Learning Reactive Power Control Polices in Distribution Networks Using Conditional Value-at-Risk and Artificial Neural Networks

Krishna Sandeep Ayyagari, Reynaldo Gonzalez, Yufang Jin, Miltiadis Alamaniotis, Sara Ahmed, and Nikolaos Gatsis

***Abstract*—Scalable coordination of photovoltaic (PV) inverters, considering the uncertainty in PV and load in distribution networks (DNs), is challenging due to the lack of real-time communications. Decentralized PV inverter setpoints can be achieved to address this issue by capitalizing on the abundance of data from smart utility meters and the scalable architecture of artificial neural networks (ANNs). To this end, we first use an offline, centralized data-driven conservative convex approximation of chance-constrained optimal power flow (CVaR-OPF) in which conditional value-at-risk (CVaR) is used to compute reactive power setpoints of PV inverter, taking into account PV and load uncertainties in DNs. Following that, an artificial neural network (ANN) controller is trained for each PV inverter to emulate the optimal behavior of the centralized control setpoints of PV inverter in a decentralized fashion. Additionally, the voltage regulation performance of the developed ANN controllers is compared with other decentralized designs (local controllers) developed using model-based learning (regression-based controller), optimization (affine feedback controller), and case-based learning (mapping) approaches. Numerical tests using real-world feeders corroborate the effectiveness of ANN controllers in voltage regulation and loss minimization.**

***Index Terms*—Chance constraint, decentralized control, distributed energy resource (DER), data-driven control, neural network, voltage regulation.**

I. Introduction

In recent years, distributed energy resources (DERs) have complicated distribution network (DN) operations. Maintaining nodal voltages within operating tolerances is particularly difficult given the uncertain and intermittent nature of DERs. As a result, step-voltage regulators and shunt capacitors must work harder to maintain voltages in DNs [1]. This consequence degrades their operating mechanism, reduces device lifetime, and can cause power quality, stability, and reliability issues. In order to overcome this challenge, photovoltaic (PV) inverters are allowed to operate at a non-unity power factor to provide reactive power support for voltage regulation. Earlier efforts have focused on developing local control strategies such as volt-var and watt-var curves. These strategies entail each PV inverter adjusting its reactive power output based on local measurements [2], [3]. However, these strategies require extensive tuning to find an appropriate curve for every PV inverter. Hence, they may become impractical in real time, where hundreds of PV inverters could be installed over a DN. Therefore, it is challenging to design a scalable framework to compute reactive power setpoints of PV inverters under limited real-time communications in DNs.

Reactive power setpoints of PV inverters can be computed from optimal power flow (OPF) problems in DNs. The non-convex nature of OPF renders the optimization problem difficult to solve. With recent theoretical advancements in optimization, different convex relaxations have been proposed [4]. Additionally, there exists a related stream of literature on different variants of OPF techniques (e.g., robust and stochastic), which aim at developing PV inverter controls in a centralized, decentralized, or distributed framework to compute the optimal setpoints in real time (e.g., [1], [5]-[12]).

Centralized control strategies [1] generally yield optimal operating costs, although they require extensive monitoring and communication infrastructure for system-wide optimal operation. Decentralized schemes, on the other hand, require no communication and only use local information to modify the DER behavior [5]-[10]. Distributed approaches in [11], [12] use limited communication between neighboring DERs to achieve close-to-optimal operation. However, such approaches may still suffer from communication delays and errors.

As a part of the recent transition to the smart grid, there is an abundance of readily available historical data from utility smart meters [13]. This has led to an increase in research in data-driven approaches for OPF using machine learning techniques (e.g., [14] for a survey). Particularly, [15]-[20] propose data-driven approaches for voltage regulation in DNs. Reference [16] proposes multiple linear regression

Manuscript received: July 30, 2022; revised: October 24, 2022; accepted: December 1, 2022. Date of CrossCheck: December 1, 2022. Date of online publication: January 26, 2022.

This work was supported by the National Science Foundation (No. ECCS-1847125, No. 2115427).

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

K. S. Ayyagari (corresponding author), R. Gonzalez, Y. Jin, M. Alamaniotis, S. Ahmed, and N. Gatsis are with the Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, USA (e-mail: KrishnaSandeep. Ayyagari@utsa. edu; Reynaldo. Gonzalez@utsa. edu; Yufang. Jin@utsa. edu; Miltos. Alamaniotis@utsa. edu; Sara. Ahmed@utsa. edu; Nikolaos. Gatsis@utsa.edu).

DOI: 10.35833/MPCE.2022.000477

models to find a local control policy that maps the local historical data of each PV inverter to the optimal reactive power setpoints computed by using an OPF. However, the previous reference uses a deterministic OPF to compute the reactive power setpoints, while neglecting the uncertainty in user load and PV generation. Reference [17] introduces support vector machines to design volt-var curves utilizing an offline centralized algorithm based on chance-constrained OPF while considering only PV uncertainty. As it is using chance-constrained optimization, an approach is needed to estimate the tightening, which represents the uncertainty margins. In the previous reference, the authors propose using Monte-Carlo simulations to implement the tightening. Another approach in [18] uses kernels to learn nonlinear PV inverter control policies and calculate real-time reactive power injections based on a linearized OPF problem. However, determining the best kernels can be a challenging task.

Notwithstanding the increasing availability of data and machine learning approaches that could be leveraged to map local historical data to optimal PV inverter setpoints, it remains a difficult task to take PV and load uncertainty into account in OPF for any data-driven learning design. For instance, an effective approach to mitigate DER and load uncertainty is to enforce probabilistic specifications for violations of voltage and PV reactive power constraints, leading to chance-constrained (CC) OPF formulations. The CC-OPF is nonconvex and challenging to solve. In order to bypass the nonconvexity, the Gaussianity assumption has been traditionally invoked to model the uncertainty distribution (e. g., [8]), which is not usually valid. Convex surrogates can also replace the chance constraints, such as the conditional-value-at-risk (CVaR) or distributionally robust formulations (e. g., [19], [21]-[23]). The advantage of CVaR-OPF is that it is data-driven and distribution agnostic. Nevertheless, data-driven machine learning approaches for PV reactive power control are not explored in [21]-[23]. Therefore, leveraging the CVaR-OPF approach to compute the optimal PV reactive power setpoints and then using machine learning techniques to learn the mapping is an interesting research direction to capture uncertainty in historical data and learn the probabilistic guarantees associated with the voltage regulation constraint. Furthermore, using artificial neural networks (ANNs) to learn the mapping is advantageous as ANNs can accommodate any degree of nonlinearity and generally constitute a model-free approach. Next, we list the literature pertaining to the application of ANNs in DNs.

In the past, research has been published concerning the application of ANNs to solve various DN problems [24]-[26]. One such application is the coordination of distribution system assets such as tap changers, shunt capacitors, and step-voltage regulators [24]. A single neural network is trained from deterministic offline optimization approaches to infer active and reactive power setpoints of the DERs in a centralized manner, replacing the role of distribution system operator (DSO) in real time [25]. A large-scale communication infrastructure required to communicate the optimal setpoints predicted from the centralized ANN to other DERs in real time is assumed [25]. In [26], ANNs are used to estimate nodal voltages in DNs using real-time smart meter data.

