. N. Papadopoulos, " Dynamic functional connectivity graph for assessing cascading events in power system," Electric Power Systems Research, Vol. 235, 2024, 110724, https://doi.org/10.1016/j.epsr.2024.110724.

Eigen-vector centrality

Eigen-vector centrality (a) DFC graph (b) Ybus based graph



- Eigen-vector centrality of the DFC graph exhibits distinct clusters, typical of spectral clustering observed due to the dynamic phenomena
- On the contrary, eigen-vector centrality of Y-bus based graph, is dispersed “flatly" across nodes
- Useful to discover latent clusters for the power system, with respect to their vulnerability to cascades!

T. Ahmad, P. N. Papadopoulos, " Dynamic functional connectivity graph for assessing cascading events in power system,” Electric Power Systems Research, Vol. 235, 2024, 110724, https://doi.org/10.1016/j.epsr.2024.110724.

Conclusions

- Challenges related to changing nature of power system dynamics
  - New types of dynamic phenomena
  - Representation and modelling of dynamics at transmission/distribution level
  - Increasing uncertainty, complexity and computational effort
- Probabilistic small-signal analysis
  - Investigate how dynamic interactions behave across a range of operating conditions
- Nonlinear dynamics

- Machine learning methods can be useful
  - Fast assessment (100s times speed up)
  - Capture complex dynamics and topological aspects
- Moving away from the black-box paradigm for machine learning through explainability/interpretability
  - Increase trust and confidence in ML applications
  - Provide useful insights into complex dynamics
- A three-step approach to implementation
  - i) fast assessment and improved situational awareness, ii) decision support, iii) automation

Contact details

Reader (Associate Prof.) and UK Research and Innovation Future Leaders Fellow,
The University of Manchester

Email: panagiotis.papadopoulos@manchester.ac.uk
Linkedin: https://www.linkedin.com/in/panagiotis-papadopoulos-92b32a27/
Website: https://research.manchester.ac.uk/en/persons/panagiotis-papadopoulos

MANCHESTER 1824
The University of Manchester

DEFINING SYSTEM STRENGTH WHILE CONSIDERING DYNAMICS

Small signal variability as a measure of system strength

L. Benedetti, P. N. Papadopoulos and A. Egea-Àlvarez, "A Modal Contribution Metric for Quantifying Small-Signal Variability in Power Systems With Converter-Interfaced Generation," in *IEEE Transactions on Power Systems (Early Access)*, 2024. doi: 10.1109/TPWRS.2024.3500786.

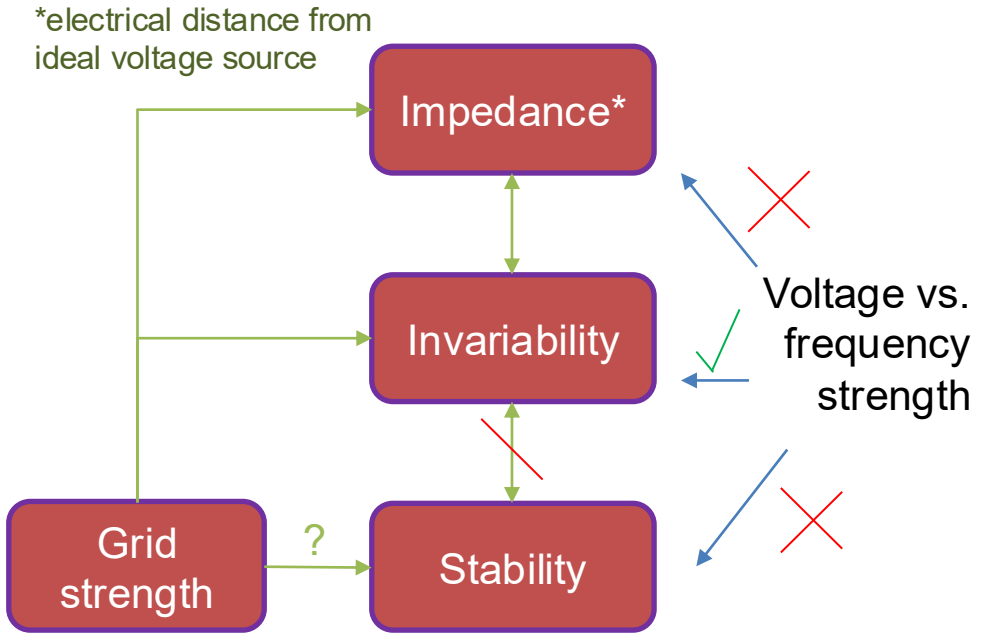
Dataset available:

