

# Data-driven decentralised control design in Active Distribution Networks

---

Petros Aristidou

Department of Electrical Engineering, Computer Engineering & Informatics,  
Cyprus University of Technology

<https://sps.cut.ac.cy>

(joint work with S. Karagiannopoulos and G. Hug, ETH Zurich)

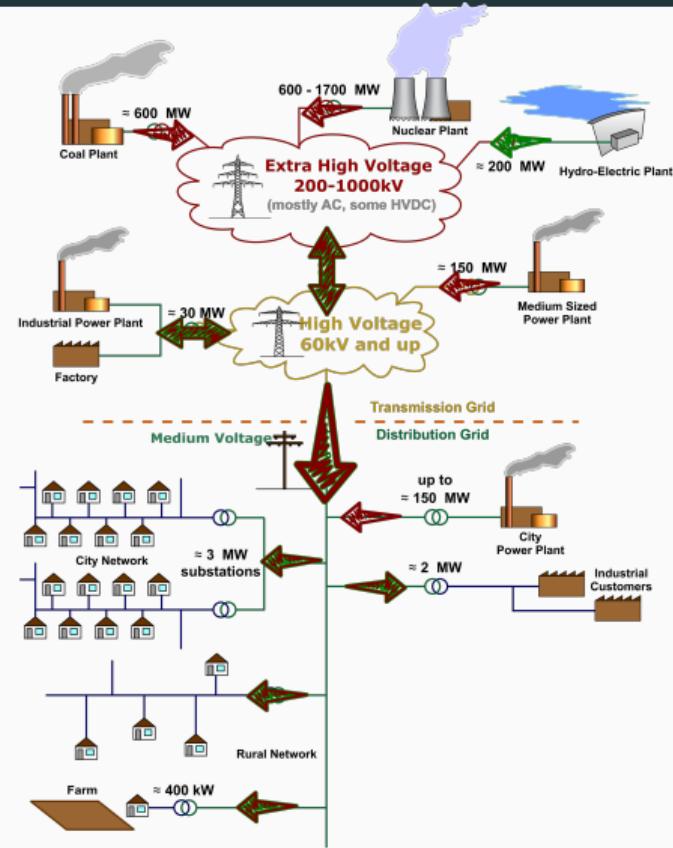


## Motivation

---

# Transformation of power systems

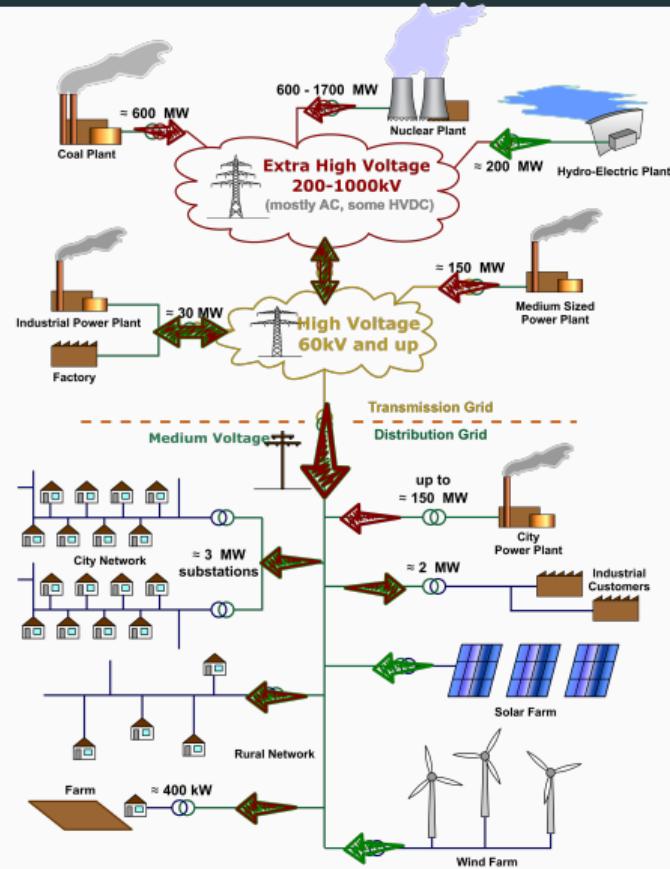
## New developments in distribution grids



# Transformation of power systems

## New developments in distribution grids

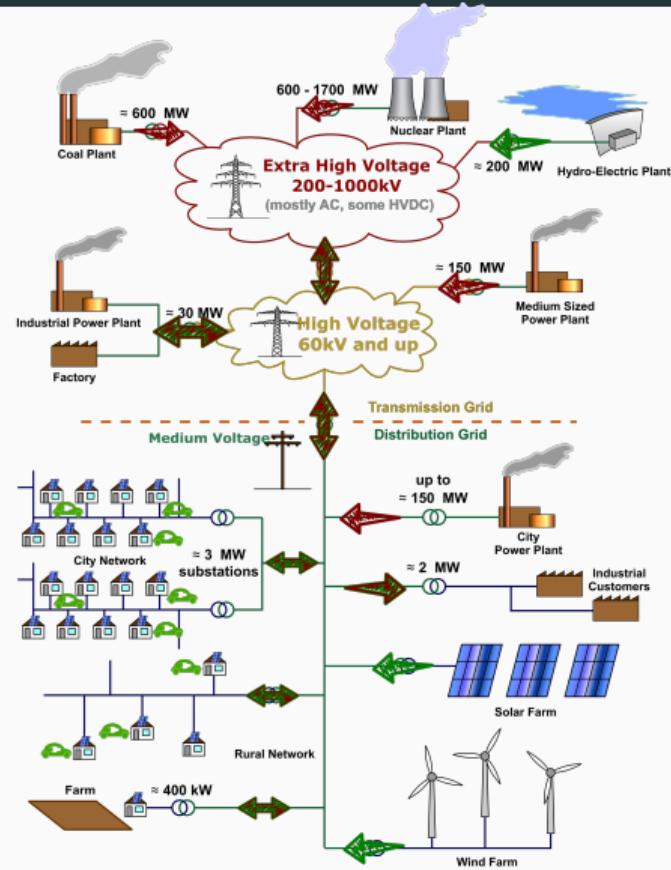
- Introduction of large distributed generators  
(renewable energy sources, etc.)



# Transformation of power systems

## New developments in distribution grids

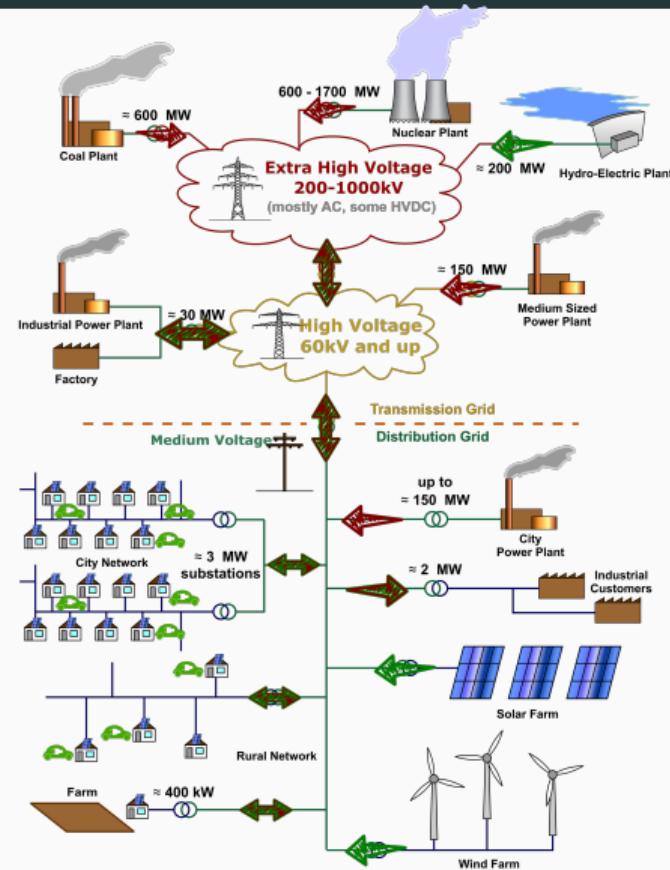
- Introduction of large distributed generators (renewable energy sources, etc.)
- Introduction of small distributed generators and energy storage systems
- Electrification of transportation (plug-in hybrid, battery electric, etc.)



# Transformation of power systems

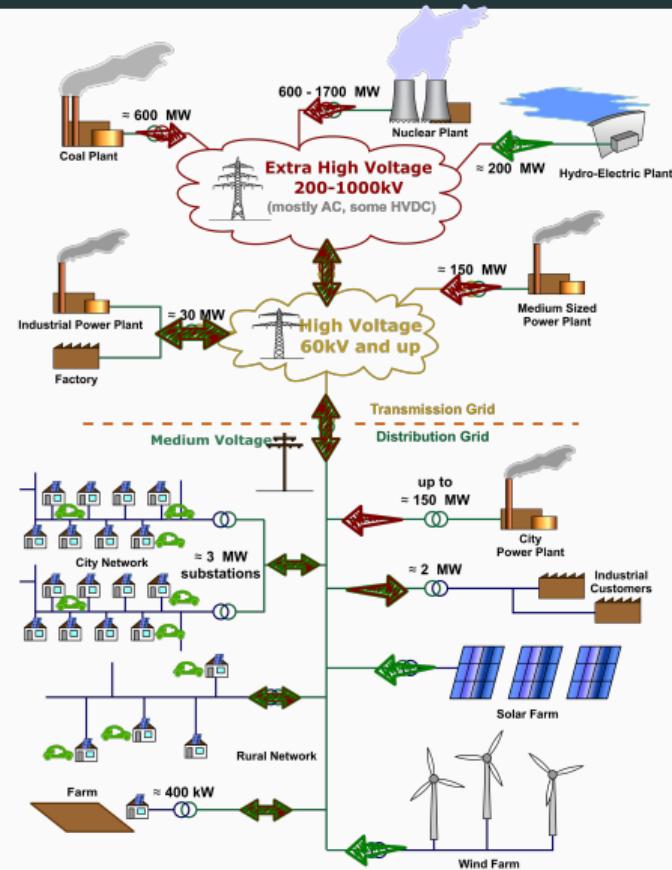
## New developments in distribution grids

- Introduction of large distributed generators (renewable energy sources, etc.)
- Introduction of small distributed generators and energy storage systems
- Electrification of transportation (plug-in hybrid, battery electric, etc.)
- Demand response schemes (reaction to price signals, emergency load reduction, peak shaving, etc.)



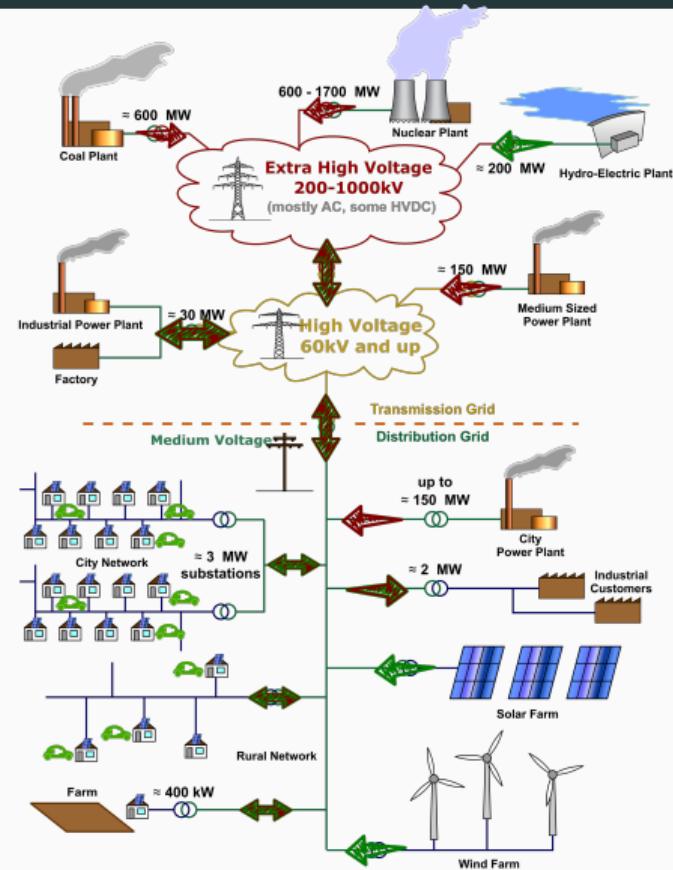
# New challenges

- Operation of the distribution grids close or above the physical limits and hosting capacity. *Distribution grids were not designed to host generation.*



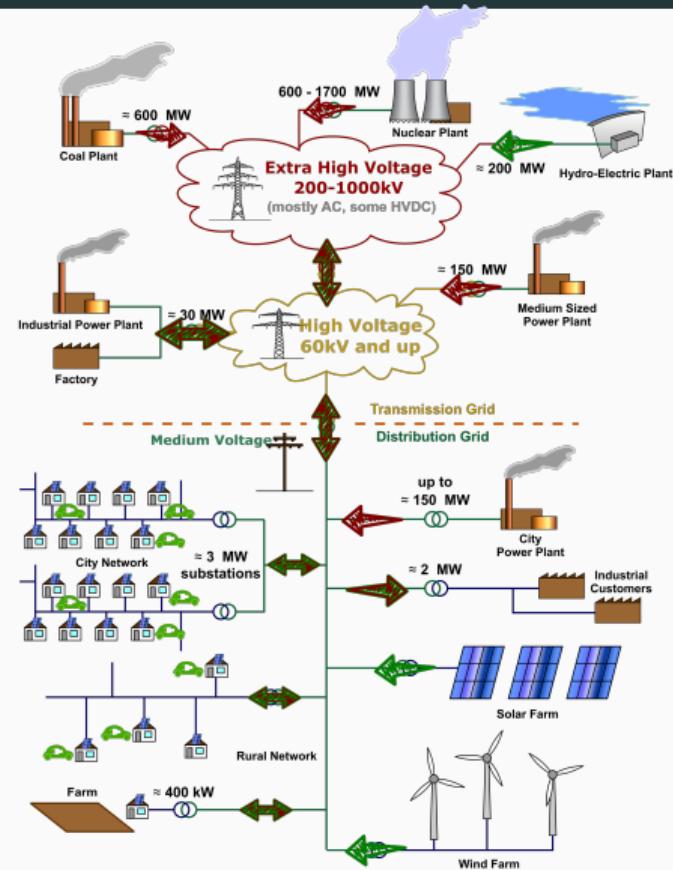
# New challenges

- Operation of the distribution grids close or above the physical limits and hosting capacity. *Distribution grids were not designed to host generation.*
- Bi-directional flows. *Most of system protections and operation practices were not designed for this.*



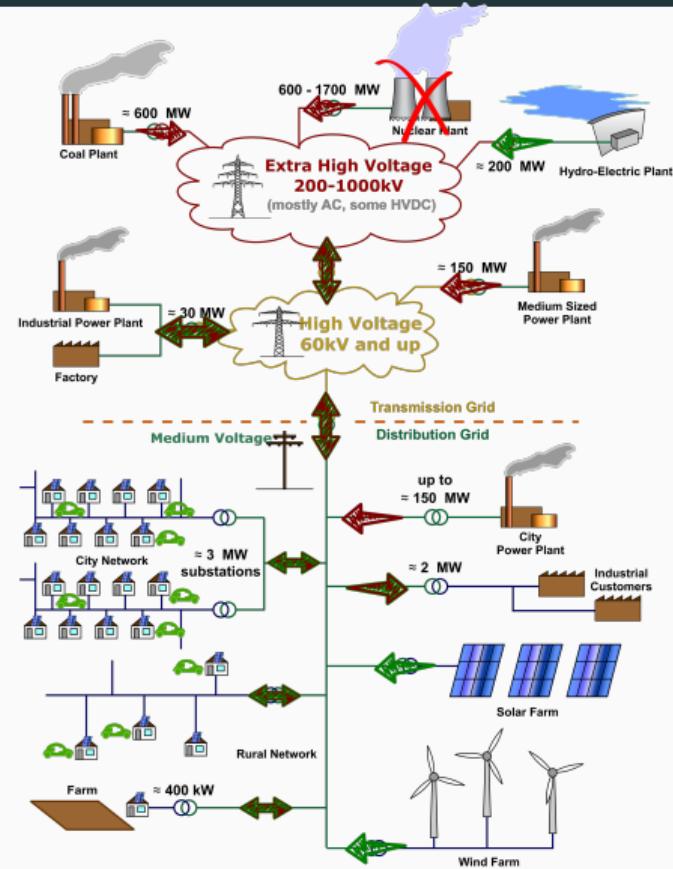
# New challenges

- Operation of the distribution grids close or above the physical limits and hosting capacity. *Distribution grids were not designed to host generation.*
- Bi-directional flows. *Most of system protections and operation practices were not designed for this.*
- Increased uncertainty. *Intermittent generation, new consumption profiles and patterns, unknown consumer response.*