Most recently, [27] integrates deep neural networks (DNNs) for PV inverter control policy directly into the CVaR-OPF by considering load and PV uncertainty. Specifically, the DNN training is directly incorporated into the CVaR-OPF problem where the DNN weights are trained using back-propagation and upon computing gradients of losses and voltages with respect to inverter reactive power injections. Gradient-free variants are also explored in order to optimize the DNN weights. Albeit DNN specific parameters can be learned directly by solving the CVaR-OPF, incorporating additional network objectives or constraints remains difficult. Reference [28] uses deterministic optimization for data-driven learning. Also, the CVaR objective along with mean-square error (MSE) is used during ANN training to infer the reactive power setpoints. Furthermore, the mini-batch gradient descent algorithm is developed by carefully selecting the mini-batches to speed up the training process of the ANNs. Therefore, evaluating the performance of ANNs for data-driven decentralized voltage regulation using the optimal reactive power setpoints computed from CC-OPF is thus required, which was carried out in the conference precursor of our work [19]. Specifically, in [19], the ANNs are trained using optimal reactive power setpoints obtained by solving the CC-OPF problem without incorporating the training process directly into the optimization, and the capability of voltage regulation and loss minimization of the developed ANN model is compared with that of a tree-based regression model for a smaller network (IEEE 13-node) and for a longer timescale (1-hour).

The contributions of this paper are listed as follows.

1) A methodology for using ANNs to learn the mapping from load and PV generation uncertainties to inverter reactive power setpoints from data optimized by the CVaR-OPF is developed. The CVaR-OPF formulation and the ANN structure with different activation functions and training process are presented in detail.

2) We extend the previous work in [19] to accommodate larger networks with reasonable number of PVs and showcase the benefits of a scalable implementation of the data-driven decentralized learning approach. Results are showcased in the Arizona SB 129-node feeder [29].

3) The decentralized controllers are implemented using a faster timescale of 15 min and are tested for under- and over-voltage test cases.

4) It is investigated whether the trained ANNs generalize the uncertainty in data sufficiently over a longer period by respecting the probabilistic specification of voltage constraints.

5) The developed ANN controllers are compared in terms of voltage regulation and thermal loss minimization with the following data-driven approaches: regression-based controllers [16], case-based learning [30], and optimized affine feedback schemes [23]. Specifically, we develop regression-based controllers by adopting the CVaR-OPF problem, which is not performed in [16]. The end result is to present the trade-offs of various decentralized designs concerning voltage regulation and thermal loss minimization while com-

plying with the probabilistic specification on the voltages and PV inverter constraints. The ANNs perform remarkably well in terms of probabilistic voltage regulation and thermal loss minimization compared with regression-based, case-based, and optimization-based affine control schemes.

The remainder of this paper is organized as follows. Section II presents the system model and the data-driven centralized offline stochastic OPF. The voltage regulation problem with generic chance constraints and their data-driven approximations used to obtain optimal DER setpoints are also formulated. Data-driven local designs for reactive power control using linear and nonlinear policies are the theme of Section III. Section IV details the numerical tests, including the network setup and data collection process. Thorough comparisons are presented between the performance of the developed ANN controller and other designs in terms of voltage regulation and thermal loss minimization. Finally, conclusions are drawn in Section V.

II. System Model and Data-driven Centralized Offline Stochastic OPF

The network and resource model adopted in this paper are detailed first, followed by the methodology to account for the uncertainty in user load and PV generation using CVaR optimization [19], [23].

A. Power Distribution System Model

Consider a single-feeder radial distribution network modeled by a tree graph with $N+1$ buses (nodes) and lines (edges) connecting these buses. Let $\mathcal{N}_0:=\{0, 1, ..., N\}$ denote the set of all buses and $\mathcal{L}:=\{1, 2, ..., L\}$ denote the set of lines. The substation is indexed by $n=0$. All nodes except the substation are included in the set $\mathcal{N}:=\{1, 2, ..., N\}$ and represent user nodes. Let v_n denote the squared voltage magnitude at bus $n\in\mathcal{N}_0$, where v_0 is fixed, and let $\boldsymbol{v}$ collect all nodal voltages for $n\in\mathcal{N}$. Let $s_n=p_n+\mathrm{j}q_n$ denote the complex power injected to bus n. For each line $n\in\mathcal{L}$, $z_n=r_n+\mathrm{j}x_n$ denotes its impedance, and $S_n=P_n+\mathrm{j}Q_n$ is the complex power flow to the bus n. Also, let b_n^{sh} be the reactive power injected at bus n (e.g., due to shunt capacitors) at nominal voltage of 1.0 p.u.. We collect all nodal quantities into vectors $\boldsymbol{p}$, $\boldsymbol{q}$, $\boldsymbol{b}^{\mathrm{sh}}$, and $\boldsymbol{v}$, and correspondingly, $\boldsymbol{r}$, $\boldsymbol{x}$, $\boldsymbol{P}$, and $\boldsymbol{Q}$ for lines. Let $\boldsymbol{z}=\boldsymbol{r}+\mathrm{j}\boldsymbol{x}$, and $\boldsymbol{S}=\boldsymbol{P}+\mathrm{j}\boldsymbol{Q}$ denote the respective complex vectors. The relationship between voltage magnitudes, power injections, and line power flows is captured by the *LinDistFlow* model [31] in p.u..

$$\boldsymbol{P}=-\boldsymbol{F}^{\mathrm{T}}\boldsymbol{p} \tag{1}$$

$$\boldsymbol{Q}=-\boldsymbol{F}^{\mathrm{T}}\boldsymbol{q}-\boldsymbol{F}^{\mathrm{T}}\cdot\mathrm{diag}(\boldsymbol{b}^{\mathrm{sh}})\cdot\boldsymbol{v} \tag{2}$$

$$\boldsymbol{A}\boldsymbol{v}=2\mathrm{Re}(\boldsymbol{Z}^{*}(\boldsymbol{P}+\mathrm{j}\boldsymbol{Q}))-\boldsymbol{a}_0 v_0 \tag{3}$$

where $\boldsymbol{A}\in\mathbb{R}^{L\times N}$ results from removing the first column of the network edge-to-node incidence matrix $\tilde{\boldsymbol{A}}=[\boldsymbol{a}_0, \boldsymbol{A}]\in\mathbb{R}^{L\times(N+1)}$; we also have $\boldsymbol{F}=-\boldsymbol{A}^{-1}$ with the property that $\boldsymbol{F}\boldsymbol{a}_0=\boldsymbol{1}_N$ and $\boldsymbol{Z}=\mathrm{diag}(\boldsymbol{z})$ [31]; v_0 is the squared voltage magnitude at the slack bus; and $\boldsymbol{a}_0$ is the first column in $\tilde{\boldsymbol{A}}$ [31]. For a given $N\times 1$ vector $\boldsymbol{z}$, $\mathrm{diag}(\boldsymbol{z})$ returns an $N\times N$ matrix with the elements of $\boldsymbol{z}$ on its diagonal. Further, $\boldsymbol{I}_N$ denotes an $N\times N$ identity matrix; and $\boldsymbol{0}_N$ and $\boldsymbol{1}_N$ are the N-dimensional vectors with all zeroes and ones, respectively.

