

Clustering Data-driven Local Control Schemes in Active Distribution Grids

Stavros Karagiannopoulos, *Member, IEEE*, Gustavo Valverde, *Senior Member, IEEE*, Petros Aristidou, *Senior Member, IEEE*, and Gabriela Hug, *Senior Member, IEEE*

Abstract—Controllable Distributed Energy Resources (DERs) in Active Distribution Grids (ADGs) provide operational flexibility to system operators thereby offering the means to address various challenges. Existing local controllers for these resources are communication-free, robust, and cheap, but with sub-optimal performance compared to centralized approaches that heavily rely on monitoring and communication. Data-driven local controls can bridge the gap by providing customized local controllers designed from historical data, off-line optimization, and machine learning methods. These local controllers emulate the optimal behavior, under expected operating conditions, without the use of communication. However, they exhibit high implementation overhead with the need of individual programming of DER controllers, especially when there are many DERs or when new units are installed at a later stage. In this paper, we propose a clustering method to decrease the implementation overhead by reducing the individual DER controls into a smaller set while still achieving high performance. We show the performance of the method on a three-phase, unbalanced, low-voltage, distribution network.

Index Terms—Time-series clustering, optimal control, data-driven control design, active distribution networks, OPF, machine learning

I. INTRODUCTION

Distribution networks (DNs) experience significant changes through the introduction of large shares of Distributed Generators (DGs), such as Photovoltaics (PVs) or wind turbines, and new load types. These renewable and intermittent resources are mostly connected to the DN through power electronic-based devices that allow controllability in terms of active and reactive power. This gives DNs the option to provide ancillary services to higher voltage levels [1].

Whereas centralized, e.g. [2], [3] and distributed, e.g. [4], schemes rely -to a certain extent- on communication among the DERs, local control schemes, e.g. [5], utilize only local information to decide on the DER response. Such local schemes exist already in grid codes [6] since they are simple to implement, robust, and cheap due to null communication

requirements. However, such a one-size-fits-all approach is suboptimal, when the same control parameters are employed to all DERs irrespective of their location and surrounding grid topology.

Data-driven control methods that design customized local controllers by using off-line optimization techniques, historical data and machine learning methods are lately explored [7]–[11]. In [7], we proposed the idea of using data to design controllers that emulate the OPF behavior, considering reactive and active power control via characteristic curves. In [8], the curves are derived by support vector machines, while [9] considered also controllable loads and storage systems using machine learning tools, such as segmented regression and support vector machines as regressors and classifiers. Initially, [7] designed the curves based on piece-wise linear segments, and accounted for the needed monotonicity and slope constraints of the curves ex-post, similar to [8]. In [9] on the other hand, we relied on segmented regression with unknown breakpoints considering the needed constraints within an optimization problem to design the curves. Finally, the authors in [10], [12] use multiple linear regression to control reactive power in an open-loop fashion, without voltage magnitudes as inputs, and [11] compares all different methods based on the input features and their response to unseen conditions. However, all these works calculate a unique characteristic curve for each DER, imposing implementation challenges in large DNs with high penetration of DERs.

Clustering the local rules that dictate the DER response would ease the transition to a customized local control scheme derived from an off-line optimal response. Clustering approaches have been used in power systems for various purposes. Many works, e.g. [13]–[16], focus on the demand side identifying consumers with similar consumption profiles, on network segmentation for preventive islanding to minimize power disruption [17], or voltage collapse area identification [18], phase identification [19] and forecasting, e.g. [20], [21]. Finally, the authors of [22] provide an extensive review of clustering methods in the renewable energy sector. In [13], the authors use a hierarchical agglomerative approach to cluster static smart meter data, and reduce the dimensions of the regression problem that calculates the system’s voltage sensitivities. In [14], [15], partitional clustering based on the k-means algorithm is applied to smart meter data, in order to group together similar consumers based on various features. However, the main focus was on the grouping itself and not on deriving a summarized consumption profile. In most cases of the reviewed papers, clustering was used to cope with the

This research is part of the activities of the Swiss Centre for Competence in Energy Research on the Future Swiss Electrical Infrastructure (SCCER-FURIES), which is financially supported by the Swiss Innovation Agency (Innosuisse - SCCER program).

S. Karagiannopoulos and G. Hug are with the Power Systems Laboratory, ETH Zurich, 8092 Zurich, Switzerland. Email: {karagiannopoulos | hug}@eeh.ee.ethz.ch.

P. Aristidou is with the Department of Electrical Engineering, Computer Engineering & Informatics, Cyprus University of Technology. Email: petros.aristidou@cut.ac.cy

G. Valverde is with the Electric Power & Energy Research Laboratory, University of Costa Rica, 11501-2060 UCR, Costa Rica. Email: gustavo.valverde@ucr.ac.cr

dimensionality problem when investigating grids with many components.

In this paper, we focus on the state-of-the-art, data-driven, local control schemes and aim at tackling implementation challenges. However, the proposed methodology can be implemented for any individual local control scheme as long as it is expected that altering the default scheme would be beneficial. First, we present the method to derive the data-driven local control schemes, using historical data, off-line Optimal Power Flow (OPF) calculations, and Machine Learning (ML) methods. The derived local schemes rely on segmented regression with unknown breakpoints [9]. Then, we propose a method of creating customized controllers for each DER, taking into account the implementation challenges in the presence of numerous DERs. The benefit of clustering the individual controllers is threefold. First, it offers a solution to DNs with hundreds of individual DERs, which is realistic in many distribution grids across the world. Instead of designing and implementing a unique control rule for each of them, we can cluster the available controllers in a finite set which can be downloaded by all DERs regularly. Although the overhead of *computing* individual optimal control laws is not significant and is considered in the off-line methodology, there are implementation challenges, i.e. how to implement the customized control laws in the possibly thousands of DER controllers. Programming each DER controller separately is a challenging task in active DNs without advanced communication infrastructure and hundreds of controllable units. However, by clustering the control laws into a finite set, we propose to communicate the clustered set to all DERs and then assign only the selected control law to each DER. Modern inverters, e.g. [23] offer the possibility of updating the firmware and operating with custom rules. Thus, receiving the clustered set of control laws e.g. every four months is straightforward. In real-time, sending an integer that will assign the final control law to each DER is relatively easy and is already a simple and common way to control appliances in many countries. For instance, in Switzerland, households' appliances such as heat pumps or hot water boilers can be blocked by this kind of ripple control, where a signal can change the status of a controller to 'blocked' or 'no blocked' [24]. This would reduce significantly implementation challenges without sacrificing the quality. Second, the clustering provides a solution to assign optimized control rules to new DERs being installed in DNs. Relying on individual control laws for each DER would require computing and assigning the new optimal control laws to every single controllable unit. Thus, the computationally expensive training methodology can be avoided when new units are connected, by assigning them to clustered control schemes according to similarity measures. Finally, the clustering method provides useful information about coherent nodes according to their optimal behavior. For example, one can identify which nodes do not face power quality issues and can be used to provide ancillary services or loss minimization and which nodes should focus on local power quality issues, e.g. operate in inductive mode to tackle overvoltage problems, e.g. [25].

More specifically, the contributions of this paper are:

- A method to cluster individual customized local control

schemes, in the form of volt/VAr and volt/watt characteristic curves, to decrease the implementation overhead in customized local controls for ADGs.

- A comparative performance analysis between using individual, customized control curves and the curves resulting from the clustering.
- An analysis of the impact of using standard (text-book) grid parameters in the design stage, when real grid parameters are not available.