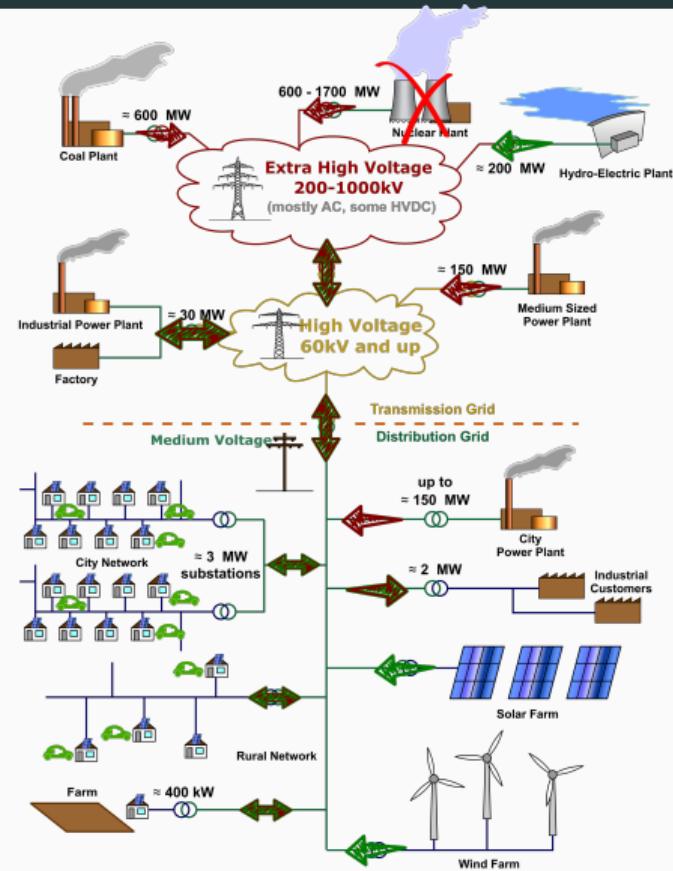
# New challenges

- Operation of the distribution grids close or above the physical limits and hosting capacity. *Distribution grids were not designed to host generation.*
- Bi-directional flows. *Most of system protections and operation practices were not designed for this.*
- Increased uncertainty. *Intermittent generation, new consumption profiles and patterns, unknown consumer response.*
- Decommission of conventional units. *Loss of traditional "dispatchable" generation and control.*



# New challenges

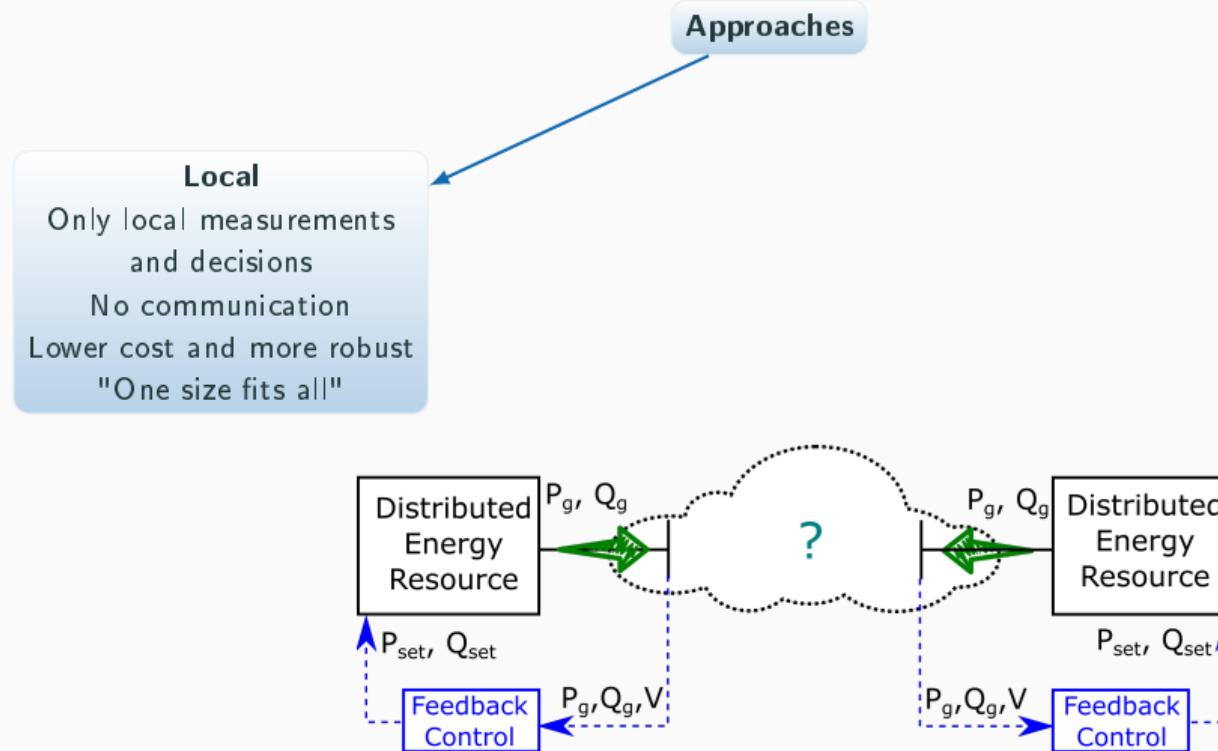
Need for Active Distribution Networks with  
real-time control



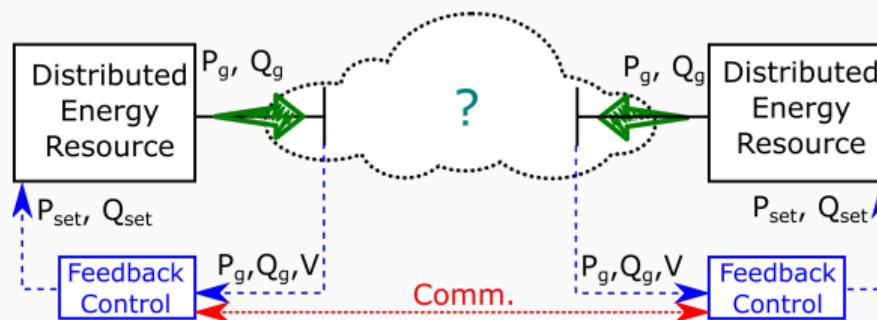
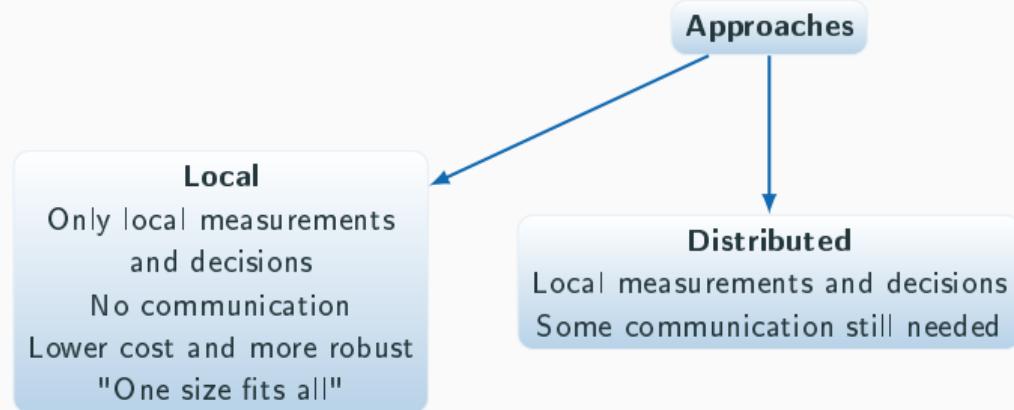
# Real-time operation of distribution grids

Approaches

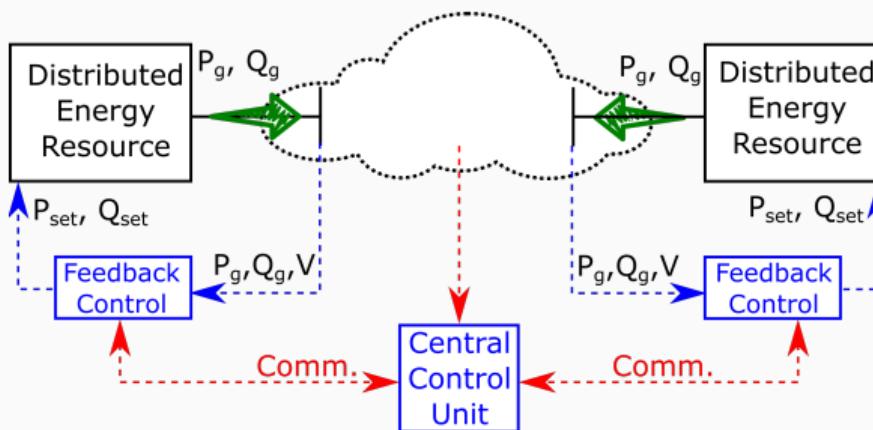
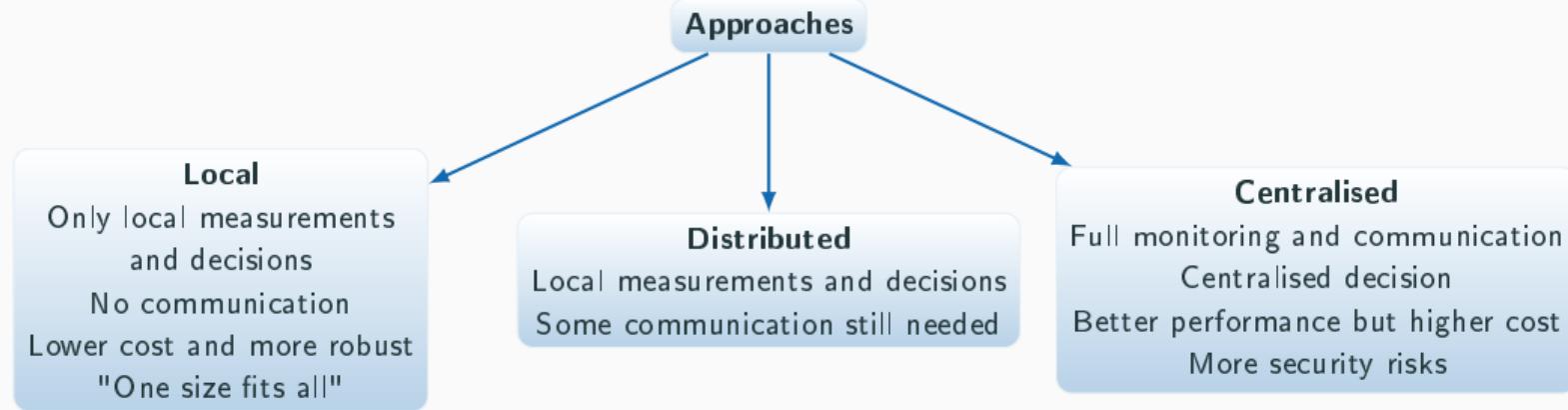
# Real-time operation of distribution grids



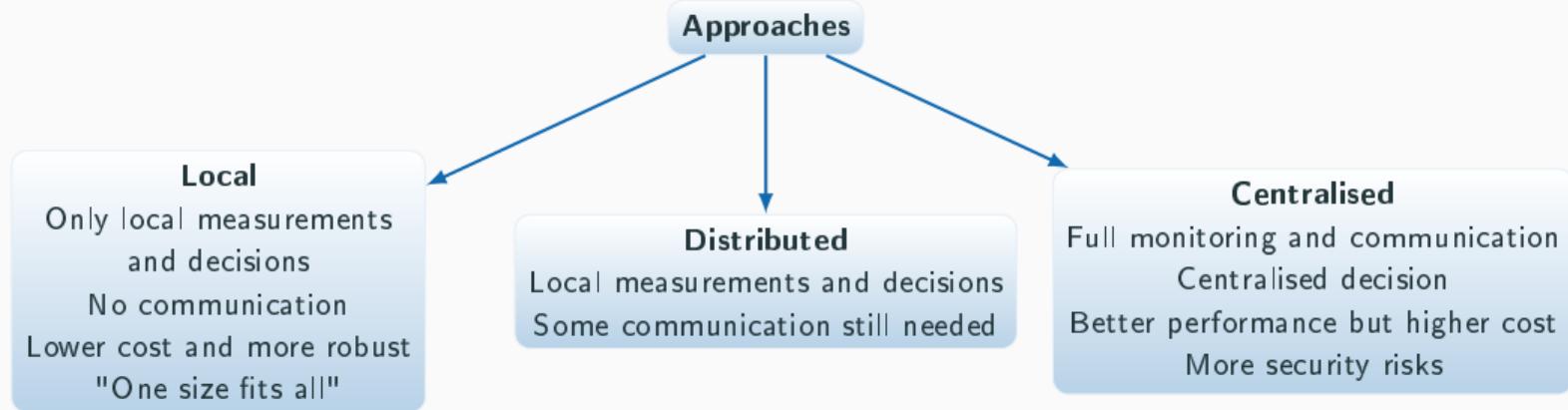
# Real-time operation of distribution grids



# Real-time operation of distribution grids

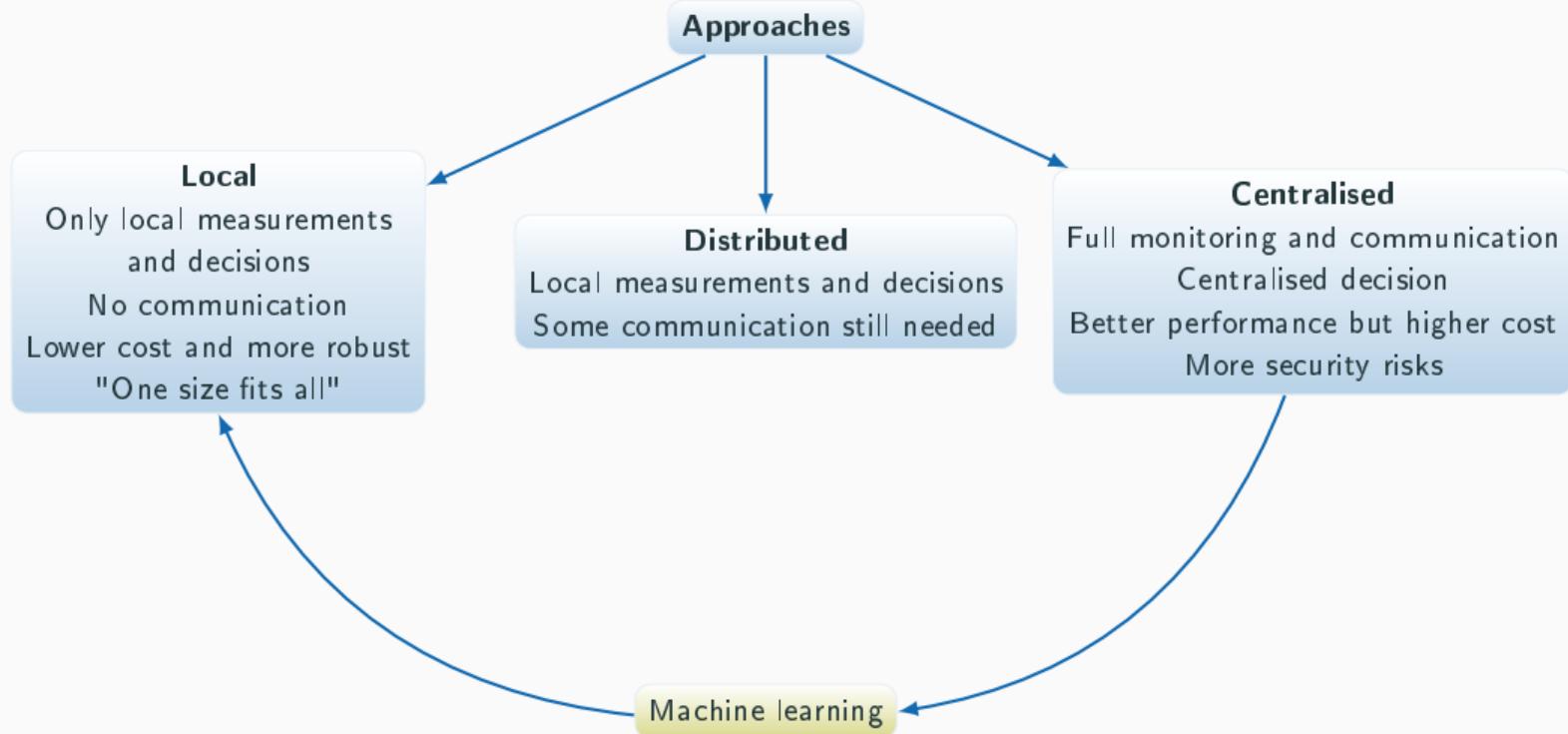


# Real-time operation of distribution grids



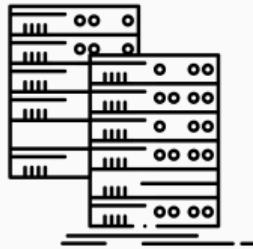
**Can we design local controls that can mimic the optimal behavior?**