L. Benedetti, A. Egea-Alvarez, and P. Papadopoulos, “Results data: A modal contribution metric for quantifying small-signal variability in power systems with converter-interfaced generation,”figshare, Jul 2024, doi:10.48420/26412331. [Online]. Available: https://figshare.manchester.ac.uk/articles/dataset/Results_Data_A_Modal_Contribution_Metric_for_Quantifying_Small-Signal_Variability_in_Power_Systems_with_Converter-Interfaced_Generation/26412331

Small-signal *variability* and system "strength"

- A lot of commonly used system strength metrics are static in nature and neglect dynamics
- Recent approaches do consider dynamics (impedance-based approaches), mostly linking strength with stability margins
- Focus on how much a system variable (e.g. voltage or frequency) changes?
  - For a given disturbance in one location, how does voltage/frequency change throughout the network?
  - Similar to traditional metrics but considering dynamics
    - Not necessarily directly linked to stability but to how much output variables change/deviate
  - Decouples strength with respect to voltage and frequency
  - Uses modal responses to output variables and calculates maximum deviation
  - Captures how new types of interactions influence variability

Different perspectives of grid "strength"



Dataset available:
https://figshare.com/articles/dataset/Results_Data_A_Modal_Contribution_Metric_for_Quantifying_Small-Signal_Variability_in_Power_Systems_with_Converter-Interfaced_Generation/26412331

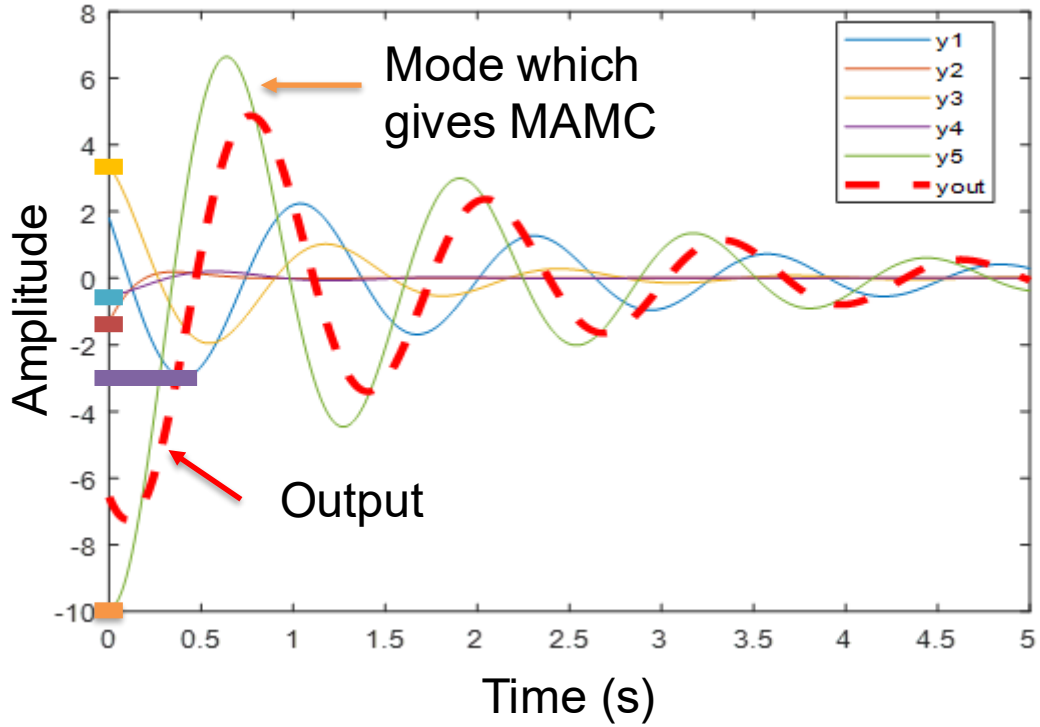
L. Benedetti, P. N. Papadopoulos, A. Egea, "A Modal Contribution Metric for Quantifying Small-Signal Variability in Power Systems with Converter-Interfaced Generation," available online, IEEE Transactions on Power Systems, https://pure.manchester.ac.uk/ws/portalfiles/portal/349631696/AModalContributionMetricforQuantifyingSmallSignalVariability.pdf .