The general overview of the proposed methodology comprises three stages and is sketched in Fig. 1. In Stage I, off-line calculations, based on OPF formulations, process the historical data (network parameters and load and generation profiles) and define the *optimal* DER control setpoints for different operating conditions. At this stage, different objectives can be formulated, such as loss minimization, maximization of self-consumption, maximization of the DN revenues by provision of ancillary services, or combinations of these. System security and power quality are ensured through hard constraints that need to be respected by the OPF. However, the final real-time operation does not provide any guarantees that network limits will not be violated. The importance of the features selection in terms of the robustness in the real-time operational is addressed in [11] which highlights the use of voltage magnitude as a local feature that brings global information. In [26] the authors propose an online self-adapting algorithm to solve local power quality issues under changing operating conditions, since constraint satisfaction cannot be guaranteed without continuously updating the training dataset with new operating conditions [27].

In Stage II, we apply ML techniques on the optimal setpoints obtained from *Stage I* to design local real-time DER controls. The goal is to derive simple and efficient, individual (i.e., one for each DER), optimized local controls that can mimic the behavior of centralized OPF-based schemes without the need of any communication infrastructure. In the last step of this stage, a clustering method is proposed to reduce the number of the optimized controllers and overcome the implementation burden of assigning unique controllers to each DER. Finally, in Stage III, we evaluate the proposed method, by comparing the real-time response against existing local schemes and a real-time OPF scheme.

The remainder of the paper is organized as follows. In Section II, we present the mathematical formulation of the centralized OPF scheme (*Stage I*). In Section III, we describe the data-driven local control design scheme and the methodology to cluster these local curves (*Stage II*). In Section IV, we present a case study and the simulation results, and, finally, we conclude in Section V.

II. STAGE I - CENTRALIZED OFFLINE OPF SCHEME

The centralized scheme is used *offline* in order to calculate the optimal response of various conditions (normal or not) that will be captured by the optimized local schemes described in Section III-A. Thus, in real-time operation no monitoring and communication infrastructure is required.

As input, we require the topology of the network as well as the installed capacity of DERs and loads, information that

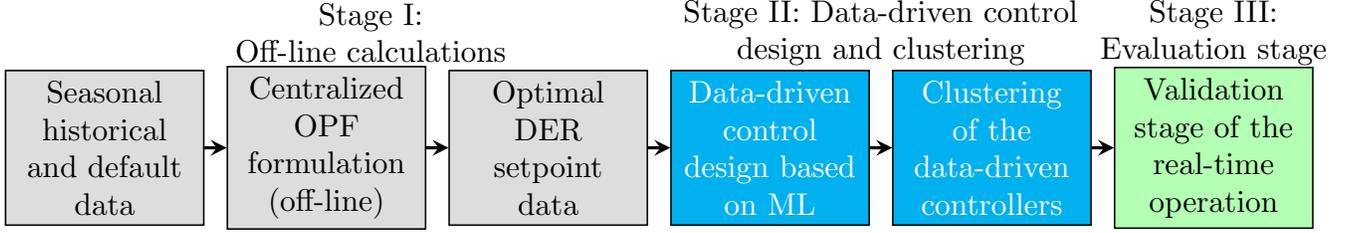


Fig. 1: General overview of the data-driven scheme and the proposed clustering methodology.

is usually available to the operators. Some other required data might not be available (e.g. normalized solar radiation, load profiles, or cable impedances). In this case, we can use typical/default values from the literature, but the impact of this uncertainty needs to be examined and quantified. As output, we derive optimal setpoints which will comprise the training dataset for the data-driven approach of Stage II.

There are various centralized schemes that can be used to derive the optimal DER setpoints. In this work, we use an OPF framework derived from [2], [28], [29] that integrates a backward-forward sweep (BFS) power flow into the optimization problem. The main advantages of this scheme include high computational performance, AC feasibility, and handling of weakly meshed grids. The interested reader is referred to [9], [29] for the three-phase modeling of the DN and the DERs. It is important to emphasize that any other suitable OPF formulation for ADGs can be used to derive the optimal DER setpoints.

A. OPF formulation

The task of the DN operator is to guarantee safe grid operation minimizing the system losses and operating costs. Here, we assign cost to the curtailment of active energy and provision of reactive power support by DGs. The objective function is evaluated by considering the DER control cost over all network nodes N_b , branches N_{br} and the entire time horizon T_{OPF} , i.e.

$$\min_{\mathbf{u}} \sum_{t=1}^{N_{OPF}} \sum_{z \in \{a,b,c\}} \left\{ \sum_{j=1}^{N_b} \left(C_P \cdot P_{c,j,z,t} + C_Q \cdot Q_{ctrl,j,z,t} \right) + \sum_{i=1}^{N_{br}} C_P \cdot P_{loss,i,z,t} \right\} \cdot \Delta t \quad (1)$$

where \mathbf{u} denotes the vector of the available active control measures and Δt the length of each time period. The curtailed power of the DERs connected at phase z , at node j and time t is calculated by $P_{c,j,z,t} = P_{g,j,z,t}^{\max} - P_{g,j,z,t}$, where $P_{g,j,z,t}^{\max}$ is the maximum available active power and $P_{g,j,z,t}$ the active power injection of the DERs. The use of reactive power provided by the DERs connected to phase z of node j and time t , i.e. $Q_{ctrl,j,z,t} = |Q_{g,j,z,t}|$ is minimized, where $Q_{g,j,z,t}$ represents the DER reactive power injection or absorption. The DER cost of curtailing active power and providing reactive power support (DER opportunity cost or contractual agreement) is represented by the coefficients C_P and C_Q , respectively. Priority is given to the utilization of reactive power control, i.e. we set $C_Q \ll C_P$. Finally, the total losses in the three-phase framework are calculated by using the difference between

input and output power in each branch of each phase [30]. Thus, $P_{loss,i,z,t} = Re(S_{j_s,z,t} + S_{j_r,z,t})$, where $S_{j_s,z,t}$ and $S_{j_r,z,t}$ represent the apparent power flowing into branch i from each end; j_s and j_r are the sending and receiving ends of the branch.

The power injections at every node j , phase z and time step t are given by

$$P_{inj,j,z,t} = P_{g,j,z,t} - P_{l,j,z,t}, \quad (2a)$$

$$Q_{inj,j,z,t} = Q_{g,j,z,t} - P_{l,j,z,t} \cdot \tan(\phi_{load}), \quad (2b)$$

where $P_{l,j,z,t}$ and $P_{l,j,z,t} \cdot \tan(\phi_{load})$ are active and reactive node demands of constant power type, with $\cos(\phi_{load})$ being the power factor of the load.