# Real-time operation of distribution grids



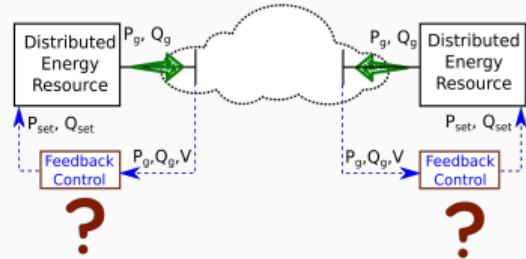
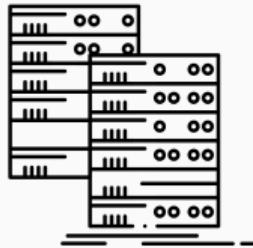
# Data-driven local control design

Historical timeseries data



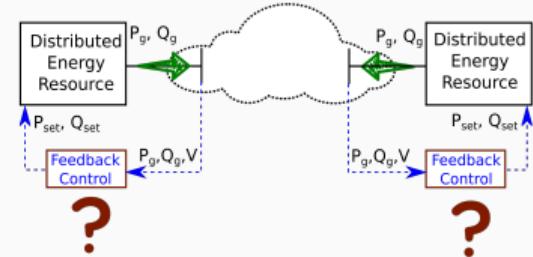
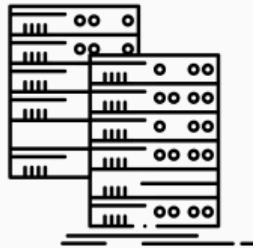
# Data-driven local control design

Historical timeseries data



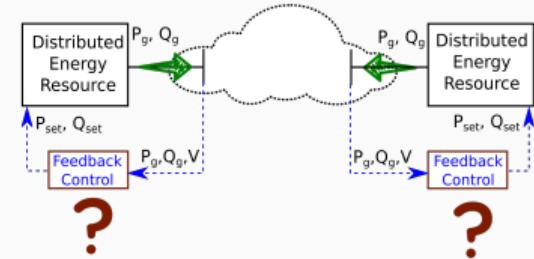
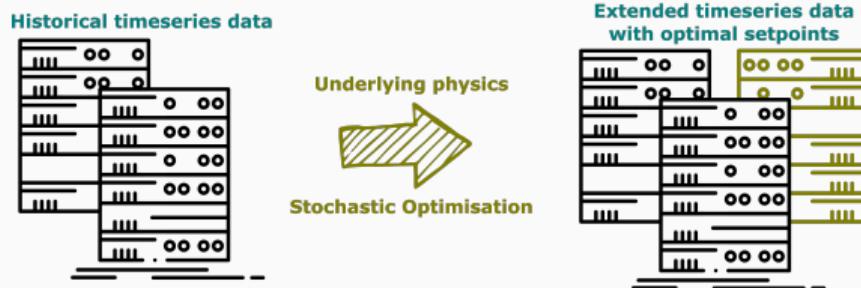
# Data-driven local control design

Historical timeseries data



Two-stage process

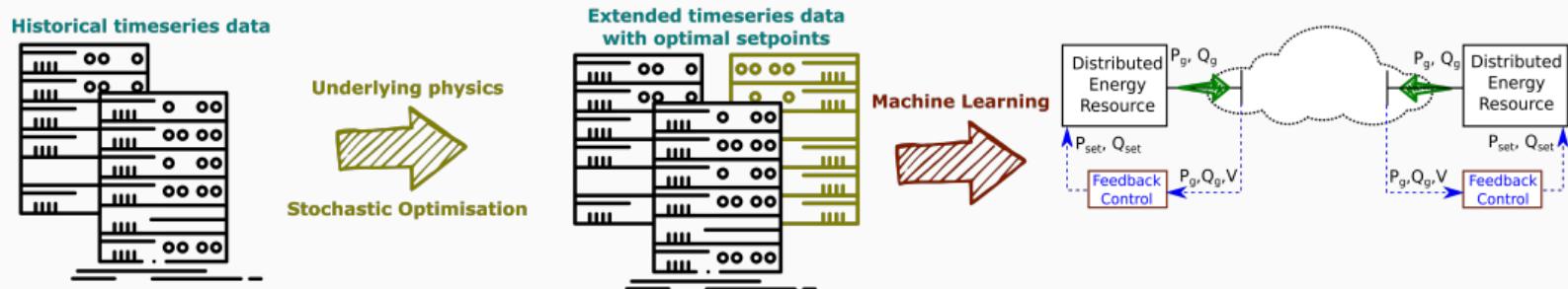
# Data-driven local control design



## Two-stage process

- **Step 1:** Process historical data and extend with optimal setpoints

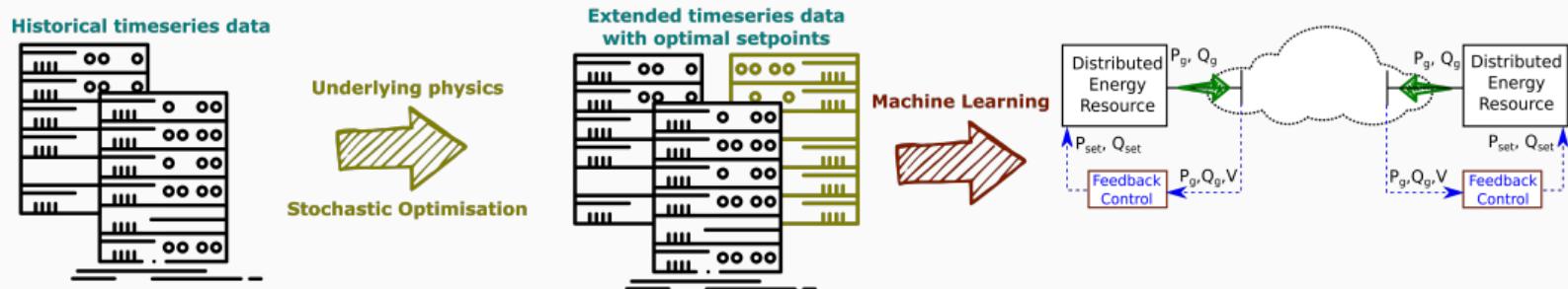
# Data-driven local control design



## Two-stage process

- **Step 1:** Process historical data and extend with optimal setpoints
- **Step 2:** Use ML on the extended dataset to design local controls

# Data-driven local control design



## Two-stage process

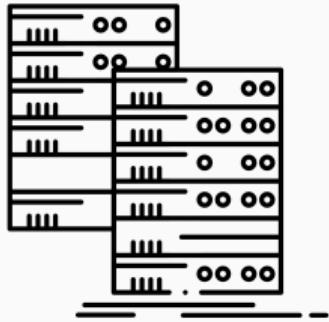
- **Step 1:** Process historical data and extend with optimal setpoints
- **Step 2:** Use ML on the extended dataset to design local controls
- Testing and validation

## Data-driven local control design

---

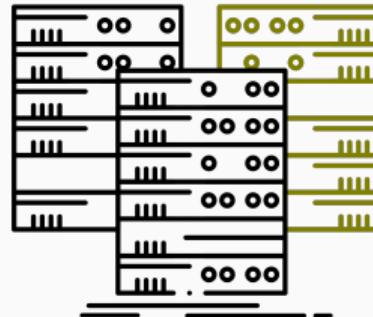
# Extending historical data with optimal setpoints

Historical timeseries data



Underlying physics  
Stochastic Optimisation

Extended timeseries data  
with optimal setpoints



## Multi-period OPF problem formulation

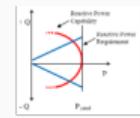
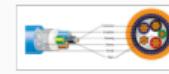
$$\min_{\boldsymbol{u}} \sum_t (c_{op}^T \boldsymbol{u} + c_{el}^T losses)$$

## Multi-period OPF problem formulation

$$\min_{\boldsymbol{u}} \sum_t (c_{op}^T \boldsymbol{u} + c_{el}^T losses)$$

Controls ( $\boldsymbol{u}$ ):

- Active power curtailment (APC)
- Reactive power control (RPC)
- Battery Energy Storage Systems (BESS)
- Controllable loads (CLs)
- On-Load Tap Changers (OLTC)



# OPF-based data processing

## Multi-period OPF problem formulation

$$\min_{\boldsymbol{u}} \sum_t (c_{op}^T \boldsymbol{u} + c_{el}^T losses)$$

### Subject to:

- AC power-flow constraints
- Voltage limits
- Thermal loading limits
- DER limits
- Balancing constraints
- Controllable load constraints
- BESS dynamics

### Controls ( $\boldsymbol{u}$ ):

- Active power curtailment (APC)
- Reactive power control (RPC)
- Battery Energy Storage Systems (BESS)
- Controllable loads (CLs)
- On-Load Tap Changers (OLTC)



# OPF-based data processing

## Multi-period OPF problem formulation

$$\min_{\boldsymbol{u}} \sum_t (c_{op}^T \boldsymbol{u} + c_{el}^T losses)$$

Subject to:

- AC power-flow constraints
- Voltage limits
- Thermal loading limits
- DER limits
- Balancing constraints
- Controllable load constraints
- BESS dynamics

Controls ( $\boldsymbol{u}$ ):

- Active power curtailment (APC)
- Reactive power control (RPC)
- Battery Energy Storage Systems (BESS)
- Controllable loads (CLs)
- On-Load Tap Changers (OLTC)



# OPF-based data processing

## AC power-flow constraints

- Non-convex and non-linear

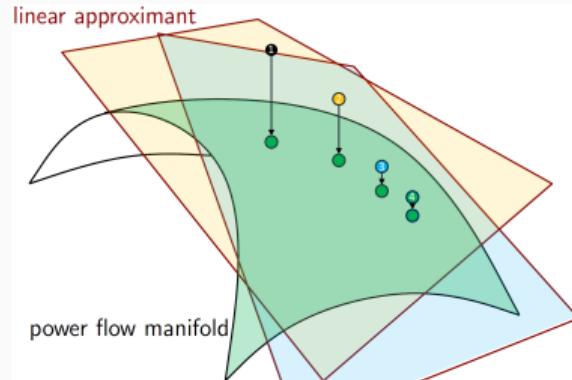
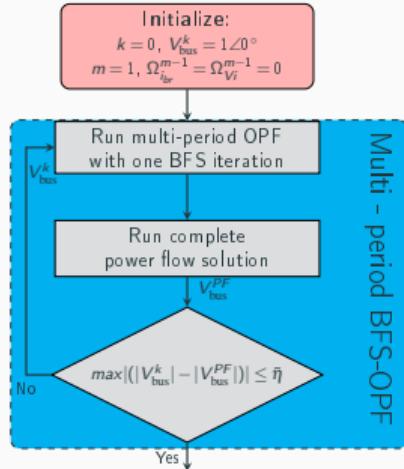
## AC power-flow constraints

- Non-convex and non-linear
  - Linear approximations of the AC power flows, heuristics, convex relaxations (Venzke et al 2018, Shchetinin & Hug 2017, Capitanescu & Bilibin 2016, Sankur et al. 2016, Bolognani & Dörfler 2015, Dall' Anese et al. 2015, Molzahn & Hiskens 2015, Lavaei & Low 2012, Paudyal et al. 2011)
  - Backward/Forward Sweep (BFS) power flow (Fortsenbacher et al. 2016)
    - ▶ Iterative procedure
    - ▶ Exploit the radial grid structure
    - ▶ Weakly meshed treatment

## AC power-flow constraints

- Non-convex and non-linear
  - Linear approximations of the AC power flows, heuristics, convex relaxations (Venzke et al 2018, Shchetinin & Hug 2017, Capitanescu & Bilibin 2016, Sankur et al. 2016, Bolognani & Dörfler 2015, Dall' Anese et al. 2015, Molzahn & Hiskens 2015, Lavaei & Low 2012, Paudyal et al. 2011)
  - **Backward/Forward Sweep (BFS) power flow (Fortenbacher et al. 2016)**
    - ▶ Iterative procedure
    - ▶ Exploit the radial grid structure
    - ▶ Weakly meshed treatment
- **Use a single BFS iteration for the OPF problem**

# OPF-based data processing



Qualitative illustration of the iterative BFS-OPF scheme

## Tackling Uncertainty

- Branch current flows and voltages are functions of the power injections and are hence influenced by **renewable generator & load power uncertainty**

## Tackling Uncertainty

- Branch current flows and voltages are functions of the power injections and are hence influenced by **renewable generator & load power uncertainty**
- ▶ Formulation of Chance Constraints

$$\mathbb{P}\{|V_{\text{bus},j,t}| \leq V_{\max}\} \geq 1 - \varepsilon$$

$$\mathbb{P}\{|V_{\text{bus},j,t}| \geq V_{\min}\} \geq 1 - \varepsilon$$

$$\mathbb{P}\{|I_{\text{br},i,t}| \leq I_{i,\max}\} \geq 1 - \varepsilon$$

## Tackling Uncertainty

- Branch current flows and voltages are functions of the power injections and are hence influenced by **renewable generator & load power uncertainty**
- ▶ Formulation of Chance Constraints
- ▶ Reformulate into deterministic constraints through “tightenings”

$$\mathbb{P}\{|V_{\text{bus},j,t}| \leq V_{\max}\} \geq 1 - \varepsilon$$

$$\mathbb{P}\{|V_{\text{bus},j,t}| \geq V_{\min}\} \geq 1 - \varepsilon$$

$$\mathbb{P}\{|I_{\text{br},i,t}| \leq I_{i,\max}\} \geq 1 - \varepsilon$$

$$V_{\min} + \Omega_{V,j,t}^{\text{lower}} \leq |V_{\text{bus},j,t}^k| \leq V_{\max} - \Omega_{V,j,t}^{\text{upper}}$$

$$|I_{\text{br},i,t}^k| \leq I_{i,\max} - \Omega_{I_{\text{br},i}}$$

# OPF-based data processing

## Tackling Uncertainty

- Branch current flows and voltages are functions of the power injections and are hence influenced by **renewable generator & load power uncertainty**
- Formulation of Chance Constraints
- Reformulate into deterministic constraints through “tightenings”