Substituting (1) and (2) into (3) and premultiplying (3) with $-\boldsymbol{F}$ yields:

$$\boldsymbol{v}=\boldsymbol{K}_{\mathrm{v}}(\boldsymbol{R}\boldsymbol{p}+\boldsymbol{X}\boldsymbol{q}+\boldsymbol{1}_N v_0) \tag{4}$$

where $\boldsymbol{R}:=2\boldsymbol{F}\cdot\mathrm{diag}(\boldsymbol{r})\cdot\boldsymbol{F}^{\mathrm{T}}$; $\boldsymbol{X}:=2\boldsymbol{F}\cdot\mathrm{diag}(\boldsymbol{x})\cdot\boldsymbol{F}^{\mathrm{T}}$; and $\boldsymbol{K}_{\mathrm{v}}:=(\boldsymbol{I}_N-\tilde{\boldsymbol{X}})^{-1}$, $\tilde{\boldsymbol{X}}:=\boldsymbol{X}\cdot\mathrm{diag}(\boldsymbol{b}^{\mathrm{sh}})$. It is assumed that the network parameters in $\boldsymbol{X}$ and $\boldsymbol{b}^{\mathrm{sh}}$ render $(\boldsymbol{I}_N-\tilde{\boldsymbol{X}})$ invertible. The model in (4) approximates squared voltage magnitude as affine functions of power injections $\boldsymbol{p}$ and $\boldsymbol{q}$, and generalizes [31] to include shunt capacitors.

B. Generation and Load Model

The network has N_{pv} distributed PV generator units whose connection to the buses is described by the PV-to-node incidence matrix $\boldsymbol{\Gamma}\in\mathbb{R}^{N\times N_{\mathrm{pv}}}$. Due to solar intermittency, the real power p_k^{pv} of PV unit $k=1, 2, \ldots, N_{\mathrm{pv}}$ can be modeled as a random variable, while its reactive power injection q_k^{pv} can be actively controlled. Further, we collect the solar generation and reactive power injections from all PV buses in vectors $\boldsymbol{p}^{\mathrm{pv}}\in\mathbb{R}^{N_{\mathrm{pv}}}$ and $\boldsymbol{q}^{\mathrm{pv}}\in\mathbb{R}^{N_{\mathrm{pv}}}$, respectively. If $S_{k,\max}^{\mathrm{pv}}$ is the apparent power capacity for inverter k, the reactive power injections respect the capacity constraints:

$$\left|q_k^{\mathrm{pv}}\right|\le q_{k,\max}^{\mathrm{pv}}:=\sqrt{(S_{k,\max}^{\mathrm{pv}})^2-(p_k^{\mathrm{pv}})^2} \tag{5}$$

The DN also includes N_{c} constant-power loads, whose connection to network buses is given by the load-to-node incidence matrix $\boldsymbol{\Psi}\in\mathbb{R}^{N\times N_{\mathrm{c}}}$. The load active and reactive power consumptions p_k^{c} and q_k^{c} $(k=1, 2, \ldots, N_{\mathrm{c}})$ are modeled as random variables. The nodal active and reactive power consumptions are collected in vectors $\boldsymbol{p}^{\mathrm{c}}=[p_1^{\mathrm{c}}, p_2^{\mathrm{c}}, \ldots, p_{N_{\mathrm{c}}}^{\mathrm{c}}]^{\mathrm{T}}\in\mathbb{R}^{N_{\mathrm{c}}}$ and $\boldsymbol{q}^{\mathrm{c}}=[q_1^{\mathrm{c}}, q_2^{\mathrm{c}}, \ldots, q_{N_{\mathrm{c}}}^{\mathrm{c}}]^{\mathrm{T}}\in\mathbb{R}^{N_{\mathrm{c}}}$, respectively. Vector $\boldsymbol{w}=[(\boldsymbol{p}^{\mathrm{c}})^{\mathrm{T}}, (\boldsymbol{q}^{\mathrm{c}})^{\mathrm{T}}, (\boldsymbol{p}^{\mathrm{pv}})^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{N\times(2N_{\mathrm{c}}+N_{\mathrm{pv}})}$ collects all system disturbances, which are uncontrollable. Finally, we express the net active and reactive power injections $\boldsymbol{p}$ and $\boldsymbol{q}$ in terms of controlled input $\boldsymbol{u}$ and disturbance $\boldsymbol{w}$ as follows:

$$\begin{cases}\boldsymbol{p}=\boldsymbol{B}_{\mathrm{w}}\boldsymbol{w}\\ \boldsymbol{q}=\boldsymbol{\Gamma}\boldsymbol{u}+\boldsymbol{K}_{\mathrm{w}}\boldsymbol{w}\end{cases} \tag{6}$$

where $\boldsymbol{B}_{\mathrm{w}}=[-\boldsymbol{\Psi}, \boldsymbol{0}_{N\times N_{\mathrm{c}}}, \boldsymbol{\Gamma}]\in\mathbb{R}^{N\times(2N_{\mathrm{c}}+N_{\mathrm{pv}})}$; $\boldsymbol{u}=[q_1^{\mathrm{pv}}, q_2^{\mathrm{pv}}, \ldots, q_{N_{\mathrm{pv}}}^{\mathrm{pv}}]\in\mathbb{R}^{N_{\mathrm{pv}}}$; and $\boldsymbol{K}_{\mathrm{w}}=[\boldsymbol{0}_{N\times N_{\mathrm{c}}}, -\boldsymbol{\Psi}, \boldsymbol{0}_{N\times N_{\mathrm{c}}}]\in\mathbb{R}^{N\times(2N_{\mathrm{c}}+N_{\mathrm{pv}})}$. Upon substituting (6) into (4), it can be observed that the nodal voltages are reformulated as linear functions of $\boldsymbol{u}$ and $\boldsymbol{w}$:

$$\boldsymbol{v}(\boldsymbol{u}, \boldsymbol{w})=\boldsymbol{D}\boldsymbol{u}+\boldsymbol{E}\boldsymbol{w}+\tilde{\boldsymbol{v}}_0 \tag{7}$$

where $\boldsymbol{D}=\boldsymbol{K}_{\mathrm{v}}\boldsymbol{X}\boldsymbol{\Gamma}\in\mathbb{R}^{N\times N_{\mathrm{pv}}}$; $\boldsymbol{E}=\boldsymbol{K}_{\mathrm{v}}(\boldsymbol{R}\boldsymbol{B}_{\mathrm{w}}+\boldsymbol{X}\boldsymbol{K}_{\mathrm{w}})\in\mathbb{R}^{N\times(2N_{\mathrm{c}}+N_{\mathrm{pv}})}$; and $\tilde{\boldsymbol{v}}_0=\boldsymbol{K}_{\mathrm{v}}\boldsymbol{1}_N v_0\in\mathbb{R}^N$.

C. Objective Function

This paper considers the objective of minimizing the thermal losses on the lines, which are approximated by $\sum_{n=1}^{N} r_n\frac{P_n^2+Q_n^2}{v_0}$. Utilizing the fact that P_n and Q_n can be written as linear functions of p_n and q_n (cf. (1) and (2)), it follows that the losses are quadratic in $\boldsymbol{p}$ and $\boldsymbol{q}$. Furthermore, it can be observed from (6) that $\boldsymbol{p}$ and $\boldsymbol{q}$ are linear functions of $\boldsymbol{u}$ and $\boldsymbol{w}$. Therefore, the thermal losses can be expressed as

quadratic functions of $\boldsymbol{u}$ and $\boldsymbol{w}$ as follows:

$$\sum_{n=1}^{N} r_n \frac{P_n^2+Q_n^2}{v_0} = J(\boldsymbol{u},\boldsymbol{w}) \tag{8}$$

$$J(\boldsymbol{u},\boldsymbol{w}) = \frac{1}{2v_0}[\boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{u}}\boldsymbol{u}+\boldsymbol{w}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{w}}\boldsymbol{w}+\boldsymbol{w}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{wu}}\boldsymbol{u}+\boldsymbol{u}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{uw}}\boldsymbol{w}+\boldsymbol{s}_{\mathrm{u}}^{\mathrm{T}}\boldsymbol{u}+\boldsymbol{s}_{\mathrm{w}}^{\mathrm{T}}\boldsymbol{w}+h] \tag{9}$$

where $\boldsymbol{R}_{\mathrm{u}}$, $\boldsymbol{R}_{\mathrm{w}}$, $\boldsymbol{R}_{\mathrm{uw}}$, and $\boldsymbol{R}_{\mathrm{wu}}$ are the appropriate matrices; $\boldsymbol{s}_{\mathrm{u}}$ and $\boldsymbol{s}_{\mathrm{w}}$ are the appropriate vectors; and h is a scalar.