A single iteration of the BFS power flow problem is considered to represent the power flow constraints, following our previous work [9]. That is: $\mathbf{I}_{br,t} = \mathbf{BIBC} \cdot \left(\frac{(P_{inj,j,z,t} + jQ_{inj,j,z,t})^*}{\bar{V}_{j,z,t}^*} \right)$ and $\mathbf{V}_t = \mathbf{V}_{slack} + \mathbf{BCBV} \cdot \mathbf{I}_{br,t}$, where $\bar{V}_{j,z,t}^*$ is the voltage of phase z , at node j at time t , $*$ indicates the complex conjugate and the bar indicates that the value from the previous iteration is used (the interested reader is referred to [9] for more details); $\mathbf{I}_{br,t}$ is the vector of the three-phase branch flow currents; and, \mathbf{BIBC} (Bus Injection to Branch Current) is a matrix with ones and zeros, capturing the three-phase topology of the DN (including any single-phase laterals); $\Delta \mathbf{V}_t = \mathbf{BCBV} \cdot \mathbf{I}_{br,t}$ is the vector of voltage drops over all branches and phases; \mathbf{BCBV} (Branch Current to Bus Voltage) is a matrix with the complex impedance of the lines as elements (including mutual coupling); \mathbf{V}_{slack} is the three-phase voltage in per unit at the slack bus (here assumed to be $\{1 < 0^\circ, 1 < -120^\circ, 1 < 120^\circ\}$). Thus, the constraint for the current magnitude for all branches i and phase z at time t is

$$|\mathbf{I}_{br,t}| \leq \mathbf{I}_{max}, \quad (3)$$

where \mathbf{I}_{max} is the maximum thermal limit of the three phases.

For the voltage magnitude constraint, we follow [29] and rotate the three voltage phases $\{a, b, c\}$ by $\mathcal{R} = \{1 < 0^\circ, 1 < 120^\circ, 1 < -120^\circ\}$ to avoid the non-convex $V_{min} \leq |V_{j,z,t}| \leq V_{max}$ constraints yielding

$$|\mathcal{R}V_{j,z,t}| \leq V_{max}, \quad \text{Re}\{\mathcal{R}V_{j,z,t}\} \geq V_{min}. \quad (4)$$

Finally, the limits of the inverter-based PVs are given by

$$P_{g,j,z,t}^{\min} \leq P_{g,j,z,t} \leq P_{g,j,z,t}^{\max}, \quad (5a)$$

$$Q_{g,j,z,t}^2 \leq (S_{inv,j}^{\max})^2 - P_{g,j,z,t}^2, \quad (5b)$$

where $P_{g,j,t}^{\min}$, $P_{g,j,t}^{\max}$ are the lower and upper limits for active DER power, and $S_{inv,j}^{\max}$ is the capacity of inverter j .

After we obtain the optimal OPF setpoints, we perform an exact power flow calculation to derive an AC feasible operating point. The voltages of this point are used in the next OPF iteration, and the loop is repeated until we reach convergence in terms of voltage magnitude mismatch. The interested reader is referred to [9] for more details regarding the three-phase OPF algorithm.

III. STAGE II - DATA-DRIVEN LOCAL CONTROL DESIGN

In this section, we describe the mathematical model of the used local data-driven control schemes, detailed in [7], [9]. This stage uses as input the optimal setpoints derived from the OPF problem of the previous Stage, and provides as output first the individual customized control laws for each DER and finally the clustered set which is the main focus of this paper. Here, we focus on a closed-loop scheme, using the *voltage magnitude* as a local feature to control active and reactive power of the DERs. These volt/watt and volt/VAr curves are similar to the ones used today in modern grid codes. In contrast to existing standards, the proposed curves are composed of an arbitrary number of piece-wise linear segments and they are optimized for each DER based on its location and the DN objectives.

There are various data-driven methods which differ in terms of the number and type of local measurements they rely on. The voltage magnitude measurement is a very important local feature since it *carries* global information from the whole network due to the physics of the system [11]. Therefore, this fact can be used in order to design controllers that can perform adequately in conditions that were not met in the training phase. Thus, many require only one feature, e.g. local voltage [7], or a set of measurements, such as local demand, generation and maximum capacities [10]. Although, a larger set of measurements can better map the OPF setpoints into a model, the inclusion of voltage magnitudes is essential for secure system operation under unexpected conditions.

Regarding notation, the real-time response of the j^{th} inverter-based DER ($j \in [1, 2, \dots, N_j]$) in terms of reactive power control $q_t^{(j)}$ and active power curtailment $c_t^{(j)}$ is derived from the N_{OPF} optimal setpoints ($t \in [1, 2, \dots, N_{OPF}]$) obtained in the offline calculations which are described in Section II. The scope of data-driven schemes is to derive models which will mimic the optimal setpoints using only local input measurements (features). In other words, we try to emulate the optimal setpoints of a multi-dimensional space, through rules that depend only on local features. The feature matrix $\Phi^{(j)} \in \mathbf{R}^{T_{OPF} \times N_K}$ contains as columns the N_K features and as rows the T_{OPF} observations of the k^{th} input measurement $\phi_k^{(j)} \in \mathbf{R}^{T_{OPF}}$, i.e.

$$\Phi^{(j)} = \begin{bmatrix} \phi_1^{(j)} & \phi_2^{(j)} & \dots & \phi_{N_K}^{(j)} \end{bmatrix}. \quad (6)$$

To simplify notation, we omit the subscripts of the phase z for the rest of the section. In the rest of the paper, we focus on the simplest scheme that relies only on local voltage magnitudes, and can be easily implemented in the form of characteristic

curves that define the real-time DER response in terms of active and reactive power.

A. Design of optimized local control schemes

In this part, we present the procedure to derive the piece-wise linear curves, originally presented in [9]. We calculate the characteristic curves for reactive power control and active power curtailment, by applying segmented-regression, optimizing also the placement of the break-points. Instead of solving the non-linear and non-differentiable problem of obtaining the piece-wise linear fitting problem with unknown breakpoints, we use an iterative algorithm and solve a Residual Sum-of-Squares (RSS) optimization problem inspired by [31]. The procedure is summarized below.

First, we define the number of break-points n_s , initialize them, and solve for each inverter j the following residual sum of squares problem

$$RSS_{\tilde{i}} := \min_{\tilde{x}_0, \beta, \gamma} \sum_{t \in T_{OPF}} P_{g,j,t} \cdot (x_t - \tilde{x}_t)^2 + \sum_{k=1}^{n_s} \gamma_k^2, \quad (7)$$

subject to

$$\tilde{x}^{\tilde{i}} = \tilde{x}_0 \cdot \mathbf{1}^T + \beta_0 \cdot \Phi^{(j)} + \sum_{k=1}^{n_s} \beta_k \cdot (\Phi^{(j)} - s_k^{\tilde{i}}) \cdot I(\Phi^{(j)} > s_k^{\tilde{i}}) + \sum_{k=1}^{n_s} \gamma_k \cdot I(\Phi^{(j)} > s_k^{\tilde{i}}), \quad (8)$$

$$\beta_0 \leq 0, \quad \beta_0 + \beta_1 \leq 0, \quad \dots, \quad \beta_0 + \sum_{k=1}^{n_s} \beta_k \leq 0, \quad (9)$$

$$|\beta_0| \leq \beta_{max}, \quad |\beta_k| \leq \beta_{max}, \quad (10)$$

where $\tilde{x}^{\tilde{i}} \in \{p, q\}$ refers to the active and reactive final models, and $\Phi^{(j)} = [|V_{j,t}|]$ is the vector of voltage magnitudes used as input to the fitting problem. We fit the linear model based on the known breakpoints $s_k^{\tilde{i}}, \forall k = 1, \dots, n_s$ at the current iteration \tilde{i} , the left slope β_0 and difference-in-slopes β_k . The indicator function $I(\cdot)$ becomes one when the inside statement is true. Finally, \tilde{x}_0 is the model intercept and γ a parameter which updates the location of the breakpoints towards the optimal one. The monotonicity constraint (weakly decreasing for the volt/VAr case) is imposed by (9), and the slope constraints by (10) avoids sudden changes of the control actions. After the problem is solved, we update the breakpoints $s_k^{\tilde{i}+1} = \frac{\gamma_k}{\beta_k} + s_k^{\tilde{i}}$ and iteration index $\tilde{i} = \tilde{i} + 1$, repeating the procedure until the RSS does not change between two subsequent iterations, i.e. when $RSS_{\tilde{i}} - RSS_{\tilde{i}-1} \approx 0$.