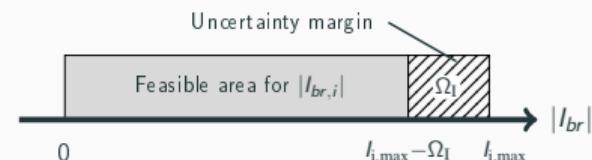
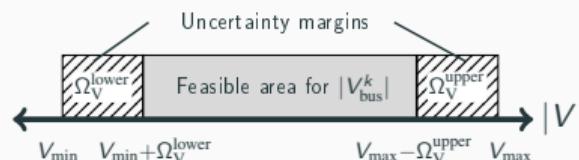
$$\mathbb{P}\{|V_{\text{bus},j,t}| \leq V_{\max}\} \geq 1 - \varepsilon$$

$$\mathbb{P}\{|V_{\text{bus},j,t}| \geq V_{\min}\} \geq 1 - \varepsilon$$

$$\mathbb{P}\{|I_{\text{br},i,t}| \leq I_{i,\max}\} \geq 1 - \varepsilon$$

$$V_{\min} + \Omega_{V,j,t}^{\text{lower}} \leq |V_{\text{bus},j,t}^k| \leq V_{\max} - \Omega_{V,j,t}^{\text{upper}}$$

$$|I_{\text{br},i,t}^k| \leq I_{i,\max} - \Omega_{I_{\text{br}},i}$$

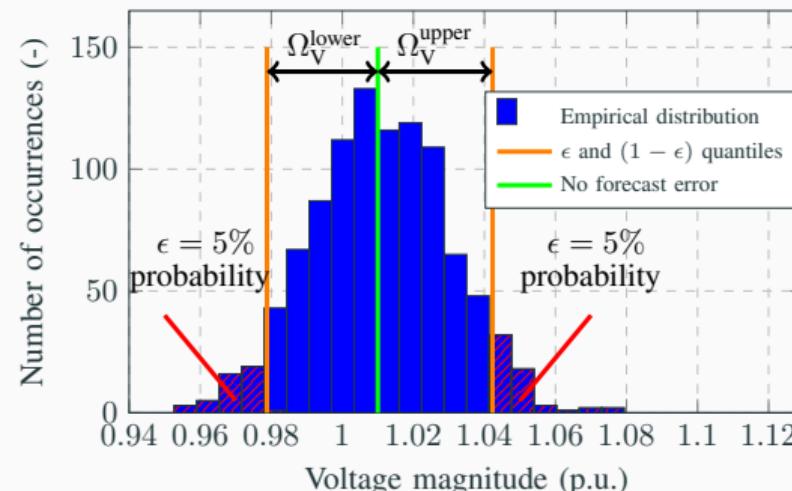


## Uncertainty margins evaluation

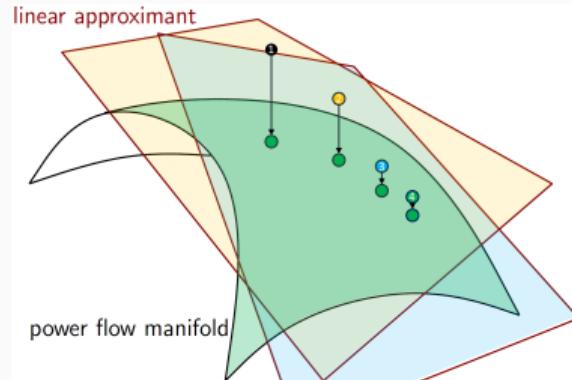
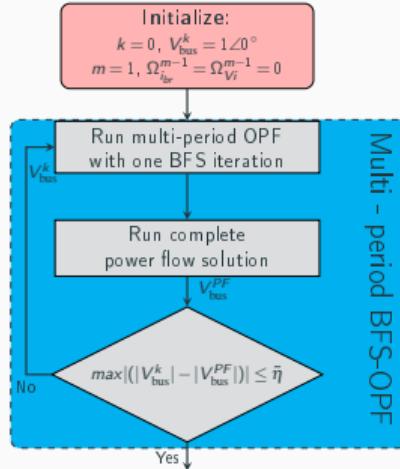
- Analytical approach → Need to know the probability distribution

## Uncertainty margins evaluation

- Analytical approach → Need to know the probability distribution
- Monte Carlo simulation using historical data from forecast errors
  - No assumptions about the uncertainty distribution
- Quantile  $\epsilon$  calculation

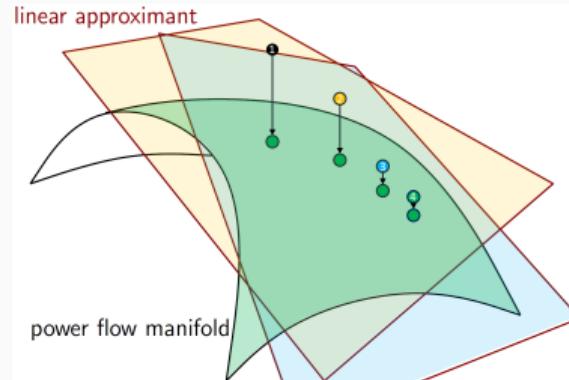
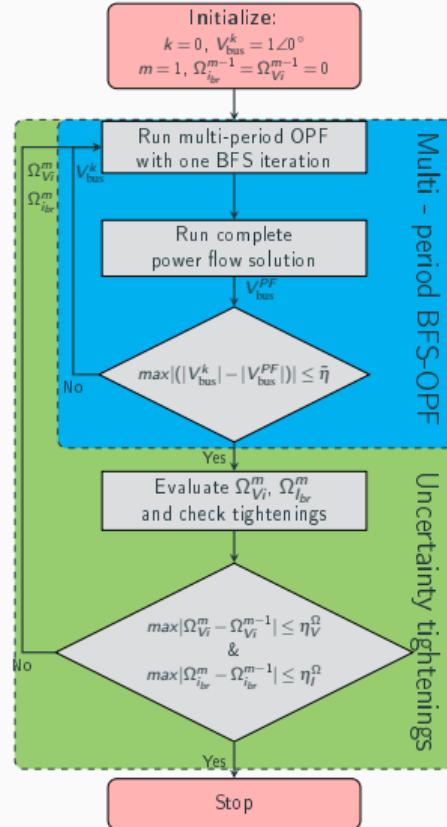


# OPF-based data processing



Qualitative illustration of the iterative BFS-OPF scheme

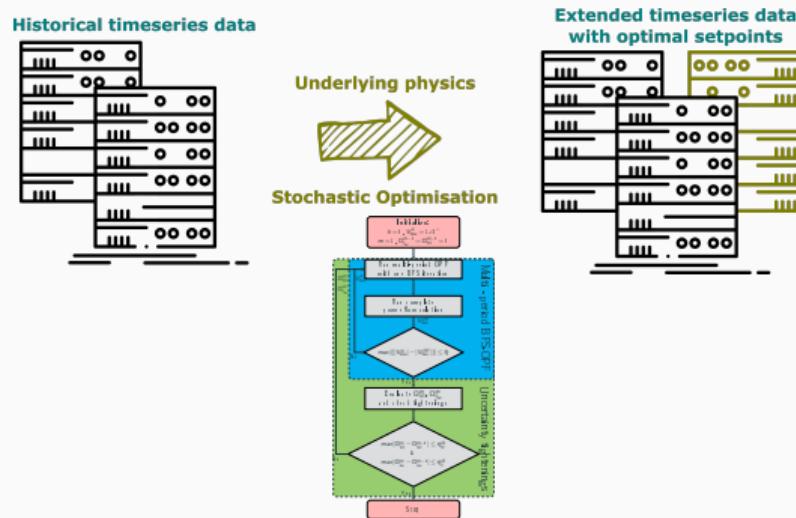
# OPF-based data processing



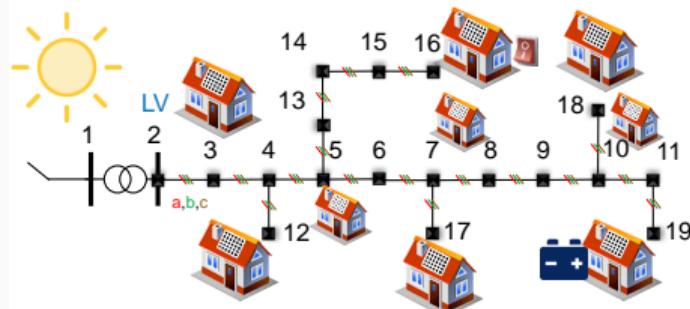
Qualitative illustration of the iterative BFS-OPF scheme



# Extending historical data with optimal setpoints



# Test system



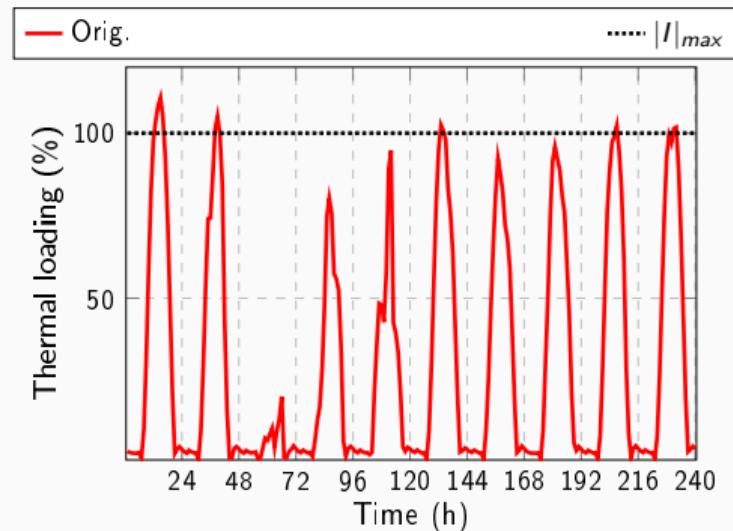
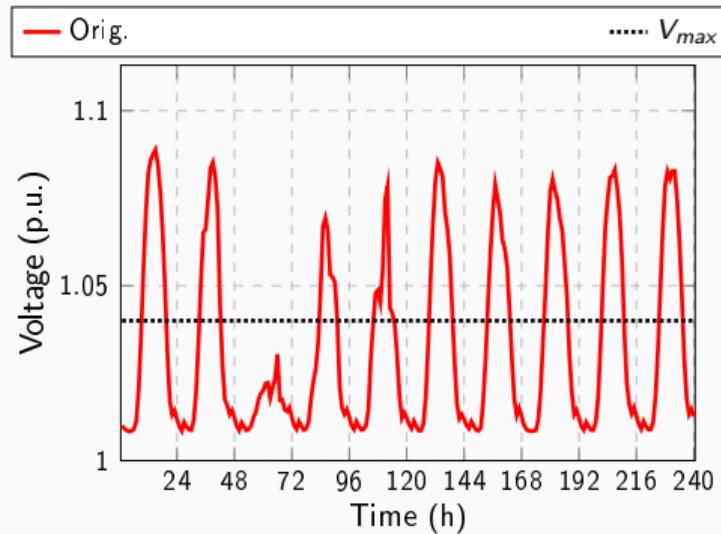
## Control actions

- Active Power Curtailment (APC)
- Reactive Power Control (RPC)
- Battery Energy Storage System (BESS)
- Controllable load (CL)
- On Load Tap Changers (OLTC)

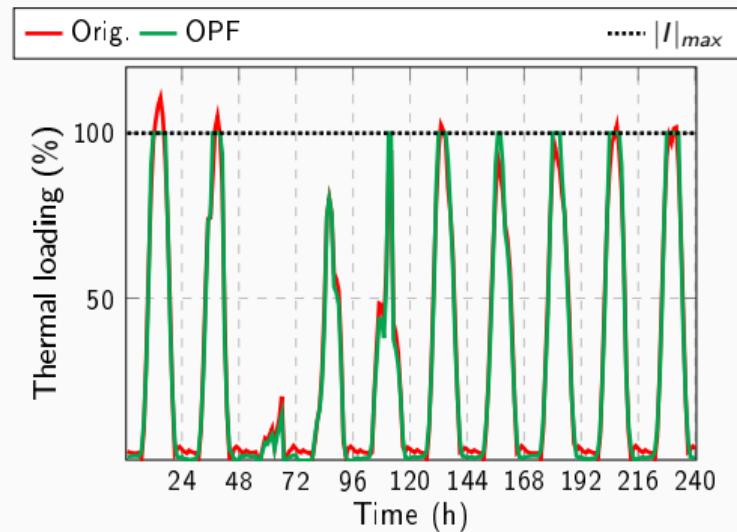
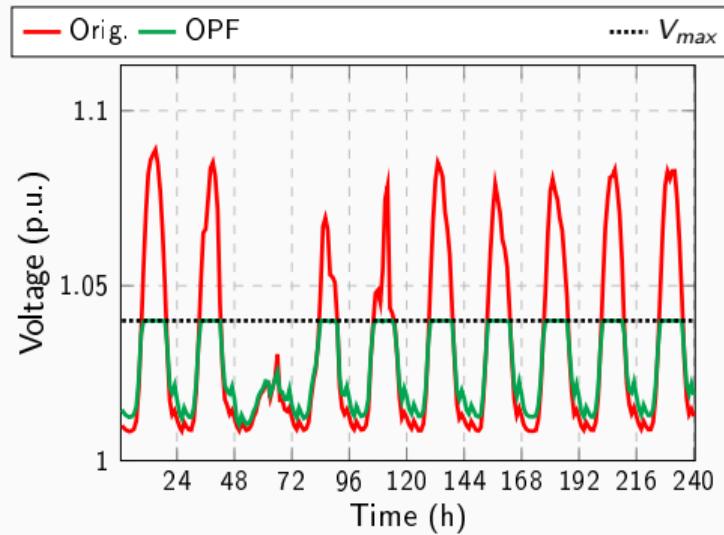
## Network description

- Based on European CIGRE LV grid
- Normalized profiles
  - PV & forecasts: Real data from Zurich
  - Load: Typical profiles based on CIGRE
- Summer day simulations
  - High solar radiation
- Acceptable limits:
  - Voltage:  $\pm 4\%$
  - Current: up to 1 p.u.
  - VUF: up to 2%

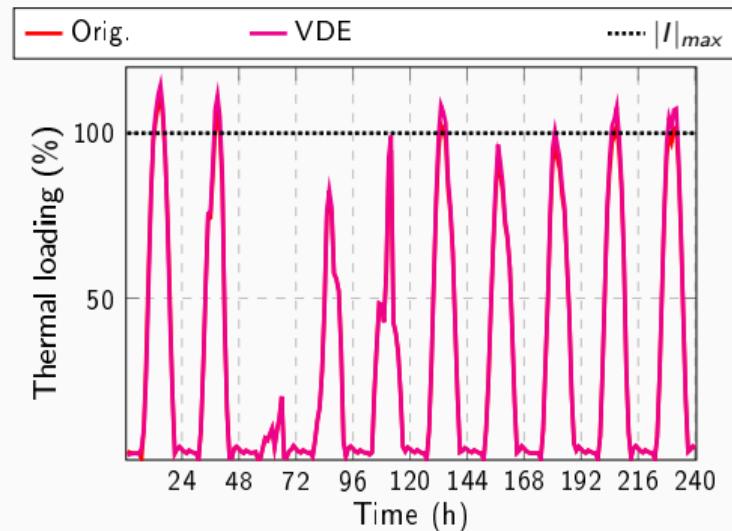
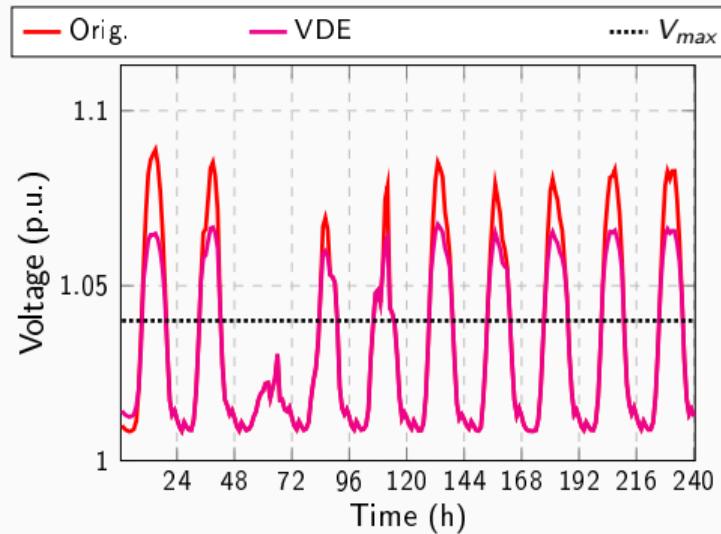
## Some results