D. Chance-constrained Voltage Regulation

Equation (7) directly demonstrates that the uncertainty in $\boldsymbol{w}$ will cause random fluctuations in the nodal voltages. Therefore, it is hard to ensure that voltages remain within bounds $v_{i,\min}$ and $v_{i,\max}$ as specified by ANSI Standard C84.1 [32], i.e., it may not be possible to ensure that $\boldsymbol{v}_{\min} \le (\boldsymbol{D}\boldsymbol{u}+\boldsymbol{E}\boldsymbol{w}+\tilde{\boldsymbol{v}}_0) \le \boldsymbol{v}_{\max}$ holds at all times. Instead, we enforce the latter constraint in a probabilistic fashion. In order to cope with the variability in $\boldsymbol{w}$, we describe the voltage regulation problem first with generic chance constraints followed by their data-driven convex approximation. Consider the following optimization problem (P1) subject to (5), (7), (11) and (12).

$$\text{(P1)} \ \min_{\boldsymbol{u},\boldsymbol{v}} \mathbb{E}(J(\boldsymbol{u},\boldsymbol{w})) \tag{10}$$

$$\Pr\{v_i \ge v_{i,\min}\} \ge \alpha_i \tag{11}$$

$$\Pr\{v_i \le v_{i,\max}\} \ge \alpha_i \quad i=1,2,\ldots,N \tag{12}$$

It follows from (9) that the dependence of thermal losses on $\boldsymbol{w}$ renders the objective function random; therefore, the expected value of the losses is minimized. In addition, constraint (5) may be enforced for all $\boldsymbol{w}$ (i.e., with probability 1) or as a chance constraint with probability β_k. The motivation for the latter is to allow for more flexible reactive power policies in the design phase; and the bounds of (5) will be respected in real time.

Unless the uncertainty has a favorable distribution, it is well-known that chance-constrained optimization is generally nonconvex and thus hard. The CVaR presents itself as a suitable risk measure that can be used as a convex surrogate of the chance constraint [33]. Specifically, CVaR can be utilized to shape the tail of a distribution. Consider a function $f(\boldsymbol{u},\boldsymbol{w})$ of the decision vector and the uncertainty, which enters in a probability constraint as $\Pr\{f(\boldsymbol{u},\boldsymbol{w}) \le 0\}$. The CVaR at level α is defined as $CVaR_{\alpha}(f(\boldsymbol{u},\boldsymbol{w})) = \inf_{t\in\mathbb{R}} \{t+(1-\alpha)^{-1}\mathbb{E}(f(\boldsymbol{u},\boldsymbol{w})-t)^{+}\}$, where $(\cdot)^{+} := \max\{\cdot, 0\}$. It can be observerd that a CVaR constraint serves as a conservative surrogate of the chance constraint, in the sense that if $CVaR_{\alpha}(f(\boldsymbol{u},\boldsymbol{w})) \le 0$, then $\Pr\{f(\boldsymbol{u},\boldsymbol{w}) \le 0\} \ge \alpha$ follows [33]. Therefore, CVaR is a conservative surrogate of the chance constraint. In other words, satisfying the CVaR constraint guarantees that the original chance constraint is respected as well. If $f(\boldsymbol{u},\boldsymbol{w})$ is convex in $\boldsymbol{u}$, CVaR offers a convex restriction to the original chance constraint as:

$$\inf_{t\in\mathbb{R}} \{t+(1-\alpha)^{-1}\mathbb{E}(f(\boldsymbol{u},\boldsymbol{w})-t)^{+}\} \le 0 \tag{13}$$

Thus, the chance constraints in the voltage regulation problem (P1) are replaced by the CVaR constraints. Then, the following operations are performed. The auxiliary variable over which the infimum is taken in (13) is included as optimization variable; the max operator in $(\cdot)^{+}$ is removed by the epigraph trick; and the expectation in (13) is replaced by its average sample approximation. To this end, a set of training scenarios $\{\boldsymbol{w}_{n_{\mathrm{s}}}\}_{n_{\mathrm{s}}=1}^{N_{\mathrm{tr}}}$ (realization of the random variable $\boldsymbol{w}$) is assumed to be available, where N_{tr} is the number of training scenarios. For notational simplicity, define $u_{k,n_{\mathrm{s}}}^{\max} = \bar{q}_{k,n_{\mathrm{s}}}^{\mathrm{pv}} = \sqrt{(S_{\max,k}^{\mathrm{pv}})^2-(p_{k,n_{\mathrm{s}}}^{\mathrm{pv}})^2}$ as the scenario-dependent maximum reactive power capacity; and $u_{k,n_{\mathrm{s}}}^{\min} = -\bar{q}_{k,n_{\mathrm{s}}}^{\mathrm{pv}} = -u_{k,n_{\mathrm{s}}}^{\max}$. The resulting CVaR-based data-driven voltage regulation problem (P2) is stated as (14), subject to (7), (15)-(23).

$$\text{(P2)} \ \min_{\substack{\boldsymbol{v},\boldsymbol{t},\boldsymbol{\vartheta},\boldsymbol{\mu},\boldsymbol{\varphi}\\ \boldsymbol{\zeta},\boldsymbol{\rho},\boldsymbol{\delta},\{u_{k,n_{\mathrm{s}}},v_{i,n_{\mathrm{s}}}\}}} \frac{1}{N_{\mathrm{tr}}}\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{tr}}} J(\boldsymbol{u}_{n_{\mathrm{s}}},\boldsymbol{w}_{n_{\mathrm{s}}}) \tag{14}$$

$$t_i + \frac{1}{1-\alpha_i}\frac{1}{N_{\mathrm{tr}}}\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{tr}}} \vartheta_{i,n_{\mathrm{s}}} \le 0 \tag{15}$$

$$-v_{i,n_{\mathrm{s}}} + v_{i,\min} - t_i \le \vartheta_{i,n_{\mathrm{s}}} \tag{16}$$

$$\mu_i + \frac{1}{1-\alpha_i}\frac{1}{N_{\mathrm{tr}}}\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{tr}}} \varphi_{i,n_{\mathrm{s}}} \le 0 \tag{17}$$

$$v_{i,n_{\mathrm{s}}} - v_{i,\max} - \mu_i \le \varphi_{i,n_{\mathrm{s}}} \tag{18}$$

$$\zeta_k + \frac{1}{1-\beta_k}\frac{1}{N_{\mathrm{tr}}}\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{tr}}} \rho_{k,n_{\mathrm{s}}} \le 0 \tag{19}$$

$$-u_{k,n_{\mathrm{s}}} + u_{k,n_{\mathrm{s}}}^{\min} - \zeta_k \le \rho_{k,n_{\mathrm{s}}} \tag{20}$$

$$\delta_k + \frac{1}{1-\beta_k}\frac{1}{N_{\mathrm{tr}}}\sum_{n_{\mathrm{s}}=1}^{N_{\mathrm{tr}}} v_{k,n_{\mathrm{s}}} \le 0 \tag{21}$$

$$u_{k,n_{\mathrm{s}}} - u_{k,n_{\mathrm{s}}}^{\max} - \delta_k \le v_{k,n_{\mathrm{s}}} \tag{22}$$

$$\{\vartheta_i,\varphi_i,\rho_k,v_k\}_{n_{\mathrm{s}}} \ge 0 \quad \forall i=1,2,\ldots,N, \forall n_{\mathrm{s}}=1,2,\ldots,N_{\mathrm{tr}}, \forall k=1,2,\ldots,N_{\mathrm{pv}} \tag{23}$$