Proposed methodology and quantification metric

Links specific modes/interactions to the variability of the output

- Modal contribution:
  - Maximum (absolute) deviation of the *decoupled modal response:* related to eigenvalues and eigenvectors
- Metric for small-signal variability of output: ***Maximum Absolute Modal Contribution (MAMC)***

Time-domain response of a linear system can be separated into a series of decoupled modal responses



$$\chi(t) = Ae^{\sigma t}cos(\omega t + \theta)$$

$$A = 2|\Phi_{j,n}c_n|$$

$$\theta = \angle\Phi_{j,n}c_n$$

$$\lambda_n = \sigma \pm j\omega$$

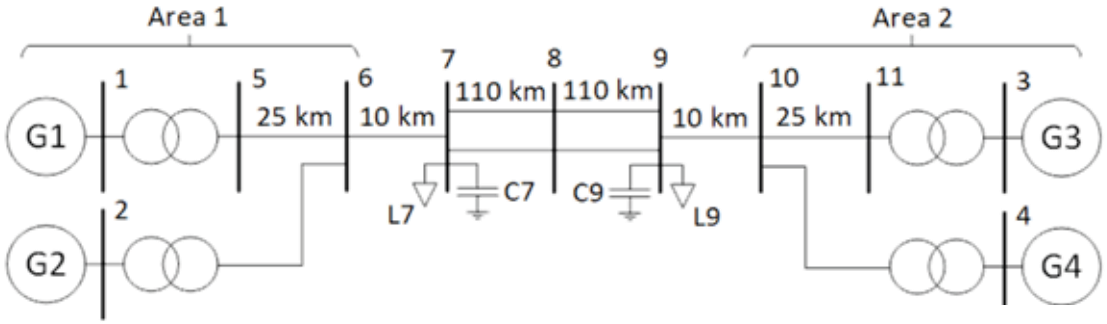
$$t_{max} = \frac{n\pi - \left(\theta + jlog\left(\pm\frac{\sqrt{-\sigma^2-\omega^2}}{\sigma+j\omega}\right)\right)}{\omega}$$ or 0 s

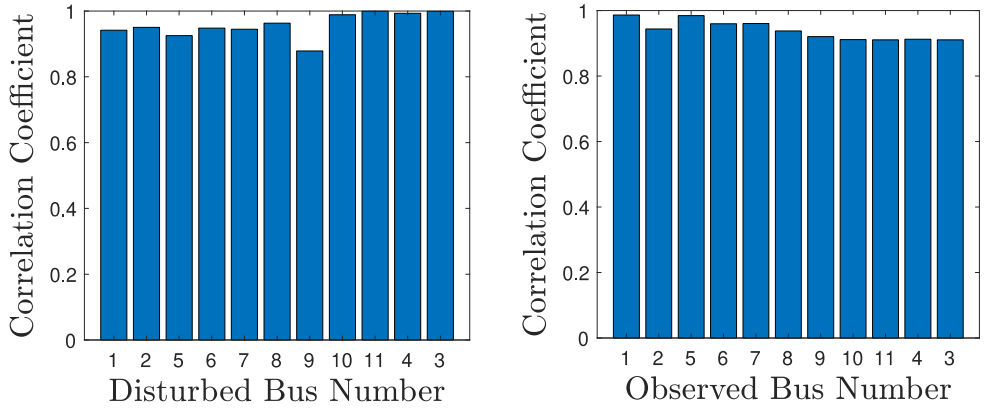
$$\Delta y_{j,n}^{max} = 2\left|\Phi_{i,n}c_n\right| e^{\sigma t_{max}} cos\left(\omega t_{max} + \angle\Phi_{i,n}c_n\right)$$