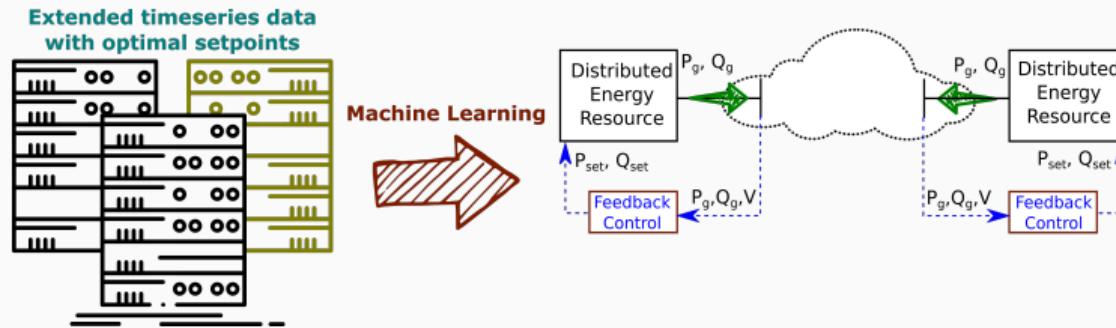
## Some results



## Some results



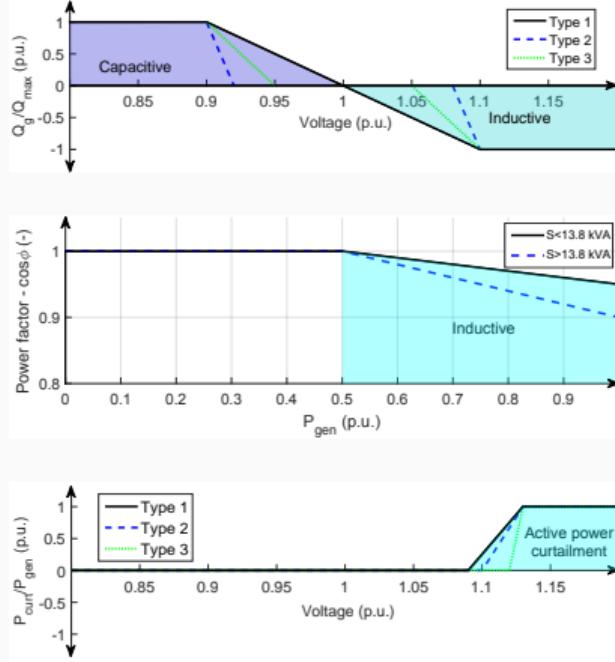
# Optimised local control schemes



# Optimised local control schemes

## Existing local control schemes

- Usually all distributed generators of same type and similar size have the same curve
- Several types, usually:  $Q = f(V)$ ,  $\cos\phi = f(P)$ ,  $P_{curt} = f(V)$
- "Open" loop or "closed" loop



# Optimised local control schemes

## Existing local control schemes

- Usually all distributed generators of same type and similar size have the same curve
- Several types, usually:  $Q = f(V)$ ,  $\cos\phi = f(P)$ ,  $P_{curt} = f(V)$
- "Open" loop or "closed" loop

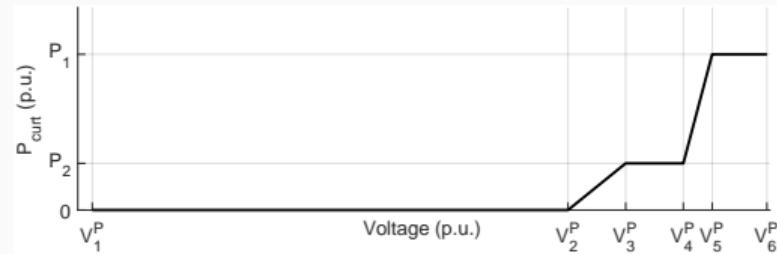
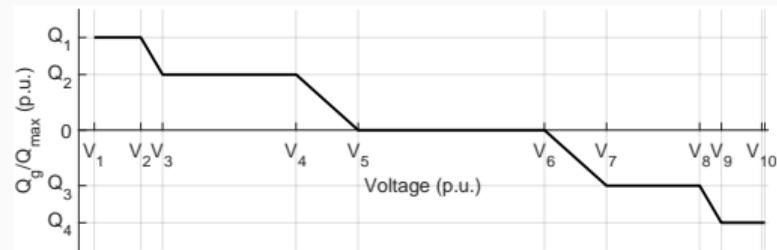
## Optimised local control schemes

- Data-driven local controls based on processed data from the previous step
- Candidate DERs
  - Distributed Generators (DGs)
  - Battery Energy Storage Systems (BESS)
  - Controlable Loads (CLs)

# Optimised local control schemes

## DGs: Piece-wise (segmented) linear fitting

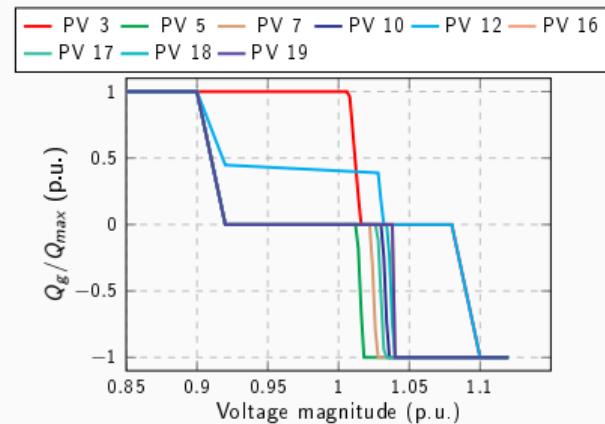
- Simple and efficient (R, sklearn, MATLAB, etc.)
  - Breakpoint selection
  - Nonlinear least squares regression
- Modified to impose stability-related monotonicity and slope constraints



# Optimised local control schemes

## DGs: Piece-wise (segmented) linear fitting

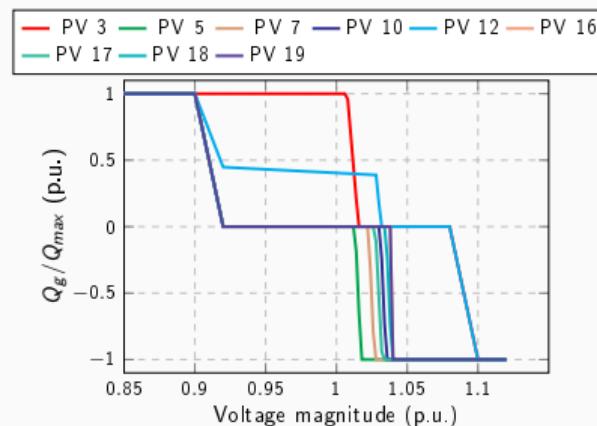
- Simple and efficient (R, sklearn, MATLAB, etc.)
  - Breakpoint selection
  - Nonlinear least squares regression
- Modified to impose stability-related monotonicity and slope constraints



# Optimised local control schemes

## Unique characteristic curve per DG

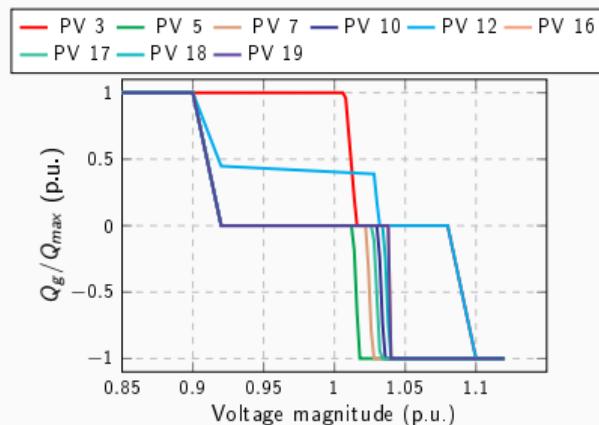
- Need to program a different curve for each agent
- Large number of inverter-based DGs



# Optimised local control schemes

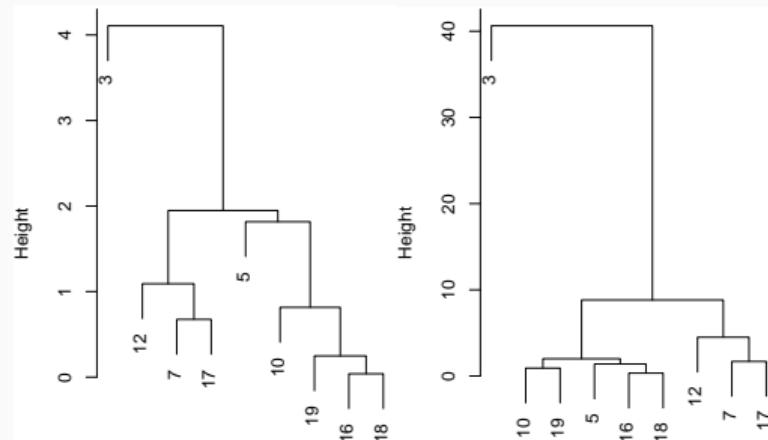
## Unique characteristic curve per DG

- Need to program a different curve for each agent
- Large number of inverter-based DGs



## Clustering of the curves

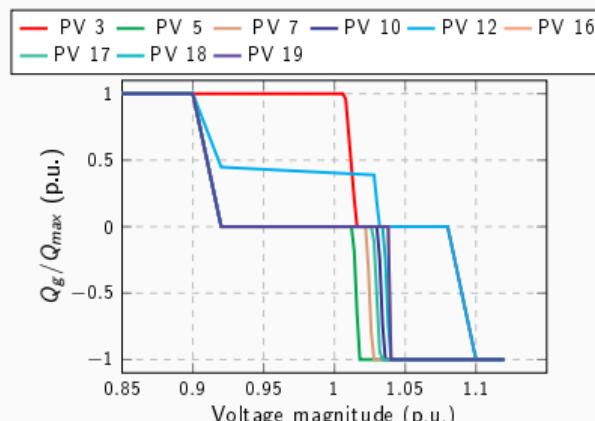
- Assign DGs to clustered curves based on “distance” with hierarchical clustering



# Optimised local control schemes

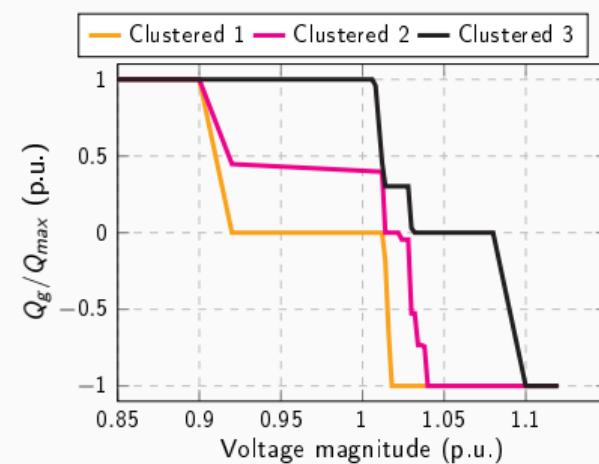
## Unique characteristic curve per DG

- Need to program a different curve for each agent
- Large number of inverter-based DGs



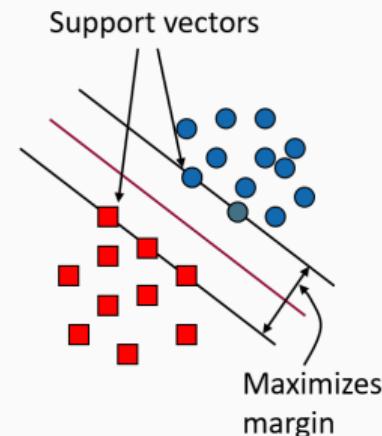
## Clustering of the curves

- Assign DGs to clustered curves based on “distance” with hierarchical clustering



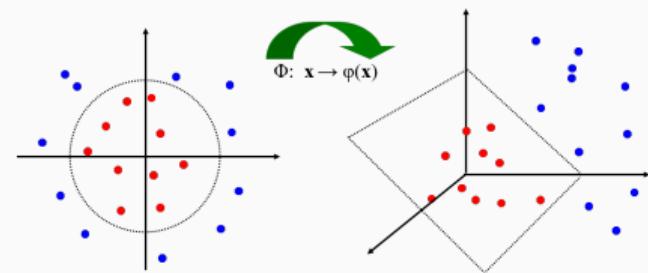
## BESS: Support Vector Machine (regression)

- Maximize the margin around the separating hyperplane
- Simple and efficient (R, sklearn, MATLAB, etc.)
  - quadratic programming problem



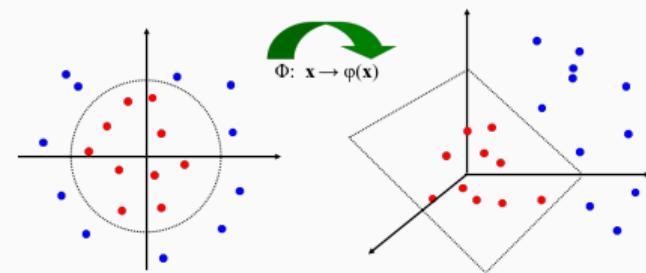
## BESS: Support Vector Machine (regression)

- Maximize the margin around the separating hyperplane
- Simple and efficient (R, sklearn, MATLAB, etc.)
  - quadratic programming problem
- Implicit mapping via kernels (Linear, Polynomial, Gaussian)
  - Map the original feature space to some higher-dimensional feature space where the training set is separable



## BESS: Support Vector Machine (regression)

- Maximize the margin around the separating hyperplane
- Simple and efficient (R, sklearn, MATLAB, etc.)
  - quadratic programming problem
- Implicit mapping via kernels (Linear, Polynomial, Gaussian)
  - Map the original feature space to some higher-dimensional feature space where the training set is separable
- Modified to impose stability-related impose monotonicity and slope constraints (e.g., monotone kernel regression methods)



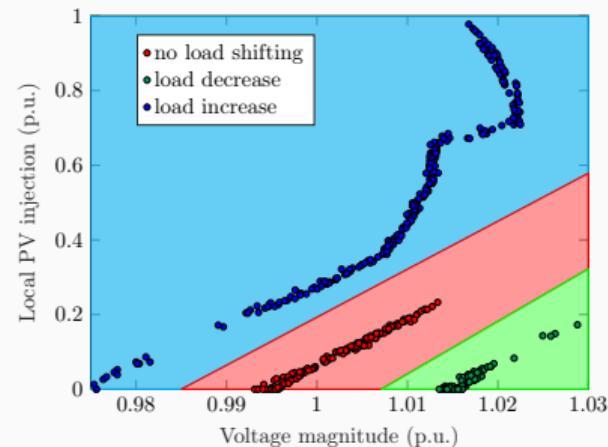
## CLs: Support Vector Machines (classification)

- Easy to implement (Python, MATLAB)
- Apply SVM with different kernel functions using the following features:
  - Voltage
  - Active power consumption
  - Reactive power consumption
  - Active PV power injection

# Optimised local control schemes

## CLs: Support Vector Machines (classification)

- Easy to implement (Python, MATLAB)
- Apply SVM with different kernel functions using the following features:
  - Voltage
  - Active power consumption
  - Reactive power consumption
  - Active PV power injection



# Optimised local control schemes

## Summary of methods

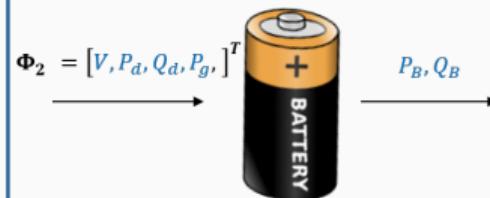
### Distributed Generators

- Segmented regression with unknown breakpoints
- Monotonicity & slope constraints
- Volt/Var & Volt/Watt curves



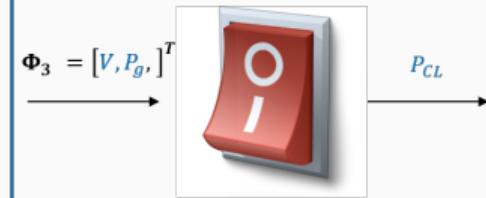
### BESS

- Support Vector Machine (SVM) regression
  - Various kernel functions
  - Energy constraints

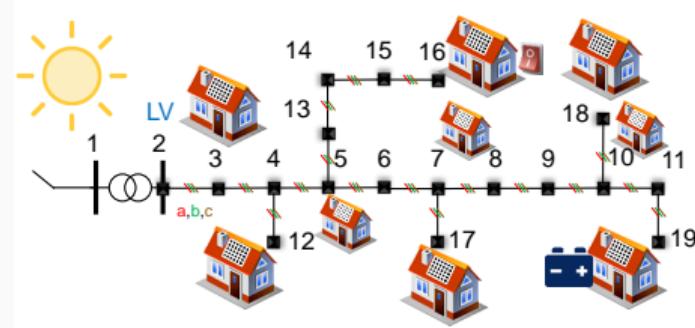


### CLs

- Support Vector Machine (SVM) classification
  - Various kernel functions
  - 3 classes: load decrease, increase, or no shifting