The same method is used for both the active power curtailment and reactive power control curves, using respectively the PV optimal active and reactive setpoints from the OPF in Section II.

B. Clustering of local control schemes

In this section, we describe the process of clustering the individual local controls of the DERs into a finite set. Clustering refers to the task of grouping similar objects together, i.e. objects within a cluster are similar to each other, and

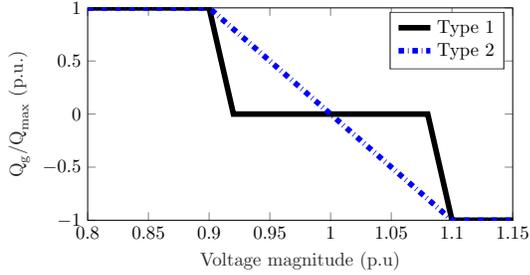


Fig. 2: Two types of local reactive power control. Type 1 provides continuous reactive power control, while Type 2 includes a deadband where no reactive power is utilized.

dissimilar to objects of other clusters. There are various clustering algorithms that differ mainly in terms of a) the applied measure to calculate distance or similarity among the objects, (e.g. Euclidean distance), b) the cluster model representation they use to describe a cluster (e.g. connectivity, medoid-based or centroid-based models), c) the input they need (e.g. the total number of clusters, the need to have equal length of trajectories) and d) their suitability to cluster static data, or trajectories (e.g. clustering over time-series data). The evaluation and suitability of these clustering algorithms depend on the application.

In this paper, we are interested in clustering different DER characteristic curves, and thus, we investigate trajectory or time-series clustering (in our case voltage-series similar to the two versions of local volt/VAr curves described in grid codes, e.g. [6], [32] and shown in Fig. 2). In most cases of time series-clustering, the procedure is similar to the static clustering algorithms, but the notions of similarity and prototype function are altered to account for the whole time-series dataset. The prototype is a time-series that effectively summarizes the most important characteristics of all series in a given cluster. This time-series is referred to as an ‘average series’, and prototyping refers to time-series averaging [33]. The interested reader is referred to [34] for a comprehensive review of time-series clustering algorithms, which organizes recent works in terms of the aforementioned types, computational complexity, applications and evaluation measures. Another general overview is given in [33], where the author not only categorizes different time-series clustering algorithms, but also provides implementation examples in R [35].

In our application, the time series are actually voltage series and we will therefore refer to the series to be clustered as voltage or data series. The voltage series to be clustered have by design the same length. Thus, in our comparison we use the two most common clustering algorithms, namely the hierarchical and the partitioning types, that are both readily applicable in this case.

1) Clustering Types:

a) *Hierarchical clustering*: This type constructs a hierarchy of clusters following one of two approaches. In the agglomerative one, each time-series starts as its own cluster, followed by pair-wise merging until we build up the hierarchy in a bottom-up scheme. On the contrary, the divisive approach follows a top-down scheme, i.e. all the data-series start in

Algorithm 1 Hierarchical agglomerative clustering algorithm

Input: \mathbf{X}^j, d

Output: Dendrogram structure of the input data

- 1: Define each voltage series \mathbf{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.

Return: Tree illustration of the clusters’ arrangement

one cluster and they are split recursively as we move down the hierarchy. Let us denote with $\mathbf{X}^j = (X^1, \dots, X^{N_j})^T$ the vector containing the N_j voltage series in terms of the active or reactive characteristic curves each of which comprising n_q elements, e.g. $X^j = [x_1^j, x_2^j, \dots, x_{n_q}^j]$, and $d(\cdot)$ a distance measure. Algorithm 1 summarizes the iterative procedure for the agglomerative approach. Hierarchical clustering is computationally costly when dealing with a lot of data, but the results can be illustrated in a dendrogram, providing intuition about similarities in the used datasets and the needed number of the final clusters, which is denoted by N_{cl} .

b) *Partitioning clustering*: In this clustering approach, the amount of final clusters is decided a-priori. The data-series are disaggregated into the decided amount of clusters each of which is then represented by the data-series centroid or medoid. The centroids are formed by taking the mean (e.g. k-means algorithm) values of the data series in the respective cluster and are calculated by an iterative approach which minimizes the total distance between all members of a cluster and their cluster prototype. Similarly, another commonly used approach is to use partition around medoids (PAM), where a medoid is a representative data-series object from a cluster, whose average distance to all other objects in the same cluster is minimal. Medoids are similar to means or centroids, but they are restricted to be members of the cluster’s dataset [33]. Algorithm 2 summarizes the iterative procedure implemented in many software platforms, such as in [35]. We denote $\mathbf{C}^{j_{cl}} = (C^1, \dots, C^{N_{cl}})^T$ as the vector containing the $N_{cl} < N_j$ cluster trajectories.

2) Distance measures:

There are many distance, or dissimilarity measures in order

Algorithm 2 Partitioning clustering algorithm

Input: \mathbf{X}^j, d, N_{cl}

Output: $\mathbf{C}^{j_{cl}}$ final clusters

- 1: Initialize N_{cl} centroids/medoids randomly.
- 2: Calculate the distances d between each voltage-series and the centroids/medoids and then assign each object to the cluster of its closest centroid/medoid.
- 3: Apply a prototyping function to each cluster and update the centroid/medoid.
- 4: Repeat Steps 2-3 until no object is changing its cluster.

Return: $\mathbf{C}^{j_{cl}} = f(\mathbf{X}^j, d, N_{cl})$

to compare and group the different time-series into clusters. The choice among the various alternatives [36] is not straightforward and depends on the scope of the grouping. As we deal with normalized data-series, i.e. from -1 p.u. to 1 p.u., data-series of the same length, and are interested in a "shape-based" clustering, we use the following distance measures:

a) *Minkowski*: This is the most commonly used distance measure that is defined by

$$d_{L_c}(X^j, X^{\tilde{j}}) = \left(\sum_{n=1}^{n_q} (x_n^j - x_n^{\tilde{j}})^c \right)^{1/c}, \quad (11)$$

where $j, \tilde{j} \in N_j$ are two time-series of X , and c is a positive integer. Typically, it is used with $c = 2$ representing the Euclidean distance and $c = 1$ for the Manhattan distance. Although this measure provides very good results in many applications, it is sensitive to shifting and time scaling (stretching or shrinking of the time axis). Using this measure, we can construct the cluster prototypes, i.e. the clustered volt/Var curve, based on the mean ($c = 1$) or median ($c = 2$) values of the voltage-series within a cluster.

b) *Dynamic time warping*: To avoid the drawbacks of the Minkowski measure, we also use the Dynamic Time Warping (DTW) distance measure, where the sequences are warped such that their distance is minimized. Following [36], we denote M as the set of all possible sequences of m pairs, that warp the time-series and are used to calculate the minimum distance. We preserve the observations order in the form $r = ((x_{\alpha_1}^j, x_{\beta_1}^{\tilde{j}}), \dots, (x_{\alpha_m}^j, x_{\beta_m}^{\tilde{j}}))$ with $\alpha_\zeta, \beta_\zeta \in \{1, \dots, n_q\}$, such that $\alpha_1 = \beta_1 = 1$, $\alpha_m = \beta_m = n_q$, and $\alpha_{\zeta+1} = \alpha_\zeta$ or $\alpha_\zeta + 1$ and $\beta_{\zeta+1} = \beta_\zeta$ or $\beta_\zeta + 1$ for $\zeta \in \{1, \dots, m-1\}$. Thus, the DTW distance is given by

$$d_{DTW}(X^j, X^{\tilde{j}}) = \min_{r \in M} \left(\sum_{\zeta=1, \dots, m} |(x_{\alpha_\zeta}^j - x_{\beta_\zeta}^{\tilde{j}})| \right). \quad (12)$$

In the case study, we evaluate the hierarchical and partitioning clustering schemes using the Euclidean and the DTW distances.