Upon solving optimization problem (P2), the optimal control setpoints of the k^{th} PV inverter for scenario n_{s}, i.e., $q_{k,n_{\mathrm{s}}}^{*}$, are projected within the interval $[-\bar{q}_{k,n_{\mathrm{s}}}^{\mathrm{pv}}, \bar{q}_{k,n_{\mathrm{s}}}^{\mathrm{pv}}]$ to respect (5) and are given as $\boldsymbol{q}_{k,n_{\mathrm{s}}}^{\mathrm{proj}}$.

Notice that the reactive power setpoints of the k^{th} PV inverter computed from problem (P2) are adaptive, i.e., the reactive power setpoints $\boldsymbol{q}_{k,n_{\mathrm{s}}}^{\mathrm{proj}}$ correspond to each scenario $\boldsymbol{w}_{n_{\mathrm{s}}}$ without any restriction on the reactive power control policy. In addition, to dispatch the PV reactive power setpoints in real time, the DSO repeatedly solves the optimization problem (P2), which can be taxing computationally and communication-wise if $\boldsymbol{w}_{n_{\mathrm{s}}}$ changes more frequently, and therefore, deploying the control rules in real time becomes obsolete.

To overcome the above-mentioned issues and to expedite the process of adjusting the DER setpoints adaptively based on time-varying $\boldsymbol{w}_{n_{\mathrm{s}}}$, we focus on developing local control policies, where the reactive power of the k^{th} PV inverter is captured by previously optimized inputs/outputs, i.e., $(\boldsymbol{d}_{k,n_{\mathrm{s}}}, \boldsymbol{q}_{k,n_{\mathrm{s}}}^{\mathrm{proj}})$,

where $\boldsymbol{d}_{k,n_s}$ contains only the local historical information of $\boldsymbol{w}_{n_s}$ to be defined shortly. To accomplish this task, we leverage: ① machine learning approaches, precisely ANN and regression-based approaches that learn the nonlinear mapping between $\boldsymbol{d}_{k,n_s}$ and $\boldsymbol{q}^{\text{proj}}_{k,n_s}$; and ② optimization-based approaches, wherein the linear control policy is included during optimization to compute inverter specific coefficients. This is the theme of the ensuing section.

III. Data-driven Local Designs for Reactive Power Control Using Linear and Nonlinear Policies

This section details the various designs for individual local open-loop linear and nonlinear control policies for each inverter. Specifically, we develop an ANN-based controller for each PV inverter, trained using the optimal PV setpoints with their local historical information $(\boldsymbol{d}_{k,n_s}, \boldsymbol{q}^{\text{proj}}_{k,n_s})$. Similarly, we design linear and nonlinear control policies, variations of which have been pursued in the literature to compute PV inverter set points (e.g., [8]-[10], [16]), but not brought together under a unifying umbrella for comparison.

Let us consider the training data set corresponding to the k^{th} PV inverter obtained from training optimization (P2) as $\{\boldsymbol{d}_{k,n_s}, \boldsymbol{q}^{\text{proj}}_{k,n_s}\}_{n_s=1}^{N_{\text{tr}}}$. Vector $\boldsymbol{d}_{k,n_s}$ is the local input to the k^{th} PV inverter whose entries are given by the following base variables: net real power demand $(p^{\text{net}}_{k,n_s} = p^{\text{c}}_{k,n_s} - p^{\text{pv}}_{k,n_s})$, reactive power demand $\boldsymbol{q}^{\text{proj}}_{k,n_s}$, and the maximum reactive power capacity $\bar{q}^{\text{pv}}_{k,n_s}$ given by (5). It should be noted that the voltage v_{k,n_s} which is dynamically coupled to the local control action $\boldsymbol{q}^{\text{proj}}_{k,n_s}$ may also be appended to $\boldsymbol{d}_{k,n_s}$; however, the stability of the resulting controller is difficult to analyze (e.g., [34]). Therefore, the nodal voltage v_{k,n_s} is not an input to the local policies developed in the present work. The goal of machine learning techniques (ANN and regression) used in Designs I-IV is to find a (generally nonlinear) control policy Ω: $f_k(\boldsymbol{d}_{k,n_s}) \rightarrow \boldsymbol{q}^{\text{proj}}_{k,n_s}$ that maps the local information of the k^{th} PV inverter to its projected optimal setpoint using the previously optimized available input-outputs $f_k(\{\boldsymbol{d}_{k,n_s}\}_{n_s=1}^{N_{\text{tr}}}) \rightarrow \{\boldsymbol{q}^{\text{proj}}_{k,n_s}\}_{n_s=1}^{N_{\text{tr}}}$.

Once the control policy has been designed, it can be locally applied in real time given the present net real power demand, reactive power demand, and available reactive power capacity, to determine the reactive power setpoint for each inverter. Specifically, the real and reactive power demands p^{c}_k and q^{c}_k are typically available by a smart meter, while the real power generation p^{pv}_k is determined by the maximum power point tracking of PV generator or similar algorithm. In the present paper, the performance of the local policies is assessed using a set of N_{tst} test (not previously seen) scenarios. The set of local inputs corresponding to the test data is denoted by $\{\boldsymbol{d}'_{k,n_s}\}_{n_s=1}^{N_{\text{tst}}}$.

In this paper, we devise ANN controllers to approximate the mapping Ω. The ANN structure amounts to a two-layer feed-forward network that consists of one hidden layer (HL) and one output layer (OL) for all PV inverters considered in this paper, as shown in Fig. 1. The first layer includes four hidden neurons and with parameters $(\boldsymbol{\Theta}^{\text{HL}}_k, \boldsymbol{\lambda}^{\text{HL}}_k)$ which represent weights and biases corresponding to the inputs of the k^{th} PV inverter, respectively. Similarly, $(\boldsymbol{\theta}^{\text{OL}}_k, \boldsymbol{\lambda}^{\text{OL}}_k)$ represent the output layer parameters. We use a nonlinear activation function ϕ^{HL} for computation at the HL; and for the OL we use a linear activation function ϕ^{OL}. Given training input vector $\boldsymbol{d}_{k,n_s}$, the layer-wise computations are given as follows.

$$\boldsymbol{\psi}^{\text{HL}}_{k,n_s} = \boldsymbol{\phi}^{\text{HL}}_k(\boldsymbol{\Theta}^{\text{HL}}_k \boldsymbol{d}_{k,n_s} + \boldsymbol{\lambda}^{\text{HL}}_k) \tag{24}$$

$$\boldsymbol{\psi}^{\text{OL}}_{k,n_s} = (\boldsymbol{\theta}^{\text{OL}}_k)^{\text{T}} \boldsymbol{\psi}^{\text{HL}}_{k,n_s} + \boldsymbol{\lambda}^{\text{OL}}_k \tag{25}$$

where the vector-valued function $\boldsymbol{\phi}^{\text{HL}}_k$ is applying the nonlinear activation function $\phi(\cdot)_k$ elementwise.