# Test system



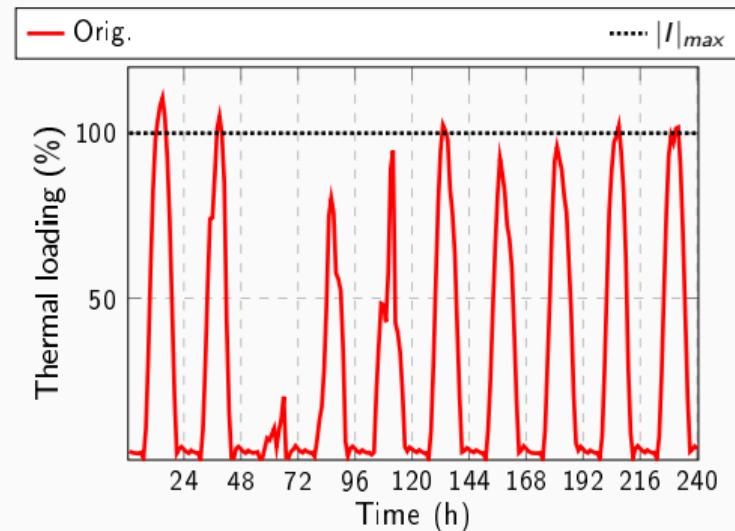
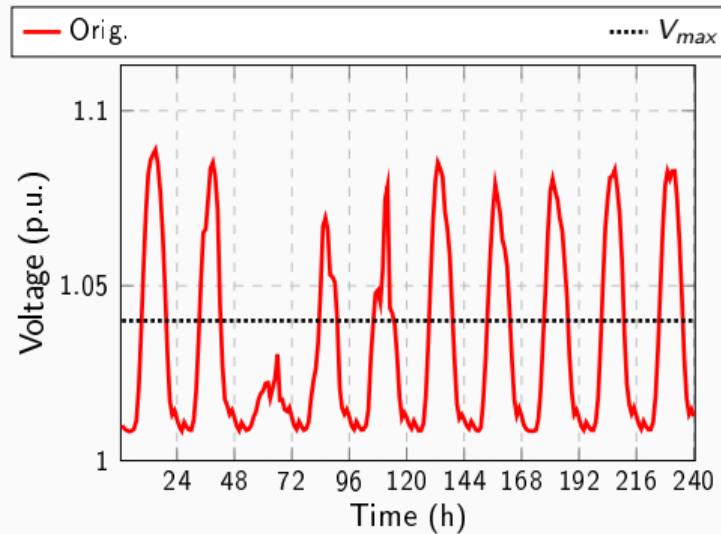
## Control actions

- Active Power Curtailment (APC)
- Reactive Power Control (RPC)
- Battery Energy Storage System (BESS)
- Controllable load (CL)
- On Load Tap Changers (OLTC)

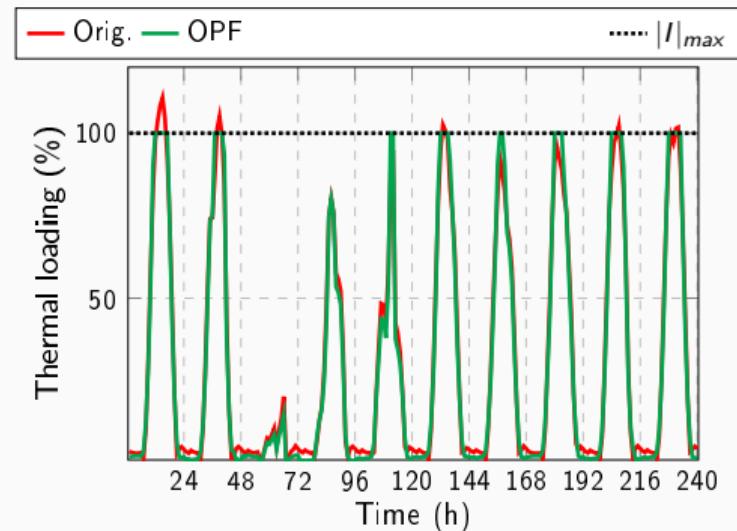
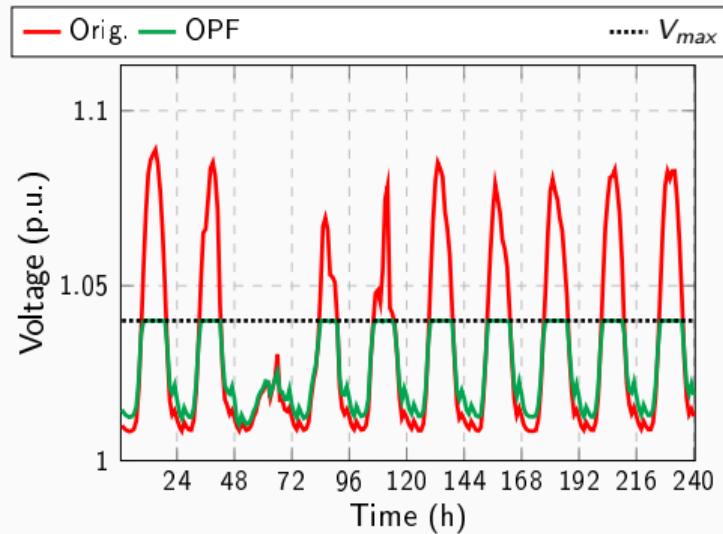
## Network description

- Based on European CIGRE LV grid
- Normalized profiles
  - PV & forecasts: Real data from Zurich
  - Load: Typical profiles based on CIGRE
- Summer day simulations
  - High solar radiation
- Acceptable limits:
  - Voltage:  $\pm 4\%$
  - Current: up to 1 p.u.
  - VUF: up to 2%

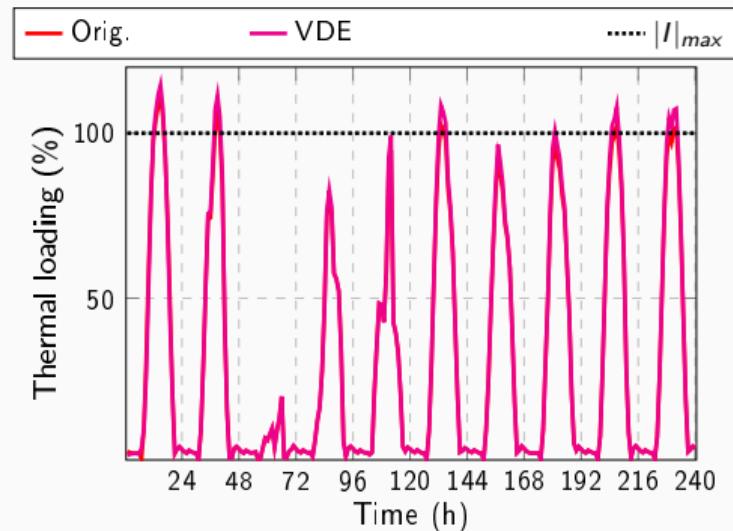
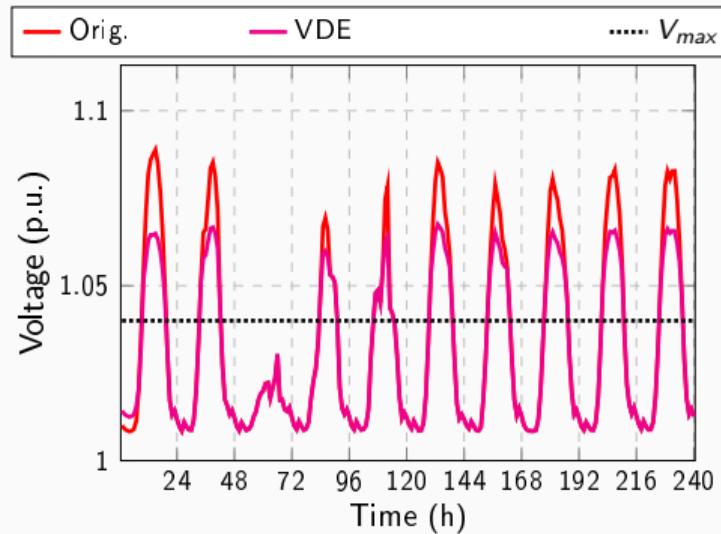
## Some results



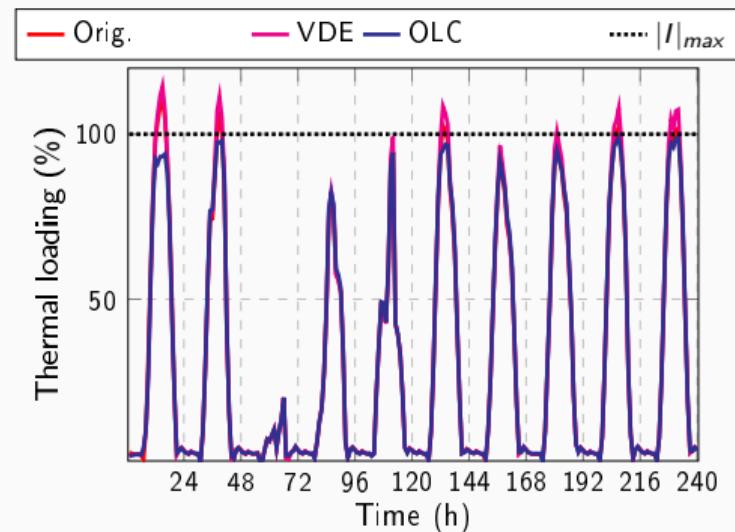
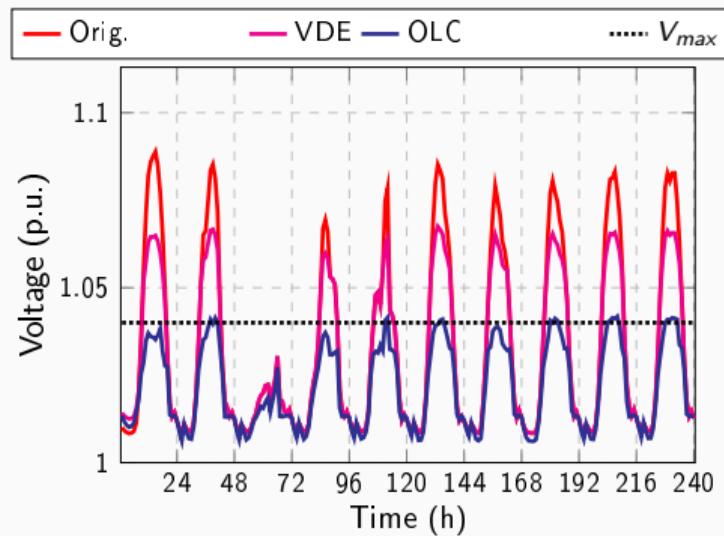
## Some results



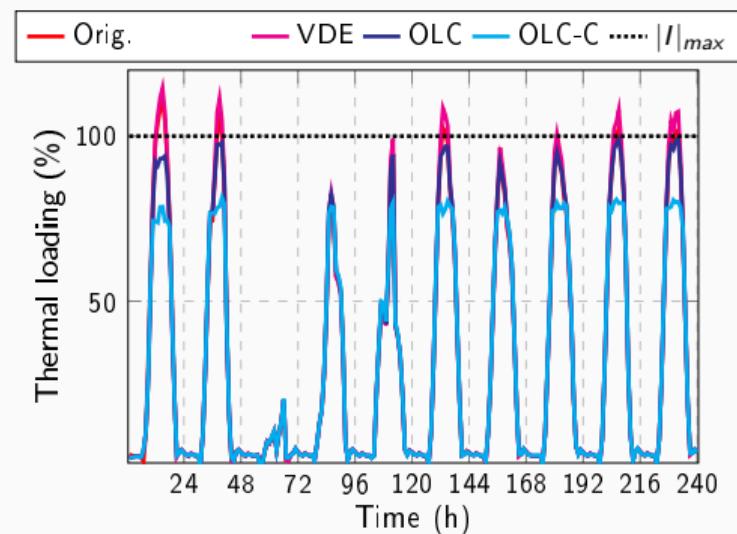
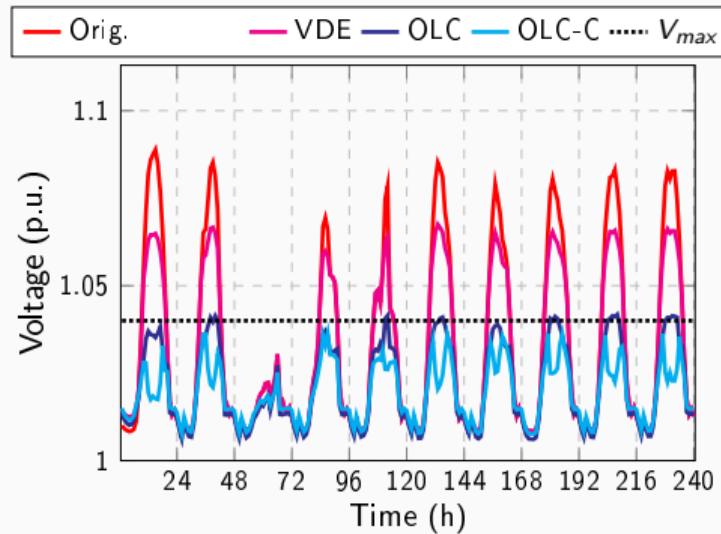
## Some results



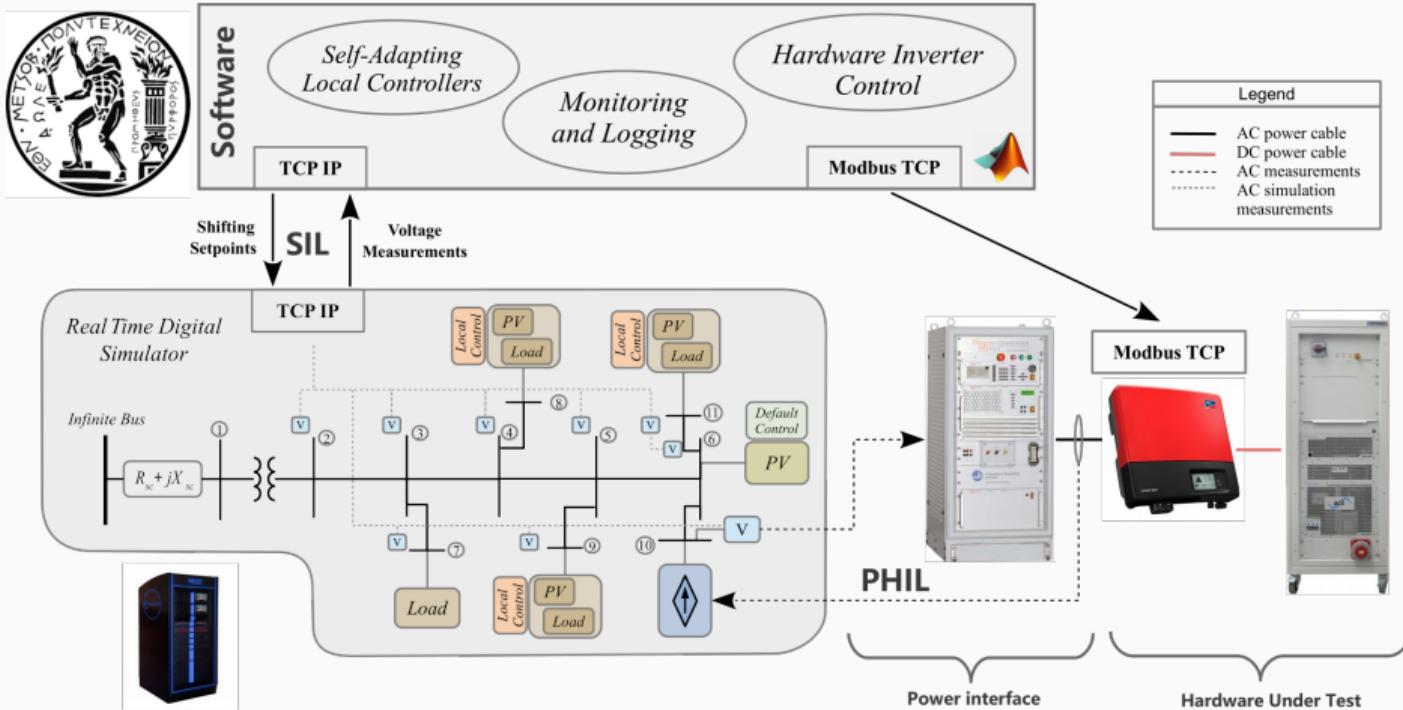
## Some results



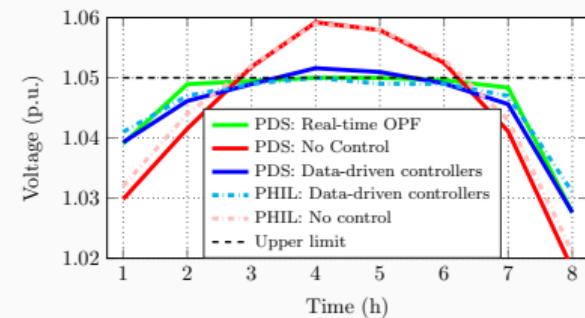
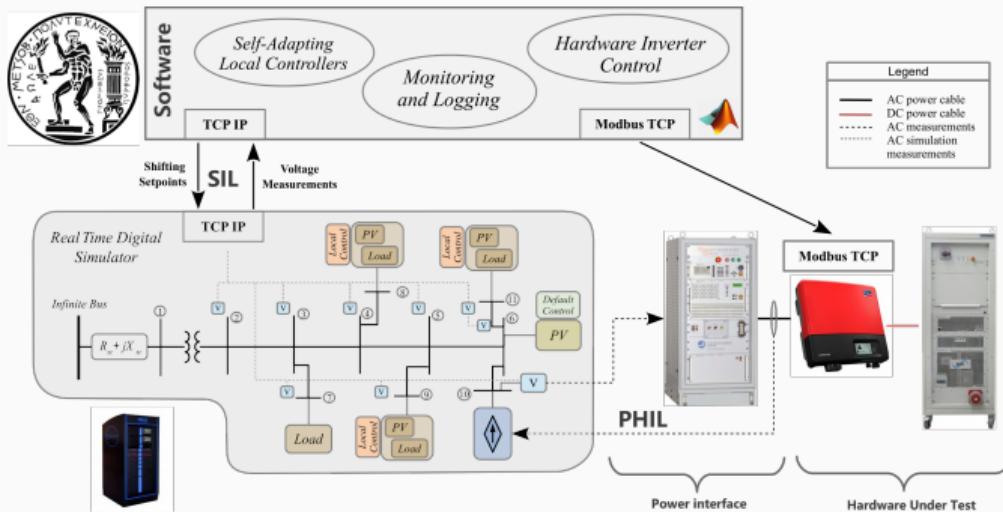
## Some results