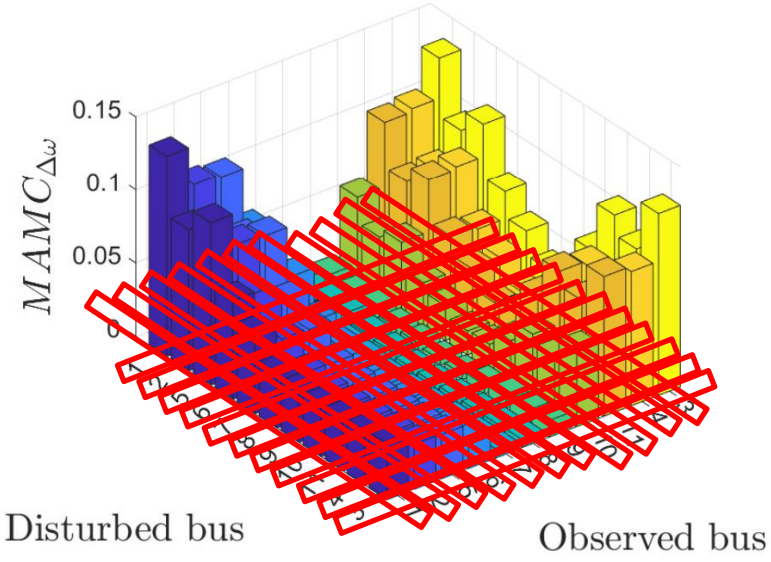
Validation

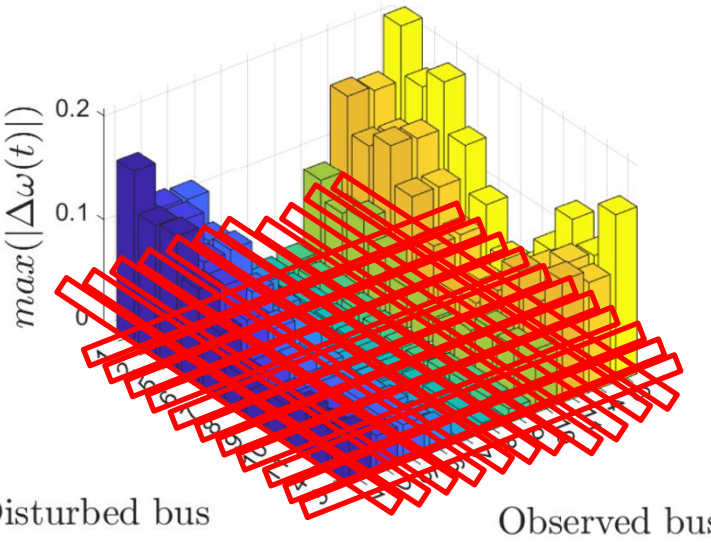
SMIB system (SG classical model)
Single mode system => MAMC should match the maximum deviation *exactly*

Input / Output	$MAMC$	Time-Domain Maximum (TDM)	Error
ω_{ib} / ω_r	5.5493 rad/s	5.5493 rad/s	0
ω_{ib} / T_e	292.3498 pu	292.3498 pu	0
$V_{m,ib}$ / ω_r	0.9044 rad/s	0.9044 rad/s	0
$V_{m,ib}$ / T_e	34.3085 pu	34.3085 pu	0









- Trends mostly captured in base case
- Generally good matches: Pearson's correlation coefficient of 0.879 or above
- Slight reduction (proportionally to other areas) of disturbance in area 1 to observation in area 2