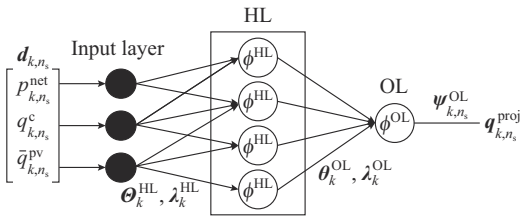

Fig. 1. ANN architecture for k^{th} PV inverter.

For notational simplicity, let $\boldsymbol{\Pi}_k$ collect the trainable parameters $(\boldsymbol{\Theta}, \boldsymbol{\lambda})$ of all layers for the ANN corresponding to the k^{th} PV inverter. In the task of supervised learning, the ANN is trained using back-propagation algorithms based on gradient descent which minimize the training loss defined as:

$$\min_{\boldsymbol{\Pi}_k} \frac{1}{N_{\text{tr}}} \sum_{n_s=1}^{N_{\text{tr}}} \left\| \boldsymbol{q}^{\text{proj}}_{k,n_s} - g_k(\boldsymbol{d}_{k,n_s}; \boldsymbol{\Pi}_k) \right\|_2^2 \tag{26}$$

where $g_k(\boldsymbol{d}_{k,n_s}; \boldsymbol{\Pi}_k)$ represents the composite mapping given by (3). The choice of training loss is task specific and in this study we use the MSE. Upon training, the optimal parameters $\boldsymbol{\Pi}^*_k$ are available, and the trained ANN is used to estimate the reactive power setpoints for the actual test data $\{\boldsymbol{d}'_{k,n_s}\}_{n_s=1}^{N_{\text{tst}}}$, which is expressed as:

$$q^{\text{pv}}_{k,n_s} = g_k(\boldsymbol{d}'_{k,n_s}; \boldsymbol{\Pi}^*_k) \tag{27}$$

1) Design I: ANN-based controller with tangent-sigmoid activation function. In this design, for nonlinear activation function $\phi^{\text{HL}}_k(\cdot)$ of each neuron in the HL, we use the tangent-sigmoid function *tanh*. One advantage of tangent-sigmoid neurons is that the negative inputs are strongly negative and the zero inputs are close to zero, allowing them to have outputs over a wide range of input space. A linear activation function is used for the neuron in the output layer.

2) Design II: ANN-based controller with rectified linear unit (ReLU) activation function. In this design, the transfer function in the HL of the neural network developed in Design I is replaced by the ReLU activation function, which is used by the majority of ANN applications in recent years. This is because of the fact that ReLU is computationally less expensive as it involves less mathematical operations and is easier to implement [35].

The training algorithm for Designs I and II is presented herein. In this paper, the Bayesian regularization algorithm is used as a training algorithm implemented using the command *TrainBr* in MATLAB [36] for Designs I and II to up-

date the weight and bias values of the ANN models. The Bayesian regularization algorithm minimizes the combination of squared errors and weights, and then determines the correct combination so as to produce a network that generalizes well [37]. Since the ANN training starts with random initial weights, different results are obtained by different training algorithm runs. Therefore, every training is rerun until the best solution is obtained in terms of performance [26]. It is noteworthy to mention that, since we consider only three base variables to estimate the optimal setpoints, the sensitivity of the estimation to the number of neurons in the hidden layer stabilizes for four neurons for all the test cases. The training parameters and corresponding values for Designs I and II are listed in Table I for all test cases.

TABLE I
TRAINING PARAMETERS AND CORRESPONDING VALUES FOR DESIGNS I AND II

Parameter	Explanation	Value
Epochs	The maximum number of epochs to train	1000
Goal	Performance goal (MSE)	0
$\min_{grad}$	The minimum improvement from one epoch to the next	10^{-9}
val_{fail}	The maximum validation failures	15
μ_{int}	Marquardt adjustment parameter	10^{-3}
μ_{inc}	Increase factor for μ	10
μ_{dec}	Decrease factor for μ	10^{-2}

3) Design III: regression-based controller with quadratic interactions and Bayesian information criterion (BIC). This design specifically uses regression [16] for deriving local control of individual DERs in a decentralized fashion. Also, this design assumes the quadratic transformations of the base variables which contain additional nonlinear terms including the pairwise products and quadratic terms of base variables. The base variables are selected based on the BIC criterion [16]. In this design, the mapping Ω is approximated by the function $f_k(\{\boldsymbol{d}_{k,n_s}\}_{n_s=1}^{N_{tr}}) = \tau_k^0 + \tau_k^1\{d_{k,n_s}^1\}_{n_s=1}^{N_{tr}} + \ldots + \tau_k^n\{d_{k,n_s}^n\}_{n_s=1}^{N_{tr}}$, where τ_k^0, τ_k^1, …, τ_k^n are the coefficients that are to be evaluated upon minimizing the MSE loss function, i.e., $\frac{1}{N_{tr}}\sum_{n_s=1}^{N_{tr}}\left\| \boldsymbol{q}_{k,n_s}^{proj} - f_k(\boldsymbol{d}_{k,n_s};\boldsymbol{\tau}_k)\right\|_2^2$.

4) Design IV: regression-based controller with linear interactions and sum of squared estimate of errors (SSE). This design also uses regression and replaces the quadratic transformations with linear transformations of the base variables and further uses SSE criterion for model selection. Vector $\boldsymbol{d}_{k,n_s}$ contains an intercept (constant term), linear term of each base variable, and all pairwise products of distinct base variables (no quadratic terms).

The regression-based Designs III and IV are trained using the *stepwiselm* command from the statistics and machine learning toolbox in MATLAB [36]. It is worth emphasizing that the regression-based controllers in [16] have been developed using deterministic OPF. However, the regression-based controllers in this study are developed by adopting the CVaR-based data-driven voltage regulation problem (P2), lending them scenario-based adaptivity in a probabilistic fashion.

5) Design V: Case-based learning approach. In this design, the training information along with target data $\{\boldsymbol{d}_k, \boldsymbol{q}_{k,n_s}^{proj}\}_{n_s=1}^{N_{tr}}$ is stored in a database of past cases. The actual test realization $\{\boldsymbol{d}'_{k,n_s}\}_{n_s=1}^{N_{tst}}$, for which the reactive power setpoints are to be computed, are called the present cases. The present case vector $\boldsymbol{d}'_{k,n_s}$ is compared to all past cases in the database to find the best match in the least Euclidean distance sense. Specifically, the distance is defined as $\left\| \boldsymbol{d}'_{k,n_s} - \boldsymbol{d}_{k,n_s}\right\|_2$. Then the corresponding setpoint $\boldsymbol{q}_{k,n_s}^{proj}$ from the past case with the smallest distance is used as the estimation for the present case. This technique has been implemented to predict building energy consumption in [30], and more generally, the Euclidean distance has been used in various case-based learning applications [38], [39]. This technique would be difficult to implement in a real-time setup due to the time needed to compute distances to all cases in the database and find the best estimate. It may also exhibit a large generalization error.