IV. CASE STUDY - RESULTS

A. Network description - Case study setup

To demonstrate the performance of the proposed approach, we compare four methods using the benchmark radial residential LV grid presented in [37] and shown in Fig. 3. The load and PV panels are distributed to the three phases unevenly, resulting in unbalanced conditions. The total load, taken from [37], is shared 25%-60%-15% among the three phases. The installed PV capacity, is set to $S_{\text{rated}}^{\text{PV, total}} = 150\%$ of the total maximum load of the entire feeder and connected equally to the PV nodes = [3, 5, 7, 10, 12, 16, 17, 18, 19]. The shares among the three phases are 25%-25%-50% and real data from a PV station in Switzerland are used to derive realistic PV profiles for the investigation.

In the design stage, we use the algorithm presented in Section II to process a 30-day summer dataset with forecasts of load and PV production, and generate the optimal DER

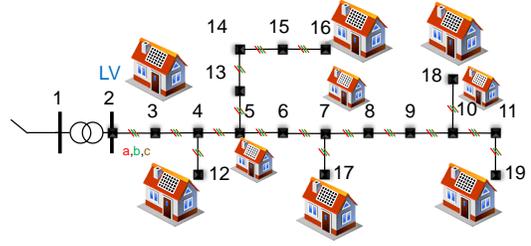


Fig. 3: European three-phase residential LV grid based on [37].

setpoints. Then, the algorithm is evaluated based on a new, unseen dataset. The operational costs of the centralized problem are assumed to be $c_p = 0.3 \frac{\text{CHF}}{\text{kWh}}$ and $c_Q = 0.01 \cdot c_p$. The implementation was done in MATLAB using YALMIP [38] as the modeling layer and Gurobi [39] as the solver. For the clustering schemes, we used R [35], and all the results were obtained on an Intel Core i7-2600 CPU and 16 GB of RAM.

B. Individual optimized local control schemes

Figure 4 presents the local volt/Var characteristic curves of all PV units in phase C, assuming perfect knowledge about the grid's parameters. The PV at node 3 in general injects reactive power to optimize losses and reduce the reactive power needs from the substation, while the remaining nodes absorb reactive power in order to keep voltages smaller than the maximum acceptable value of 1.04 p.u..

In order to examine the impact of erroneous input data regarding the impedances of the cables, we investigate the robustness of the clustered characteristic curves under erroneous impedances, i.e. assuming that the impedance of the cables deviates from the actual values reported in [37]. This can be the case in distribution grids where the operator cannot measure the actual values and uses default values instead, or when the impedance parameters have changed over time. Figure 5 shows the optimized local curves for three nodes at phase C after superimposing a uniformly distributed random error to the impedances of each branch up to $\pm 10\%$ (superscript ') and $\pm 20\%$ (superscript ''), respectively. These nodes show the greatest difference compared to the perfect information case of Fig. 4. We observe, that up to $\pm 10\%$ error in the parameters of the grid does not have a significant influence on the optimal setpoints and hence, on the final derived local curves. On the contrary, errors up to $\pm 20\%$ change the behavior of the PV unit

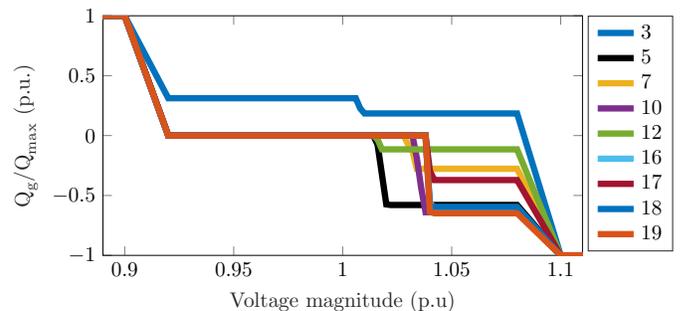


Fig. 4: Individual local characteristic curves for reactive power control of the PV units at phase C.

TABLE I: Summarized clustering results for all methods

	Clustering Type	Distance Measure	Prototype	Cluster Members	Average inter-cluster distance
1	Hierarchical	Euclidean	PAM centroids	[3], [5 10 16 18 19], [7 12 17]	[0], [0.593], [0.521]
2	Hierarchical	DTW	PAM centroids	[3], [5 10 16 18 19], [7 12 17]	[0], [0.940], [1.647]
3	Partitional	Euclidean	mean	[3], [5 10 16 18 19], [7 12 17]	[0], [0.692], [0.542]
4	Partitional	Euclidean	median	[3], [5 10 16 18 19], [7 12 17]	[0], [0.594], [0.545]
5	Partitional	Euclidean	PAM centroids	[3], [5 10 16 18 19], [7 12 17]	[0], [0.593], [0.521]
6	Partitional	DTW	mean	[3], [5 10 16 18 19], [7 12 17]	[0], [0.409], [0.287]
7	Partitional	DTW	median	[3], [5 10 16 18 19], [7 12 17]	[0], [0.190], [0.203]
8	Partitional	DTW	PAM centroids	[3], [5 10 16 18 19], [7 12 17]	[0], [0.187], [0.238]

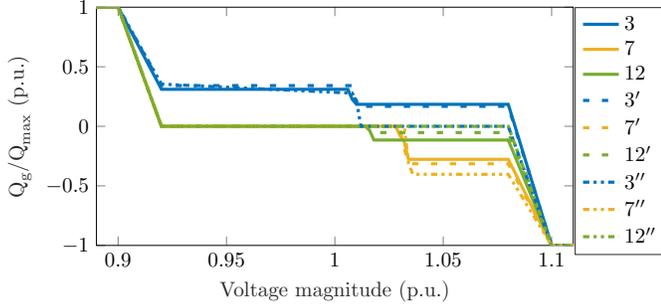


Fig. 5: Comparison of individual local characteristic curves for reactive power control of three PV units at phase C, using erroneous cable impedances.

at node 3, rendering its behavior from capacitive to inactive above the voltage magnitude of 1.01 p.u.. By observing that in all nodes the erroneous curves are shifted downwards, we can conclude that the erroneous cables' impedances are rather overestimated, i.e. the true values are lower than the used ones. Thus, they impose a behavior that favors more consumption of reactive power than the curves using the true network parameters.