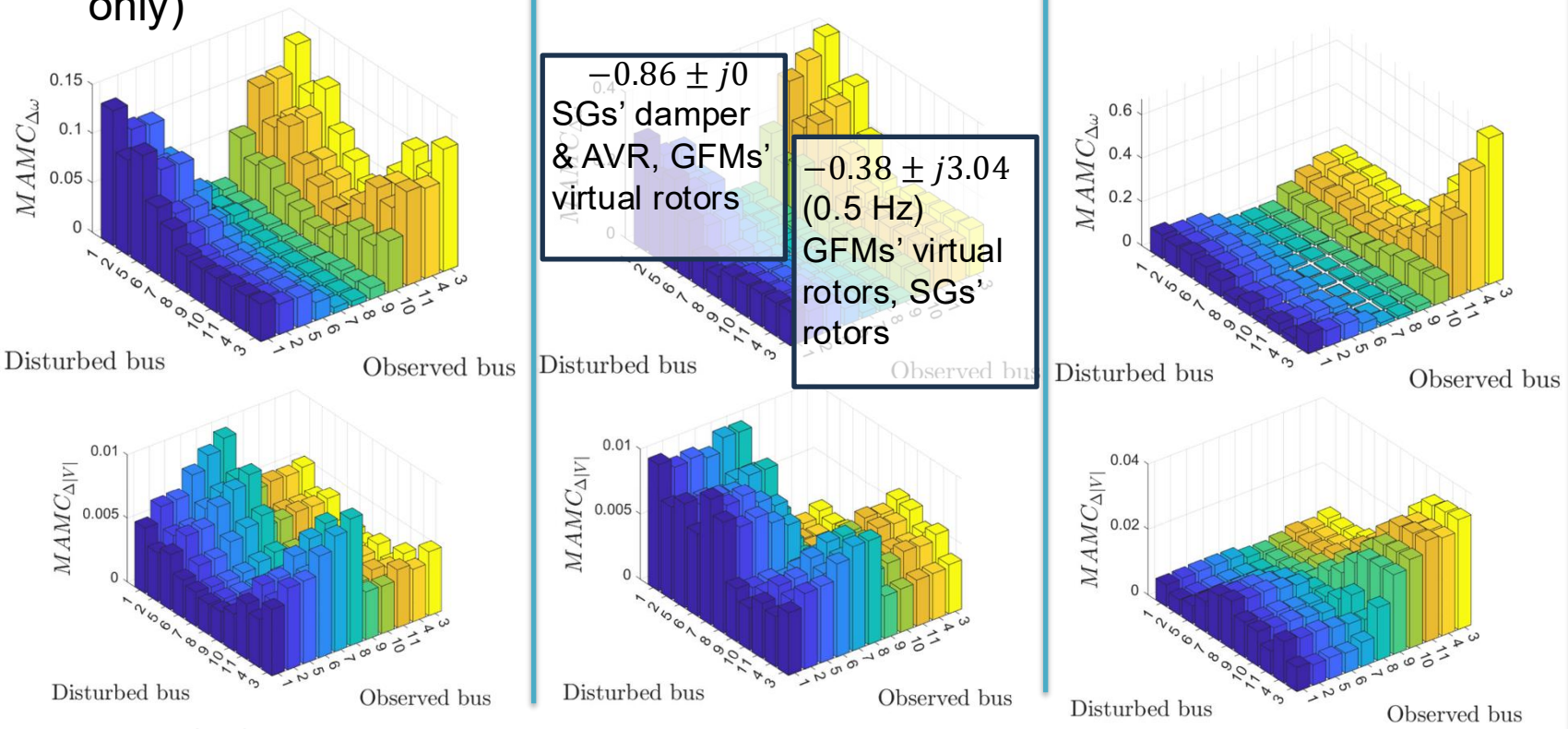
Seen as dip in correlation for disturbed bus area 1 and observed bus area 2

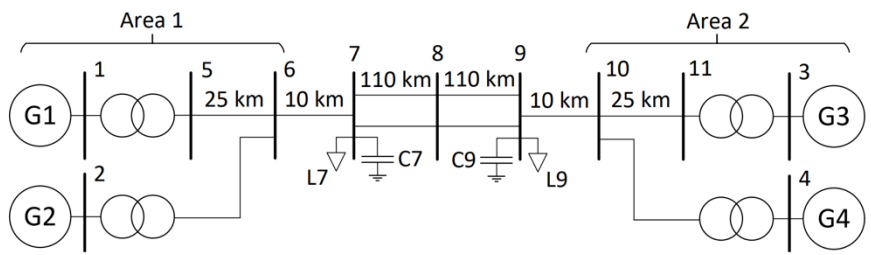
Application of the proposed methodology

Base case (SGs only)

GFMs for G3 & G4

GFLs for G3 & G4





Kundur's 2-area, 4-generator system

Key takeaways

Voltage and frequency trends are independent

When generators in area 2 are GFMs:

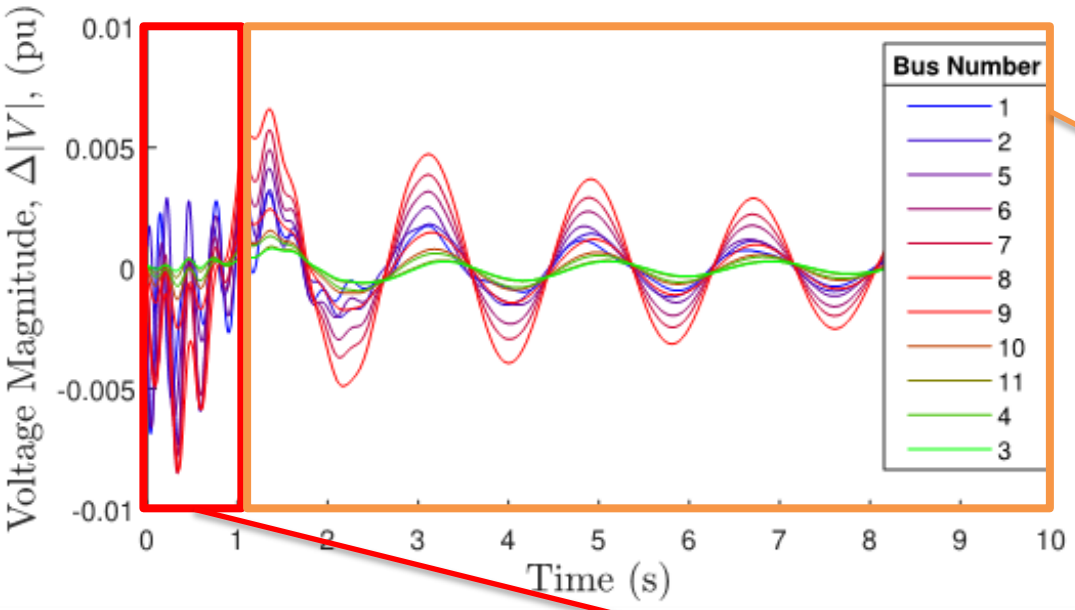
- Variability is excited most by disturbances in area 1 (SGs area)
- Variability is observed in both area 1 and area 2

When generators in area 2 are GFLs:

- Variability is excited and observed most in area 2

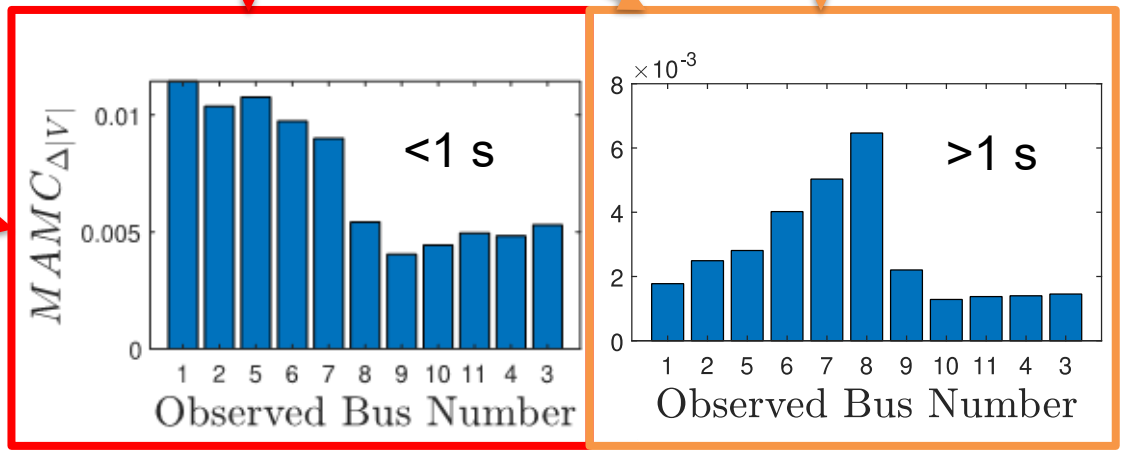
Different modes can contribute most to variability for different locations of disturbance and/or observation

Analysis over different timescales



Mode with maximum contribution:
$-1.78 \pm j20.89$ (3.32 Hz)
GFMs' inner controllers