6) Design VI: affine feedback control policy. In this design, the control policy for the k^{th} PV inverter is restricted to have a linear form $q_k^{pv}(\boldsymbol{w}) = \boldsymbol{m}_k^T\boldsymbol{w} + \tau_k$, where $\boldsymbol{m}_k$ and τ_k are the optimization variables for the k^{th} PV inverter [8]-[10]. The coefficients $\boldsymbol{m}_k$ select the entries of $\boldsymbol{w}$ corresponding to the k^{th} PV inverter (and are set to zero otherwise), yielding a decentralized linear policy. The decentralized linear policy can be optimized by including in the optimization problem (P2) additional constraints $\boldsymbol{u}_{n_s} = \boldsymbol{M}\boldsymbol{w}_{n_s} + \boldsymbol{\tau}$ [23] with appropriate sparse matrix $\boldsymbol{M} \in \mathbb{R}^{N_{pv}\times(2N_c+N_{pv})}$ and vector $\boldsymbol{\tau} \in \mathbb{R}^{N_{pv}}$.

The flow diagram depicting the open-loop local control designs and validation process is shown in Fig. 2. The load consumption and PV generation data are first collected to solve the centralized CVaR-OPF (P2) and compute the optimal PV setpoints for Designs I through V, as well as the optimal coefficients $\boldsymbol{m}_k$, τ_k per inverter for Design VI. The optimal PV setpoints produced by the CVaR-OPF solution are projected back to their feasible regions to respect the device limits and the projected data are used to train Designs I through IV. The reactive power controllers produced by all designs are then used to compute the setpoints for the actual test data. It is worth emphasizing that the test data used for validation are different from the ones used for training. The computed setpoints for the actual test data are projected back to their feasible regions and the resulting reactive power injections are then given as inputs to a nonlinear power flow solver, namely the *Z*-bus method [40], to obtain the nodal voltages across the network. The latter are used to validate the performance of the different designs in terms of voltage regulation and loss improvement. Specifically, the empirical probability of voltage violation is assessed in the ensuing section.

Figure 3 illustrates that the reactive power setpoint computed by the decentralized policy such as the ANN of Design I or Design II is fed into the traditional power controller, current controller, and switching logic module of the inverter.

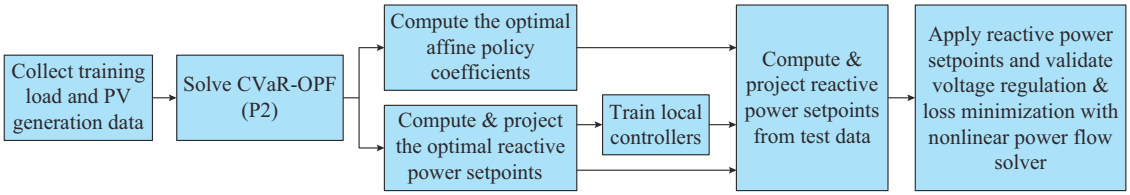

Fig. 2. Flow diagram depicting open-loop local control designs and validation process.

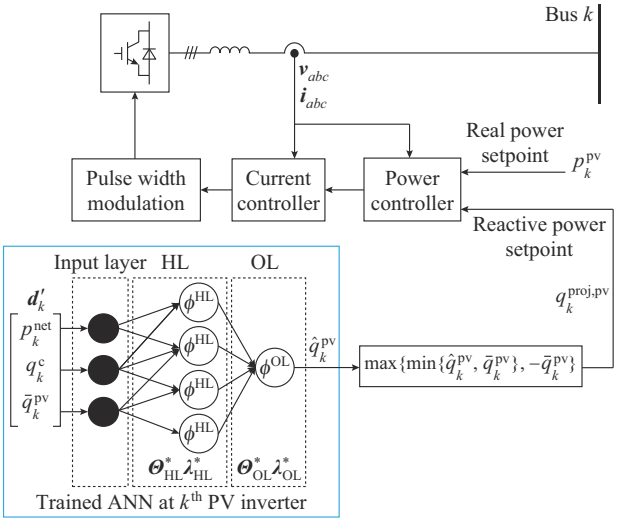

Fig. 3. Implementation of local ANN controller for the k^{th} PV inverter in Designs I and II.

Remark 1: extension to multi-phase DNs. The proposed data-driven local control designs can be extended to multi-phase DNs. The counterpart of *LinDistFlow* in (2) is *LinDist3Flow* approximation upon ignoring losses and other high-order terms [41], which we have recently extended to handle step-voltage regulator tap selection in the OPF [42]. Upon solving the resultant optimization problem, the optimal reactive power setpoints can be used to train the ANNs.

IV. Numerical Tests

A. Network and Test Case Setup

The network used in this paper is the Arizona SB 129-node test feeder, whose line parameters, nominal load values, and PV locations are adopted from [29]. Further, we assume $V_{base} = 4.16$ kV and $S_{base} = 2$ MVA for all the test cases considered in this paper. The voltage limits v_{min} and v_{max} are set to be 0.95 p.u. and 1.05 p.u., respectively, for all test cases.

B. Data Collection

The data for solving the optimization problem (P2) are collected from the homes installed with smart meters located on Pecan Street in Austin, Texas, USA [43]. Historical load consumption and PV generation data are obtained for the month of July 2015. The data are collected in both hourly and minute-based resolution. Since the Pecan Street data include only active power, the reactive power consumptions are generated by $\boldsymbol{q}^c = \boldsymbol{p}^c \tan\phi$ assuming lagging power factor. The data from each home are then aggregated to match the nominal load of the node. Details on data aggregation can be found in [19]. A variety of under- and over-voltage scenarios, control resolutions, and chance constraint specifications are investigated, as described next.

C. Test Cases

1) Test Cases A and B (Under-voltage, Slow Time-scale)

Test cases A and B are to investigate the performance of various designs for the under-voltage scenario in DNs. In test cases A and B, we consider a one-hour resolution of historical data. The optimization is performed based on the data for the first 30 days of the month ($N_{tr} = 720$ scenarios) to generate the training scenarios, and the performances of the developed control designs (Designs I-VI) are evaluated for the last day ($N_{tst} = 24$ scenarios). Further, the maximum PV generation is assumed to be 80% of the nominal consumption $\boldsymbol{p}^c$. The apparent power capacity S_{max}^{pv} is set to be 185% of P_{max}^{pv}. Moreover, a lagging power factor of 0.95 is assumed for all loads. For test case A, the voltage violation probability α is set to be 0.9 and the PV inverter reactive power capacity violation probability β is assumed to be 0.95 in the optimization problem (P2). The difference in test case B is that the probability specifications are tightened in the optimization, which poses a challenge for the decentralized designs not to exceed the desired violation probabilities. Precisely, the voltage violation probability is tightened to 0.95 and the PV inverter capacity violation probability is further tightened to 0.99.