C. Clustering of the optimized local control schemes

In this part, we apply the hierarchical and partitioning clustering algorithms in order to obtain a reduced set of local rules. The optimal number of clusters in data clustering has a crucial impact on the quality of the final solution. However, there is no consensus in terms of theoretical methods that can calculate the optimal number of needed clusters. The

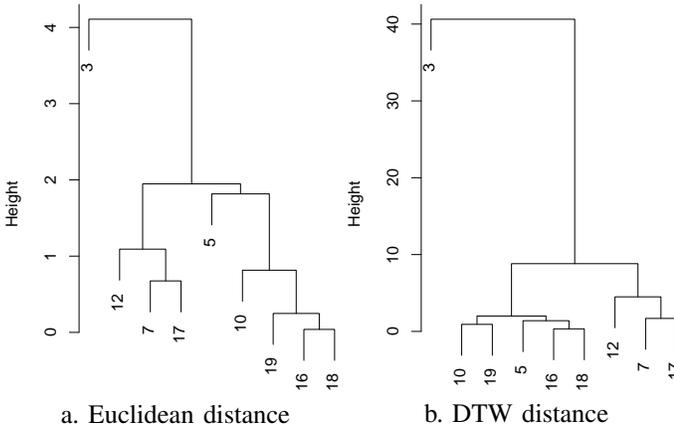


Fig. 6: Hierarchical clustering based on different distance measures.

optimal number depends on the method used for measuring similarities, the main scope of the data-driven scheme, and the parameters used in the partitioning process. There are various metrics that can evaluate the performance of clustering algorithms, and these can be used as indicators to assist in the selection of the needed cluster number [40]–[42]. The Calinski–Harabasz and the Davies–Bouldin indices introduced in [40] and [41], respectively, are based on the means of the time-series as the clustering centers, while the silhouette index presented in [42] is based on the clustering validity of each time-series separately. Typically, by inspecting the dendrogram produced using a hierarchical clustering approach in time-series data provides some intuition regarding the number of clusters that can summarize the behavior of the original data in a satisfactory way, as they are based on the same similarity measures as the aforementioned indices. Figure 6 shows the dendrograms derived by the Euclidean and DTW distance measures using the true values of the cables' impedances. We observe that although there are some differences in the derived tree diagrams, they both lead to the same conclusions irrespective of the distance measure used: a) Three clusters group the curves in the most efficient way, b) the local curve of the PV inverter at Node 3 is very dissimilar compared to the other rules, which is also observed in Fig. 4.

We choose the number of clusters $N_{cl} = 3$, construct the different clusters and assign them to the respective inverters. Table I summarizes the main characteristics of two clustering types analyzed under 8 different configurations. The approaches differ in terms of the used distance measure and the selection of the prototype, i.e. the series that effectively summarizes the most important characteristics of all series in a given cluster. PAM provides the advantage of not having to examine the monotonicity and slope constraint requirements of the medoid, since this is already considered in the design stage of the individual curves. Finally, the cluster members and the average inter-cluster distance are given for all configurations. We observe that all the methods converge to the same grouping of the characteristic curves into 3 clusters, irrespective of the selected type and distance measure. The differences in the clustered curves are marginal and result from the selection of the cluster prototype.

Figure 7 shows the clustering of the characteristic curves, and the centroid prototype of each cluster according to configuration 6. As it has been already observed from Table I, all configurations select the characteristic for the PV inverter at node 3 to comprise one cluster, which is dissimilar to all

the other curves. Furthermore, using the mean approach to construct the centroid prototypes results in the creation of new clustered curves. On the contrary, all the other methods, i.e. based on the PAM centroids or the median, select one time-series dataset as the cluster centroid. More specifically, these methods result in the following centroid prototypes: $cl_1 = [3]$, $cl_2 = [16]$, $cl_3 = [7]$, where the numbers correspond to the volt/VAr curve of the respective PV unit.

The same procedure is followed for the other two phases, and for the volt/watt curves for active power curtailment.

D. Comparison of power flow results

We compare the behavior of the individual and clustered local schemes running power flow computations, under the following configurations:

- Method 0: The DERs are operating according to the German grid-code [32], where all DERs become inductive as soon as they inject more than 50% of their installed capacity. The power factor decreases linearly from 1 to 0.95 or 0.9 based on the DER installed capacity.
- Method 1: All DERs are controlled based on the centralized OPF-based algorithm summarized in Section II. Here we assume perfect knowledge of the grid's impedances and perfect communication and monitoring infrastructure. We use this scheme as the benchmark for the best achievable performance.
- Method 2: All DERs are operating according to the individual controls according to Section III-A derived using a uniformly distributed random error up to $\pm 10\%$ of the true impedances.
- Method 3: All DERs are operating according to clustered controls according to Section III-B using the same error values as in Method 2.

Table II summarizes the results from applying these methods in real-time operation for a test period of one month. Representing the benchmark, Method 1 satisfies all security constraints in the optimal way. Method 0 (standard industry practice) results in high losses, due to increased reactive power needs by the PV units, without solving the overvoltage and overload issues. Finally, Methods 2 and 3 mitigate adequately the overvoltage and overload issues, while being capable of mimicking the OPF-based control without the need of communication. However, they result in more active power curtailment than needed, making the losses comparison inconsistent. Finally, we observe that the clustered scheme behaves

TABLE II: Summarized monthly results for all methods

Method	0	1	2	3
Losses (%)	5.75	5.38	4.84	4.76
$ V _{\max}$ (p.u.)	1.073	1.04	1.046	1.047
$ I _{\max}$ (%)	124.04	100	89.47	87.57
P_{curt} (%)	0	1.39	4.64	4.8

similar to Method 2, with marginal differences arising from the ‘‘averaged’’ and fewer local control schemes.

Figure 8 displays, as an example, the voltage magnitude evolution of Node 19, phase C, over the first evaluation week. Here, the same conclusions as before can be drawn: Methods 2 and 3 mimic the optimal centralized solution (Method 1) without communication needs, and current industrial practice (Method 0) can be insufficient in cases of large DER penetration. The Root-Sum-of-Squares Values (RSSVs) over the whole evaluation month compared to the voltage magnitudes of the centralized OPF, i.e. Method 2, is summarized as follows: $RSSV_0 = 0.2416$, $RSSV_2 = 0.0693$, $RSSV_3 = 0.0664$. We observe that Method 3 results in marginally smaller values due to the higher active curtailment which results in voltage magnitudes closer to Method 1.

E. Sensitivity analysis with respect to the number of clusters

In this part, we highlight the importance of selecting the right number of clusters, by considering two additional configurations. More specifically, we compare Methods 2 and 3 against the following setups:

- Method 4: All DERs are grouped into 2 clusters using the PAM method as the prototype according to Sect. III-B with erroneous cable impedances up to $\pm 10\%$.
- Method 5: Finally, the same setting is applied here, but the DERs are now grouped into 4 clusters.

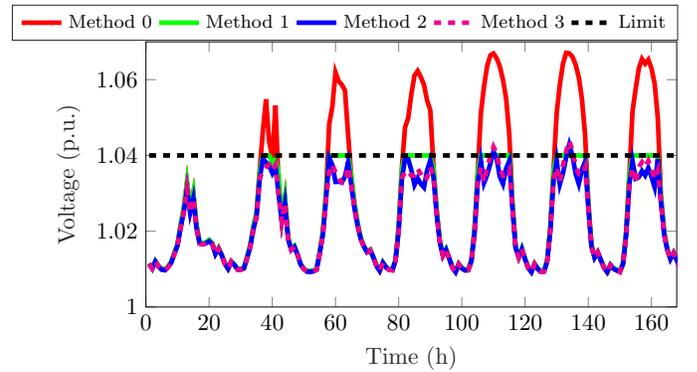


Fig. 8: Voltage magnitude evolution at Node 19, phase C.