# How about a hardware validation?



# How about a hardware validation?



## Concluding remarks

---

# Concluding remarks

## Summary

- Most of the new *Smart Grid-driven* developments are located in distribution grids
- Lack *monitoring, communication, and remote control* infrastructure

# Concluding remarks

## Summary

- Most of the new *Smart Grid-driven* developments are located in distribution grids
- Lack *monitoring, communication, and remote control* infrastructure
- **Centralised** controllers have great performance but high cost and robustness concerns
- **Local** controllers are robust and low cost but cannot cope with modern challenges

# Concluding remarks

## Summary

- Most of the new *Smart Grid-driven* developments are located in distribution grids
- Lack *monitoring, communication, and remote control* infrastructure
- **Centralised** controllers have great performance but high cost and robustness concerns
- **Local** controllers are robust and low cost but cannot cope with modern challenges

**Data-driven optimised local controllers can bridge the gap**

# Concluding remarks

## Summary

- Most of the new *Smart Grid-driven* developments are located in distribution grids
- Lack *monitoring, communication, and remote control* infrastructure
- **Centralised** controllers have great performance but high cost and robustness concerns
- **Local** controllers are robust and low cost but cannot cope with modern challenges

**Data-driven optimised local controllers can bridge the gap**

## Future steps

- Extend to different types of controls and embed grid-supporting services.
- Single-stage data-driven control design → distributionally robust optimisation?
- Self-adapting controls through online learning to handle changing conditions → (Deep)RL?

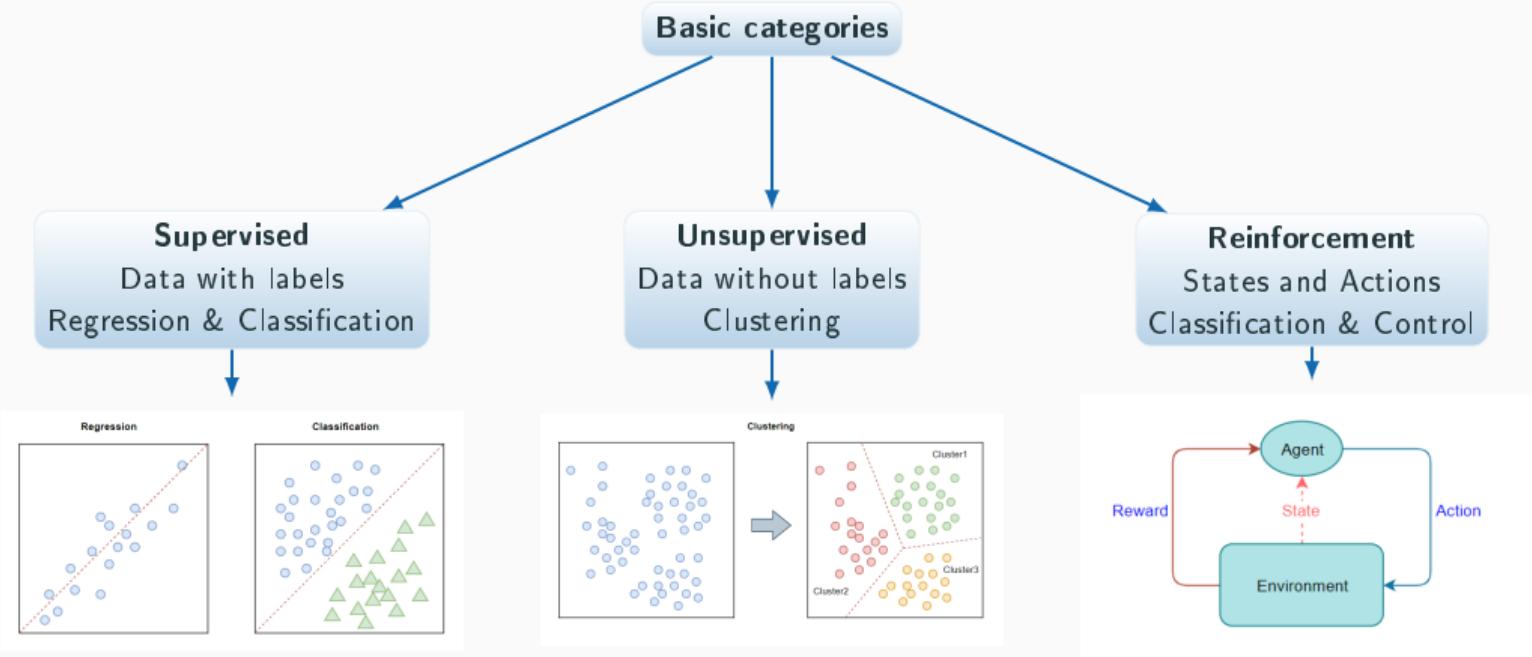
# Questions?



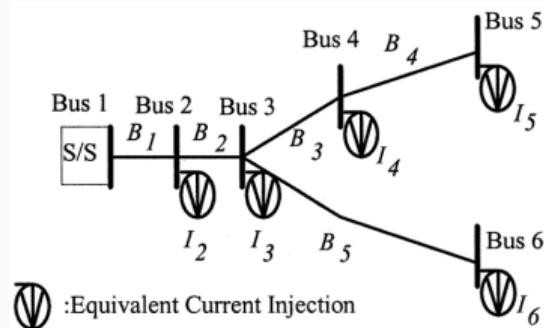
## Backup slides

---

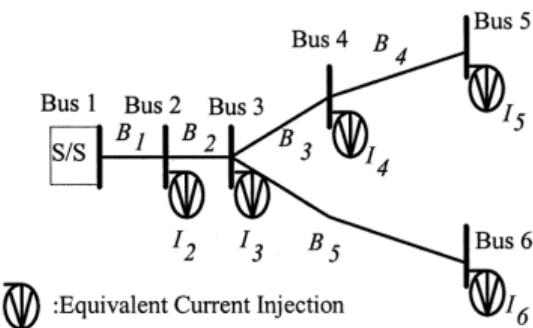
# Machine learning in power systems



# Backward/Forward Sweep (BFS) power flow



# Backward/Forward Sweep (BFS) power flow



- Bus-injection to branch-current (BIBC)
- Branch-current to bus-voltage (BCBV)

---

**Input:** BIBC, BCBV,  $P_{inj}$ ,  $Q_{inj}$ ,  $V_{slack}$

**initialize:**  $k = 1$ ,  $V_{bus}^k = 1\angle 0^\circ$

**do**

**Backward sweep:**  $I_{inj}^k = \left( \frac{(P_{inj} + jQ_{inj})^*}{V_{bus}^{k*}} \right)$

$$I_{br}^k = BIBC \cdot I_{inj}^k$$

**Forward sweep:**  $\Delta V^{k+1} = BCBV \cdot I_{br}^k$

$$V_{bus}^{k+1} = V_{slack} - \Delta V_{tap} \cdot \rho_t + \Delta V^{k+1}$$

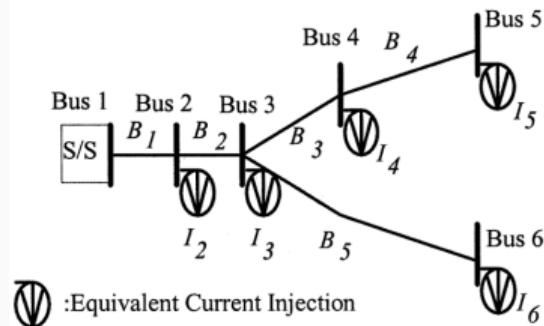
**Update iteration:**  $k += 1$

**while**  $\max(|V_{bus}^k| - |V_{bus}^{k-1}|) \geq \bar{\eta}$

---

**Output:**  $I_{br}^k$ ,  $V_{bus}^{k+1}$

# Backward/Forward Sweep (BFS) power flow



- Bus-injection to branch-current (BIBC)
- Branch-current to bus-voltage (BCBV)

---

**Input:** BIBC, BCBV,  $P_{inj}$ ,  $Q_{inj}$ ,  $V_{slack}$

**initialize:**  $k = 1$ ,  $V_{bus}^k = 1\angle 0^\circ$

**do**

**Backward sweep:**  $I_{inj}^k = \left( \frac{(P_{inj} + jQ_{inj})^*}{V_{bus}^{k*}} \right)$

$$I_{br}^k = BIBC \cdot I_{inj}^k$$

**Forward sweep:**  $\Delta V^{k+1} = BCBV \cdot I_{br}^k$

$$V_{bus}^{k+1} = V_{slack} - \Delta V_{tap} \cdot \rho_t + \Delta V^{k+1}$$

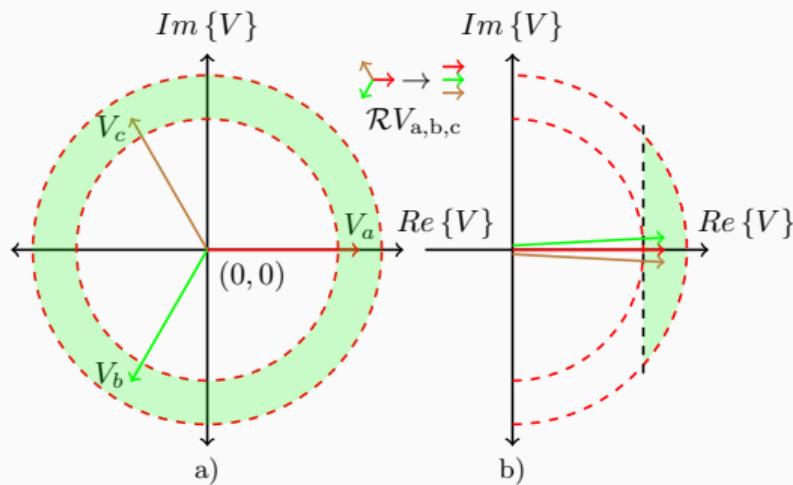
**Update iteration:**  $k += 1$

**while**  $\max(|V_{bus}^k| - |V_{bus}^{k-1}|) \geq \bar{\eta}$

**Output:**  $I_{br}^k$ ,  $V_{bus}^{k+1}$

---

# OPF voltage constraints relaxation



# Computational time of deterministic OPF solutions

## Methods

1. BFS-OPF using YALMIP and Gurobi as the interface platform and solver.
2. Standard AC OPF using the nonlinear and non-convex power flow equations, and IPOPT as the solver of the non-convex problem.
3. Standard AC OPF similar to the previous case plus providing functions for the gradient and hessian of the objective function, the Jacobian of the nonlinear constraints, and the Hessian of the Lagrangian.

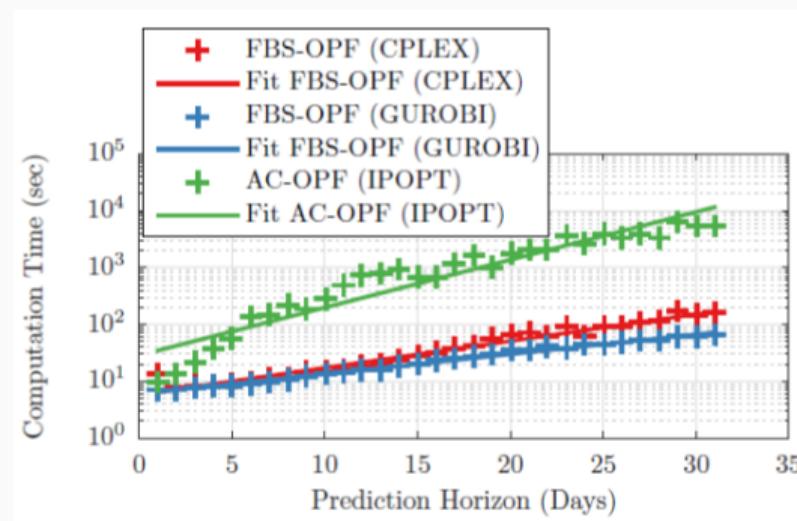
Solving Time (sec)	APC	+RPC	+BESS	+CL
Method 1	2.03	2.47	1.55	2.68
Method 2	9.47	12.53	18.04	16.99
Method 3	0.86	2.61	1.65	1.21

Solving time for the Cigre LV grid.

Solving Time (sec)	APC	+RPC	+BESS	+CL
Method 1	2.59	1.48	1.51	1.57
Method 2	21.41	41.63	41.56	34.61
Method 3	1.92	2.15	2.72	2.75

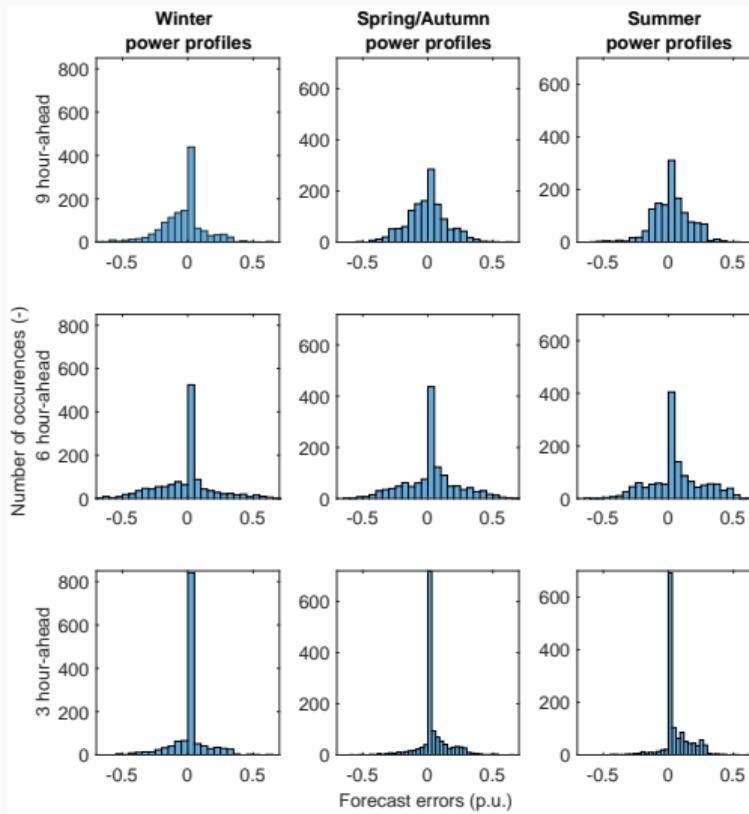
Solving time for the MV-LV grid.