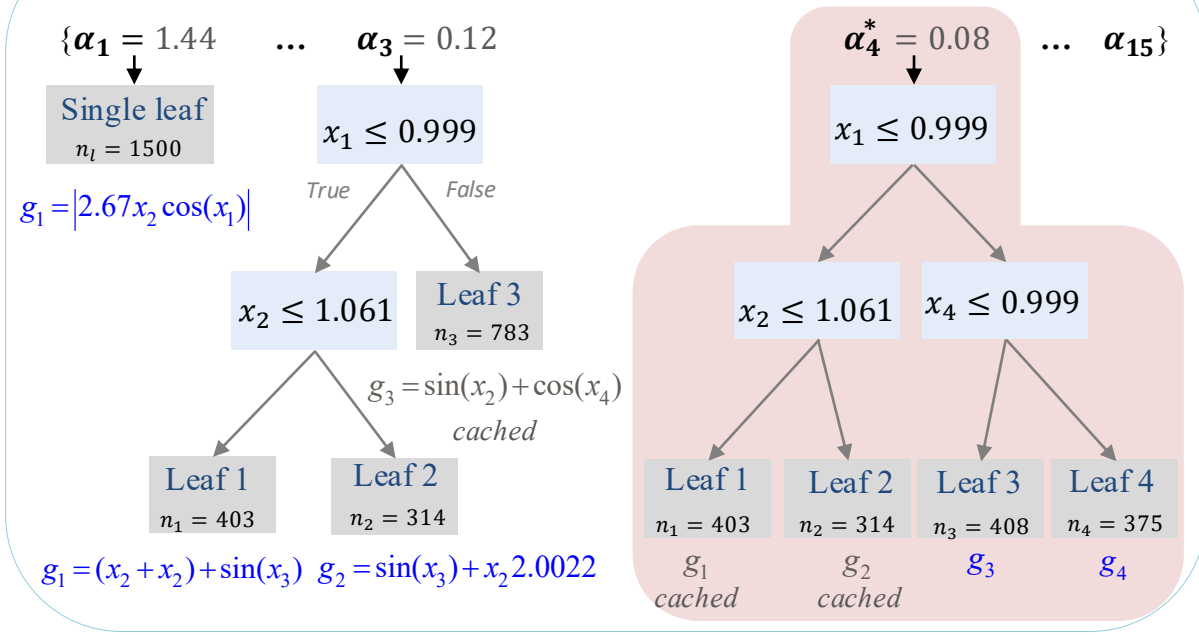
Mode with maximum contribution:
$-0.38 \pm j3.04$ (0.48 Hz)
GFMs' virtual rotors & SGs' rotors



MANCHESTER 1824
The University of Manchester

Pc-SR walk-through: CCP iteration → canonicalisation → embedding similarity → merging → recovered model

Iterate along effective $\{a_k\}$ from CCP path (simple → complex tree):



Selected Partition:

SR models	$g_1 = (x_2 + x_2) + \sin(x_3)$	$g_2 = \sin(x_3) + x_2 2.0022$	$g_3 = \log(x_2) x_4$	$g_4 = \sqrt{\lvert x_3 - x_5 \rvert}$
Canonicalization	$g_1 = 2x_2 + \sin(x_3)$	$g_2 = \sin(x_3) + x_2 2$	$g_3 = \log(x_2) x_4$	$g_4 = \sqrt{\lvert x_3 - x_5 \rvert}$

Embedding Similarity > 0.9 ⇒ Merge

Final Pc-SR model:

$$\hat{f}(x) = \begin{cases} g_1(x) = 2x_2 + \sin(x_3), & x \in \mathrm{R}_1^{\alpha^*} : \{x_1 \le 0.999\} \\ g_3(x) = \log(x_2) x_4, & x \in \mathrm{R}_3^{\alpha^*} : \{x_1 > 0.999 \wedge x_4 \le 0.999\} \\ g_4(x) = \sqrt{|\, x_3 - x_5 \,|}, & x \in \mathrm{R}_4^{\alpha^*} : \{x_1 > 0.999 \wedge x_4 > 0.999\} \end{cases}$$

- Single leaf at α_1: one global formula *fails to match any region.*
- $\alpha_4^* = \alpha \dagger = 0.08$: recovered splits **$x_1 \le 0.999$ and $x_4 \le 0.999$** *match the planted boundaries ($x_1 = 1$, $x_4 = 1$).*
- Canonicalisation **merges two equivalent leaves** . *Caching reuses unchanged leaves.*

Back