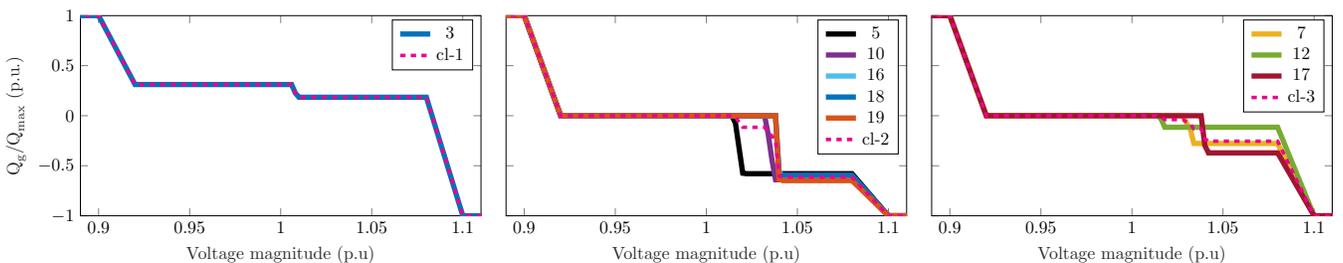


Fig. 7: Clustering and prototypes of the local characteristic curves for reactive power control of the PV units at phase C.

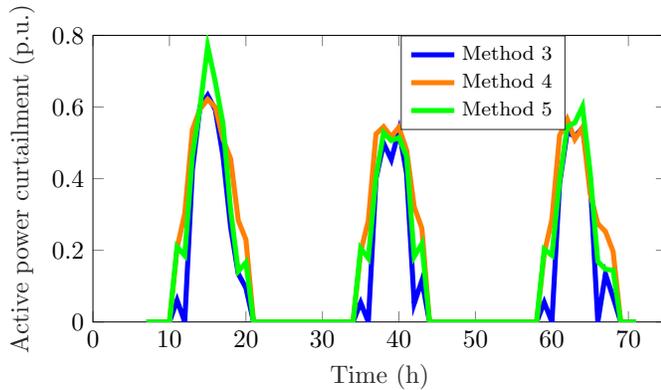


Fig. 9: Active power curtailment at Node 19, phase C varying the number of clusters in the optimized local control schemes.

The more clusters we allow, the closer the performance is to Method 2, which allows a unique model for each DER. On the contrary, less clusters lead to a behavior further away from the optimal. This is observed also in Fig. 9 that compares the hourly active power curtailment of Node 19, phase C, among Methods 3, 4, and 5 over the period of 3 days. Method 4 uses only 2 clustered curves for reactive and active power curtailment for all units, resulting in higher total curtailment than Methods 3 and 5. Over the period of the whole evaluation month, Method 4 required 75% more total active power curtailment than Method 3. Finally, although Method 5 employs more curves, it results in a marginal benefit of 2.28% less curtailment than Method 3.

V. CONCLUSION

More and more controllable DERs are connected to the grid making a safe, reliable and efficient grid operation challenging. Central OPF-based approaches for the coordination of the DERs can achieve optimal results, but rely on communication and monitoring infrastructure which raises privacy and investment concerns. On the other hand, purely local schemes are robust and cheap, but incapable of coping with the modern DN challenges. Lately, optimized data-driven local control schemes show promising results, mimicking the centralized optimal behavior using only local features. However, the individual customized controls needed in these methods increase the implementation overheads and complexity since they require setting the unique customized control law to each DER separately. In this paper, we tackle the practical challenges of requiring unique control rules, by clustering them into a limited set. All DERs can easily download the clustered control laws during firmware upgrade and the method is generic to every inverter-based DER. For the real-time operation a very simple signal can be used to set each DER to one control law of the clustered set. By comparing the performance under normal conditions we concluded that the clustered schemes still perform adequately, reducing the demand of customizing each inverter control. Future work will focus on the experimental verification of the proposed scheme both in terms of active power curtailment and reactive power

characteristic curves, as well as on more complicated models with multiple features based on machine learning tools.

REFERENCES

- [1] N. Hatzigiorgiou, O. Vlachokyriakou, T. Van Cutsem, J. Milanović, P. Pourbeik, C. Vournas, M. Hong, R. Ramos, J. Boemer, P. Aristidou, V. Singhvi, J. dos Santos, and L. Colombari, "IEEE PES Task Force on Contribution to Bulk System Control and Stability by Distributed Energy Resources connected at Distribution Network," Tech. Rep., 2017.
- [2] P. Fortenbacher, M. Zellner, and G. Andersson, "Optimal sizing and placement of distributed storage in low voltage networks," in *Proc. of the 19th PSCC, Genova*, Jun 2016.
- [3] S. Karagiannopoulos, P. Aristidou, A. Ulbig, S. Koch, and G. Hug, "Optimal planning of distribution grids considering active power curtailment and reactive power control," *IEEE PES General Meeting*, 2016.
- [4] S. Bolognani and S. Zampieri, "A distributed control strategy for reactive power compensation in smart microgrids," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2818–2833, Nov. 2013.
- [5] P. Kotsampopoulos, N. Hatzigiorgiou, B. Bletterie, and G. Lauss, "Review, analysis and recommendations on recent guidelines for the provision of ancillary services by Distributed Generation," in *IEEE International Workshop on Intelligent Energy Systems, IWIES*, 2013.
- [6] IEEE 1547, "IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems," Jul 2003.
- [7] S. Karagiannopoulos, P. Aristidou, and G. Hug, "Hybrid approach for planning and operating active distribution grids," *IET Generation, Transmission & Distribution*, pp. 685–695, Feb 2017.
- [8] F. Bellizio, S. Karagiannopoulos, P. Aristidou, and G. Hug, "Optimized local control schemes for active distribution grids using machine learning techniques," in *IEEE PES General Meeting, Portland*, June 2018.
- [9] S. Karagiannopoulos, P. Aristidou, and G. Hug, "Data-driven Local Control Design for Active Distribution Grids using off-line Optimal Power Flow and Machine Learning Techniques," *accepted in IEEE Transactions on Smart Grid*, 2019. [Online]. Available: <http://arxiv.org/abs/1808.01009>
- [10] R. Dobbe, O. Sondermeijer, D. Fridovich-Keil, D. Arnold, D. Callaway, and C. Tomlin, "Data-Driven Decentralized Optimal Power Flow," 2018. [Online]. Available: <http://arxiv.org/abs/1806.06790>
- [11] S. Karagiannopoulos, R. Dobbe, P. Aristidou, D. Callaway, and G. Hug, "Data-driven Control Design Schemes in Active Distribution Grids : Capabilities and Challenges," in *2019 IEEE PES PowerTech*, 2019.
- [12] O. Sondermeijer, R. Dobbe, D. Arnold, and C. Tomlin, "Regression-based Inverter Control for Decentralized Optimal Power Flow and Voltage Regulation," *IEEE PES General Meeting*, 2016.
- [13] G. Valverde, T. Zufferey, S. Karagiannopoulos, and G. Hug, "Estimation of Voltage Sensitivities to Power Injections using Smart Meter Data," in *2018 IEEE International Energy Conference (ENERGYCON)*. IEEE, 2018.
- [14] G. Chicco, R. Napoli, and F. Piglion, "Comparisons among clustering techniques for electricity customer classification," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 933–940, 2006.
- [15] T. Zufferey, A. Ulbig, S. Koch, and G. Hug, "Unsupervised Learning Methods for Power System Data Analysis," in *Big Data Application in Power Systems*. Elsevier Inc., 2017, pp. 107–124.
- [16] G. Le Ray, P. Pinson, and E. M. Larsen, "Data-driven demand response characterization and quantification," in *2017 IEEE Manchester PowerTech*. IEEE, 2017, pp. 1–6.
- [17] R. J. Sánchez-García, M. Fennelly, S. Norris, N. Wright, G. Niblo, J. Brodzki, and J. W. Bialek, "Hierarchical spectral clustering of power grids," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2229–2237, 2014.
- [18] A. Z. De Souza, C. A. Canizares, and V. H. Quintana, "New techniques to speed up voltage collapse computations using tangent vectors," *IEEE Transactions on Power systems*, vol. 12, no. 3, pp. 1380–1387, 1997.
- [19] F. Olivier, A. Sutera, P. Geurts, R. Fonteneau, and D. Ernst, "Phase identification of smart meters by clustering voltage measurements," in *2018 Power Systems Computation Conference (PSCC)*. IEEE, 2018.
- [20] C. Bock, "Forecasting energy demand by clustering smart metering time series," in *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Springer, 2018, pp. 431–442.
- [21] C. Feng, M. Cui, B.-M. Hodge, S. Lu, H. Hamann, and J. Zhang, "Unsupervised clustering-based short-term solar forecasting," *IEEE Transactions on Sustainable Energy*, 2018.