# Computational time of deterministic OPF solutions



Computation time grows polynomially

# Forecast errors of PV generation in Zurich data (EWZ)



# Uncertainty tightening procedure

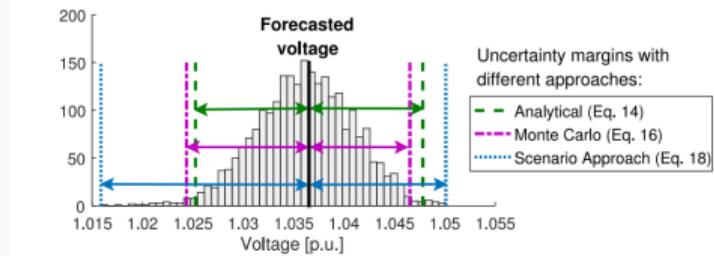
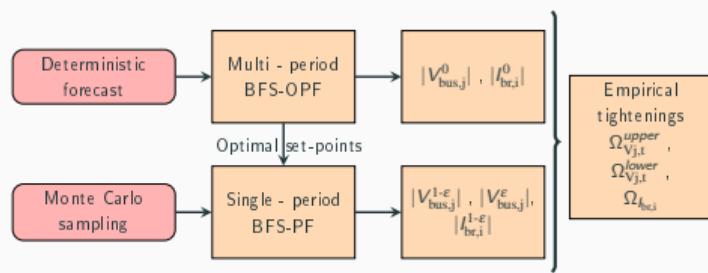
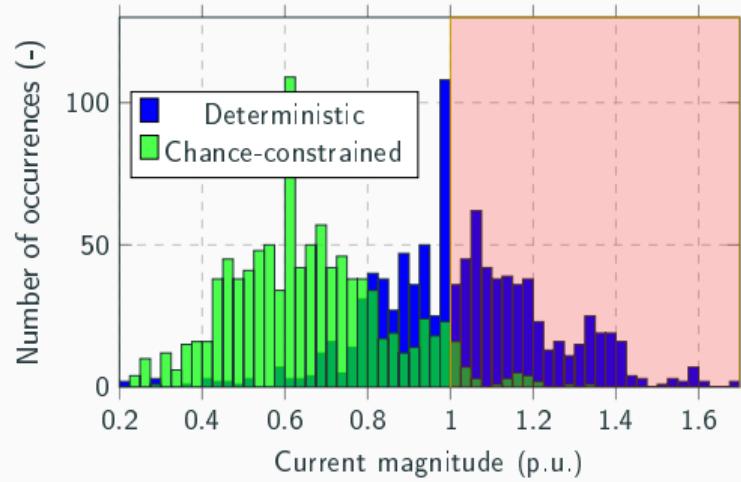
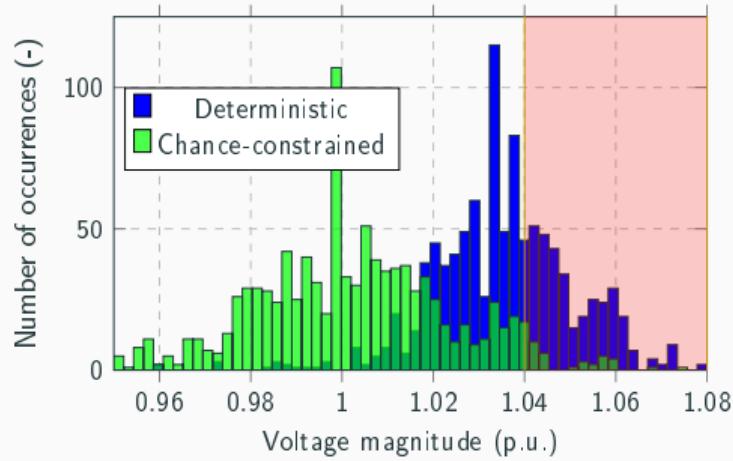
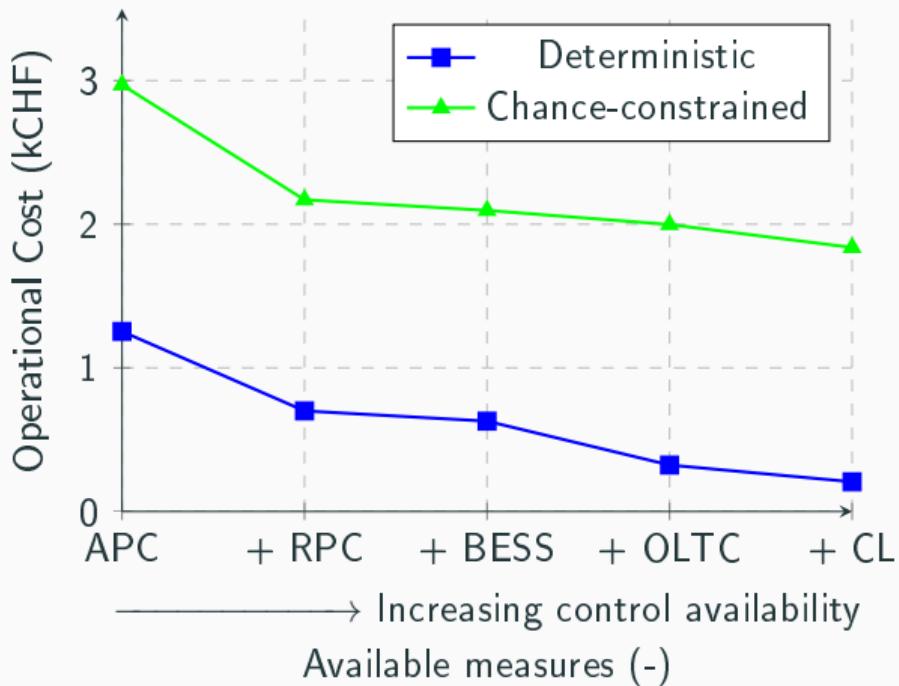


Fig. 1. Uncertainty margins  $\lambda_{V,j}^u, \lambda_{V,j}^l$  for an example voltage constraint, as obtained with a given set of samples. The black line represents the forecasted voltage, and the histogram shows the empirical voltage distribution for the given sample set. The uncertainty margins are given by the distance between the black line and the corresponding coloured lines: Analytical reformulation (green), Monte Carlo simulation (purple) and Scenario approach (blue).

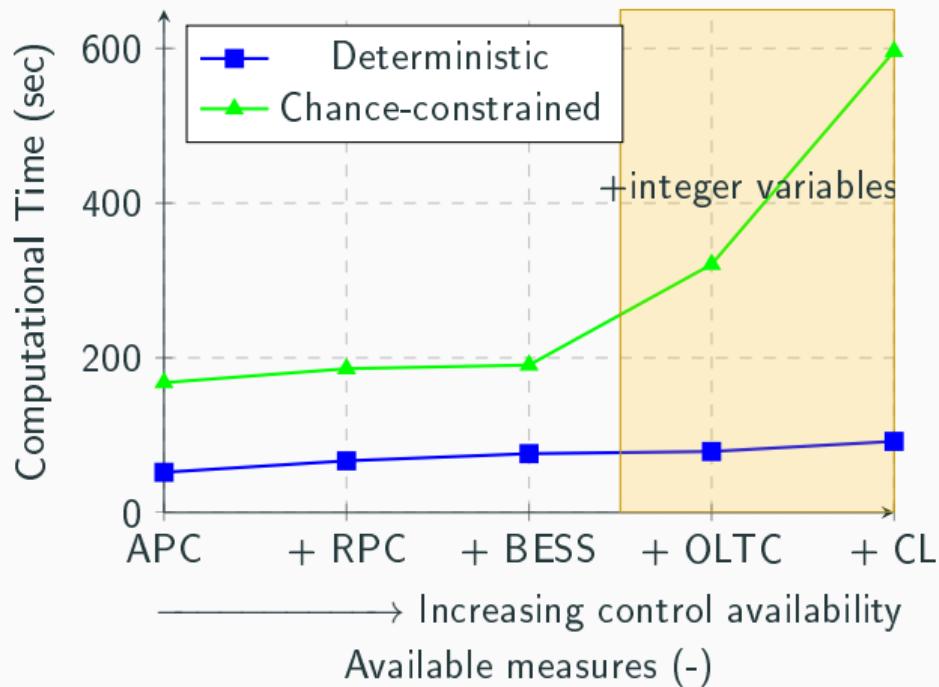
# Comparison of violations due to forecast errors



## Objective function cost



# Computational time of OPF solutions



## Some results

Method	VDE	OPF	OLC
Losses (%)	4.60	4.42	4.45
$ V _{\max}$ (p.u.)	1.069	1.04	1.045
$ I _{\max}$ (%)	119.94	100	99.49
$VUF_{\max}$ (%)	1.81	1.98	2.33
$P_{\text{curt}}$ (%)	0	1.08	2.03

# Linear SVR

Given the training data

$$(x_i, y_i) \quad i = 1, \dots, m$$

Minimise:

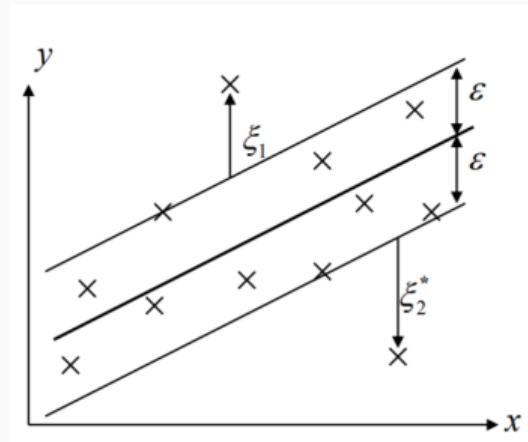
$$\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*)$$

Subject to:

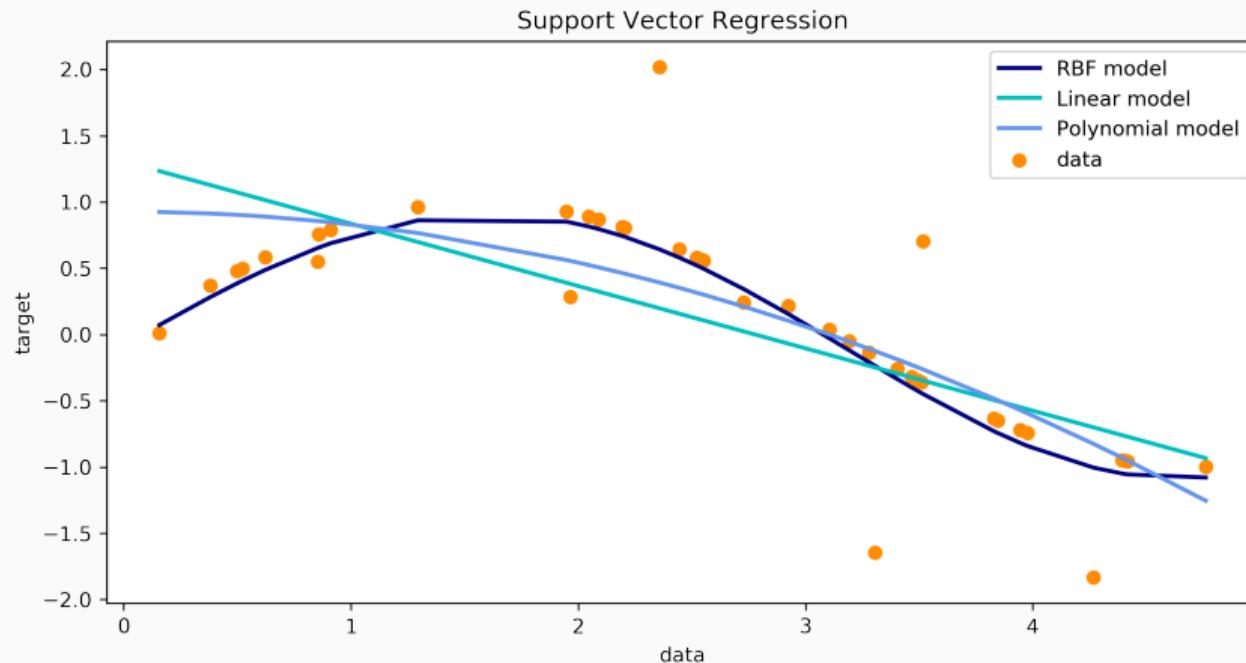
$$y_i - (\mathbf{w} \cdot \mathbf{x}_i) - b \leq \varepsilon + \xi_i$$

$$(\mathbf{w} \cdot \mathbf{x}_i) + b - y_i \leq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0 \quad i = 1, \dots, m$$



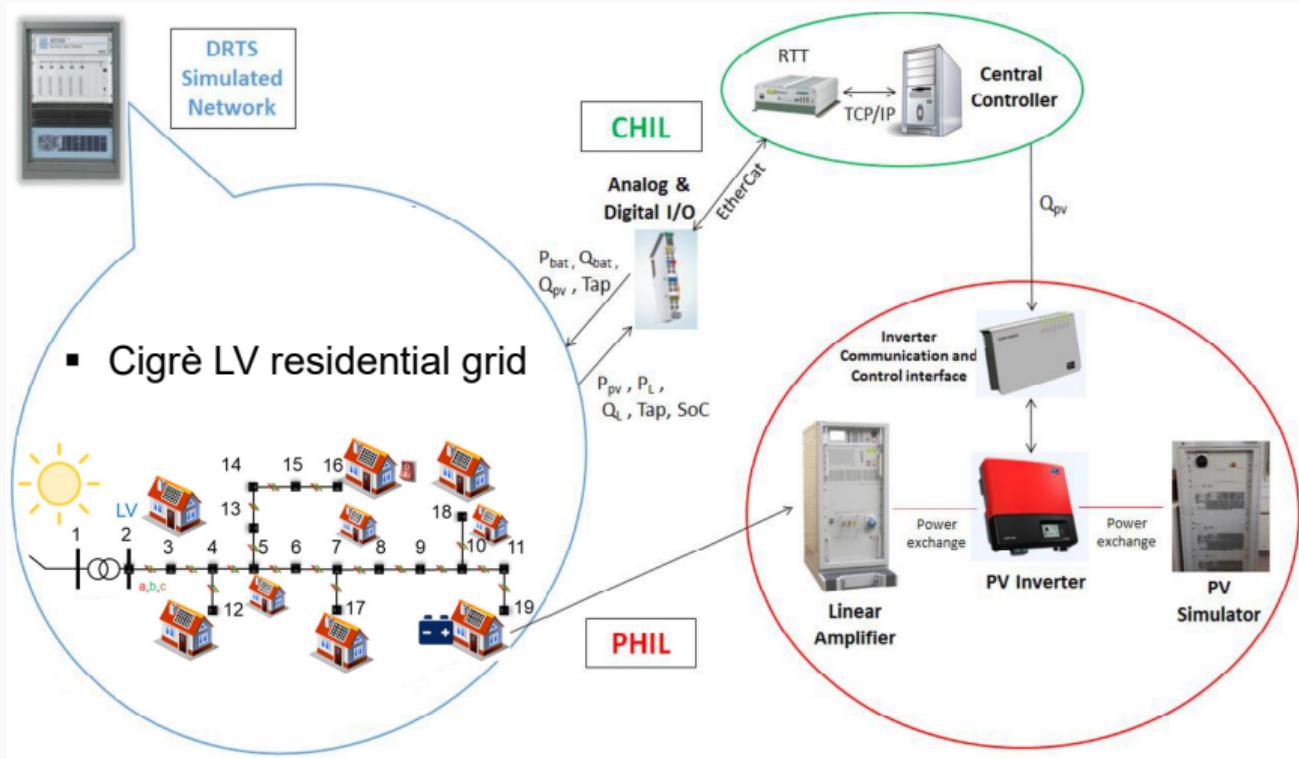
# SVR Kernels



## SVR Kernels

Node	C	Kernel Type	$\gamma$	Polynomial Order
12	4.742	RBF	0.046	—
16	710.250	polynomial	-	4
18	36.243	RBF	0.002	—
19	682.37	polynomial	—	2

# Planned testing



## Results comparison

10 sum. days	Orig.	OPF	VDE	OLC	OLC-C
losses (%)	6.49	6.17	7.07	6.59	6.45
$V_{max}$ (p.u.)	1.1073	1.04	1.0808	1.0412	1.0403
$V_{MSE}$ (-)	9.73E-05	0	5.68E-05	6.65E-06	1.88E-05
$V_{MAE}$ (-)	0.0041	0	0.0035	0.0013	0.0023
$P_{curt}$ (%)	0	5.74	0	4.61	4.34

# Hierarchical agglomerative clustering algorithm

---

**Input:**  $X^j$ ,  $d$

- 1: Define each voltage series  $X^j$  as its own cluster.
- 2: Calculate the pair-wise distances between the clusters (based on  $d$ ).
- 3: Merge the closest 2 clusters based on the minimum average linkage value.
- 4: Repeat Steps 2-3 until there is only one cluster, comprising all the voltage series

**Output:** Dendrogram structure of the input data

---