- [22] M. Pérez-Ortiz, S. Jiménez-Fernández, P. Gutiérrez, E. Alexandre, C. Hervás-Martínez, and S. Salcedo-Sanz, "A review of classification problems and algorithms in renewable energy applications," *Energies*, vol. 9, no. 8, p. 607, 2016.
- [23] SMA Solar Technology AG, *Sunny Boy 3000TL / 3600TL / 4000TL / 5000TL With Reactive Power Control*, SMA, 2015, <http://files.sma.de/dl/15330/SB30-50TL-21-BE-en-11.pdf>.
- [24] M. Geidl, B. Arnoux, T. Plaisted, and S. Dufour, "A fully operational virtual energy storage network providing flexibility for the power system," in *Proceedings of the 12th IEA Heat Pump Conference, Rotterdam, The Netherlands*, 2017, pp. 15–18.
- [25] G. Valverde, D. Shchetinin, and G. Hug-Glanzmann, "Coordination of distributed reactive power sources for voltage support of transmission networks," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1544–1553, 2019.
- [26] S. Karagiannopoulos, A. Vasilakis, P. Kotsampopoulos, N. Hatzia-ryriou, P. Aristidou, and G. Hug, "Experimental Verification of Self-Adapting Data-driven Controllers in Active Distribution Grids," in (submitted) *Proc. of the 21st PSCC, Porto*, June 2020.
- [27] P. Siano, C. Cecati, H. Yu, and J. Kolbusz, "Real time operation of smart grids via fcn networks and optimal power flow," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 4, pp. 944–952, 2012.
- [28] S. Karagiannopoulos, P. Aristidou, and G. Hug, "Co-optimisation of Planning and Operation for Active Distribution Grids," in *Proc. of the 12th IEEE PES PowerTech, Manchester*, Jun 2017.
- [29] —, "A Centralised Control Method for Tackling Unbalances in Active Distribution Grids," in *Proc. of the 20th PSCC, Dublin*, June 2018.
- [30] W. H. Kersting, *Distribution System Modeling and Analysis*, 2002.
- [31] V. M. R. Muggeo, "Estimating regression models with unknown breakpoints," *Statistics in Medicine*, vol. 22, no. 19, pp. 3055–3071, Oct 2003.
- [32] VDE-AR-N 4105, "Power generation systems connected to the LV distribution network." FNN, Tech. Rep., 2011.
- [33] A. Sardá-Espinosa, "Comparing time-series clustering algorithms in R using the dtwclust package," *R package vignette*, vol. 12, p. 41, 2017. [Online]. Available: <https://cran.r-project.org/web/packages/dtwclust/vignettes/dtwclust.pdf>
- [34] S. Aghabozorgi, A. Seyed Shirshorshidi, and T. Ying Wah, "Time-series clustering - A decade review," *Information Systems*, vol. 53, pp. 16–38, 2015.
- [35] R Core Team, "R: A language and environment for statistical computing," 2013. [Online]. Available: <http://www.R-project.org/>
- [36] P. Montero, "TSclust : An R Package for Time Series Clustering," *Journal Of Statistical Software*, vol. 62, 2014.
- [37] K. Strunz, E. Abbasi, C. Abbey, C. Andrieu, F. Gao, T. Gaunt, A. Gole, N. Hatzia-ryriou, and R. Iravani, "Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources," *CIGRE, Task Force C6.04*, no. 273, pp. 4–6, 4 2014.
- [38] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in matlab," in *Proc. of the CACSD Conference, Taiwan*, 2004.
- [39] I. Gurobi Optimization, "Gurobi optimizer reference manual," 2016. [Online]. Available: <http://www.gurobi.com>
- [40] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.
- [41] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE transactions on pattern analysis and machine intelligence*, no. 2, pp. 224–227, 1979.
- [42] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.



Stavros Karagiannopoulos (S'15-M'19) was born in Thessaloniki, Greece. He received a Diploma in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece, in 2010, and a M.Sc. degree in Energy Science and Technology from the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, in 2013. After his M.Sc., he worked at ABB Corporate Research Center in Switzerland until 2015. In 2019 he received the Ph.D. degree from the Swiss Federal Institute of Technology (ETH). Since August 2019, he is working as a postdoctoral researcher with Power Systems Laboratory at ETH Zurich, and Laboratory for Information and Decision Systems at Massachusetts Institute of Technology (MIT). His main research focuses on data-driven and optimization-based methods for active distribution grids.



Gustavo Valverde (S'08-M'12-SM16) received a B.Sc. degree in electrical engineering from the University of Costa Rica, in 2005 and the M.Sc. and Ph.D. degrees in electrical power systems from the University of Manchester in 2008 and 2012, respectively. He worked as a postdoctoral researcher at the University of Liege, Belgium, in 2012-2013 and ETH Zurich, Switzerland, in 2017-2018. Currently, he is Professor at the School of Electrical Engineering at the University of Costa Rica. His research interests include voltage control and stability, smart grids, and integration of renewables.



Petros Aristidou (S'10-M'15-SM'20) received a Diploma in Electrical and Computer Engineering from the National Technical University of Athens (Greece) in 2010, and a Ph.D. in Engineering Sciences from the University of Liege (Belgium) in 2015. After his Ph.D., he took a Postdoctoral Researcher position at the Power Systems Laboratory at ETH Zurich (Switzerland) for one year. Between 2016-2019, he was a Lecturer (Assist. Prof.) at the University of Leeds (UK) working on Smart Energy Systems. Since January 2020, he is a Lecturer in Sustainable Power Systems at the Cyprus University of Technology. His expertise is on power system dynamics and control and he has participated in several working groups looking into the challenges of low-inertia systems.



Gabriela Hug (S'05-M'08-SM'14) was born in Baden, Switzerland. She received the M.Sc. degree in electrical engineering in 2004 and the Ph.D. degree in 2008, both from the Swiss Federal Institute of Technology, Zurich, Switzerland. After the Ph.D. degree, she worked with the Special Studies Group of Hydro One, Toronto, ON, Canada, and from 2009 to 2015, she was an Assistant Professor with Carnegie Mellon University, Pittsburgh, PA, USA. She is currently an Associate Professor with the Power Systems Laboratory, ETH Zurich, Zurich, Switzerland. Her research is dedicated to control and optimization of electric power systems.