2) Test Cases C and D (Over-voltage, Fast Time-scale)

Test cases C and D are the investigations into over-voltage scenarios in the DNs (SB 129-node modified) while using the 15-min data-point resolution. For the 15-min data-point resolution, the 1-min data-point resolution load consumption and PV generation profiles are first considered. The values are then averaged every 15 min to construct 15-min based profiles. Furthermore, we assume the power factor of 0.99 (lagging) for all loads to compute the reactive power profiles. To create over-voltage scenarios, the original SB 129-node feeder for test cases C and D is modified by adding eight additional PV inverters and 6 shunt capacitors with ratings [1.5, 100, 5, 5, 5, 3]kvar. Voltage violation and PV inverter probability specifications are both set to be 0.95. Furthermore, the maximum PV generation is assumed to be 130% of the actual nominal consumption $\boldsymbol{p}^c$. The apparent power capacity S_{max}^{pv} is set to be 120% of P_{max}^{pv}. Also, nominal loads are scaled down to 10% of their actual values. The optimization for test case C is performed in a similar way as test cases A and B, i.e., the optimization is performed for 30 days of July 2015 ($N_{tr} = 2880$ scenarios) and the performance of the developed control designs is evaluated for the last day

(N_{tst} = 96 scenarios). For test case D, the optimization is performed for the first 24 days (N_{tr} = 2304 scenarios). Then, the control designs are tested on the last seven days of July 2015 (N_{tst} = 672 scenarios). In other words, the difference in test case D is that we test for the last seven days instead of the last day. The objective is to investigate the generalization over a longer period of time.

The optimization problem for test cases A, B, C, and D is programmed in MATLAB invoking CVX [44] with MOSEK Solver. Specifically, test cases A and B are solved using a PC with 32 GB RAM, and test cases C and D are solved using a high-performance computing cluster.

Table II lists the maximum voltage probability violations in percentage for the terminal node (node 129) in all designs and test cases. Voltage violations are shown in bold. Figure 4 depicts a bar graph comparing the results in Table II. More detailed analysis of the results per test case are provided next.

TABLE II
THE MAXIMUM VOLTAGE PROBABILITY VIOLATIONS FOR NODE 129

Test case	The maximum voltage probability violation (%)							Allowed voltage violation (%)
	Design I	Design II	Design III	Design IV	Design V	Design VI	No control	
A	8.33	8.33	**25.00**	**16.67**	**8.33**	4.17	**62.50**	10
B	4.17	**8.33**	**25.00**	**16.67**	**8.33**	0.00	**62.50**	5
C	1.04	1.04	1.04	**6.25**	**9.38**	0.00	**21.88**	5
D	2.83	2.38	2.83	3.13	2.38	0.00	**26.70**	5

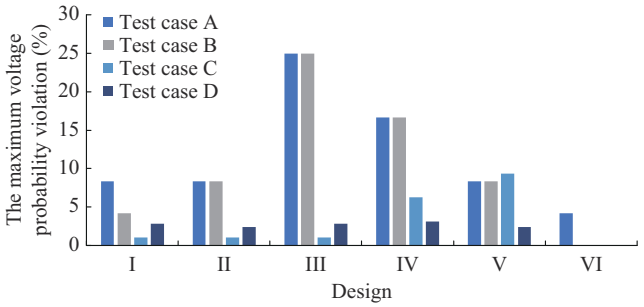

Fig. 4. The maximum voltage probability violations in percentage for node 129.

1) Test case A: it can be observed from Table II that both regression controllers (Designs III and IV) violate the voltage specification, i.e., the voltage probability violations are above 10%. From Fig. 4, it can also be observed that Designs III and IV perform significantly worse for test case A. Both ANN controllers (Designs I and II) and the case-based learning (Design V) pass the voltage specification, and have the same performance outcome of 8.33%. The affine controller (Design IV) performs the best for test case A, with a probability of 4.17%.

2) Test case B: by observing Table II, it can be observed that only two designs passed the voltage specification, i.e., the ANN with tangent-sigmoid (Design I) and the affine controller (Design VI). From Fig. 4, it can be observed again that both regression controllers perform significantly more poorly in comparison to the other designs for test case B. The ANN with tangent-sigmoid (Design I) performs better in test case B compared with test case A, with voltage probability violation of 4.17%. The affine controller (Design VI) performs the best compared with the other designs, for test case B with zero probability of voltage violations.

3) Test case C: for this test, the ANN controllers (Designs I and II) and the regression with quadratic interactions (Design III) all show good performance with 1.04% probability violation, and the affine controller (Design VI) again performs the best with zero probability of voltage violations. From Table II and Fig. 4, it can be observed that the regression without quadratic interactions (Design IV) and the case-based learning (Design V) both exceed the allowed voltage violation specification of 5%, with violation probabilities of 6.25% and 9.38%, respectively.

4) Test case D: for this test case, all designs pass the voltage violation specification of 5%. Design IV performs slightly worse to the other designs at probability violation of 3.13%. ANN with tangent-sigmoid (Design I) and affine controller (Design IV) perform the best with 2.38%.

The empirical cumulative distribution functions (CDFs) for the voltage at node 129, which is the node with the highest probability violation using different designs in various test cases, are depicted in Figs. 5 and 6. The y-axis in Fig. 5 depicts the probability of voltage violation for node 129 being below the minimum voltage limit (p.u.), i.e., $\Pr\{v_{129} \leq 0.95\}$. Similarly, the y-axis in Fig. 6 depicts the probability of voltage violation for node 129 being above the maximum voltage limit, i.e., $\Pr\{v_{129} \geq 1.05\}$. The x-axis, κ, is the voltage violation specification in Figs. 5 and 6.

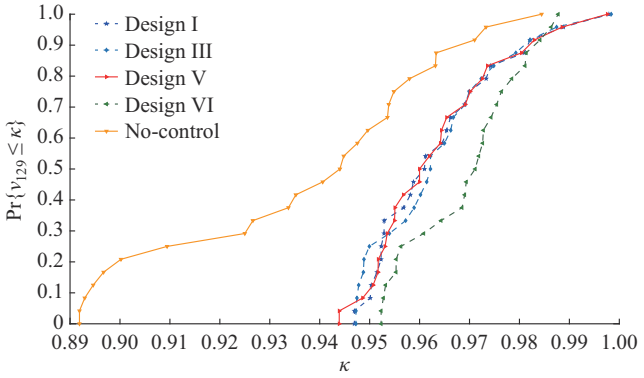

Fig. 5. Empirical CDF of voltage at node 129 for Designs I, III, V, VI, and no-control for test case B using actual test data (July 31) and upon solving nonlinear power flows with Z-bus method.

The network-wide voltage profiles using Designs I, III, V, and VI with respect to no-control for test cases B and C are

also shown in Figs. 7 and 8, respectively. The *y*-axis in Fig. 7 depicts the smallest voltage across the network for each time period in the *x*-axis. The duration over which each curve stays below the horizontal line of 0.95 p. u. corresponds to the voltage probability violation; likewise for Fig. 8 and the over-voltage case.

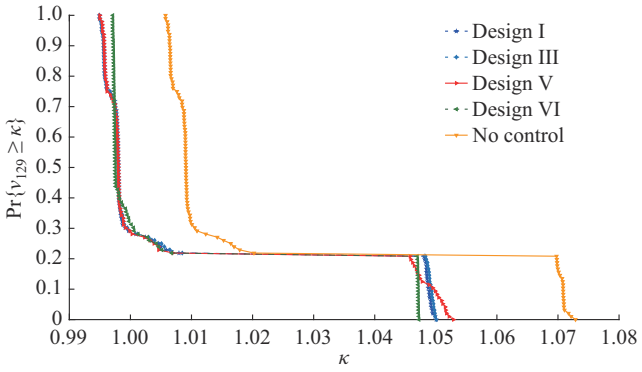

Fig. 6. Empirical CDF of voltage at node 129 for Designs I, III, V, VI, and no-control for test case C using actual test data (July 31) and upon solving nonlinear power flows with *Z*-bus method.

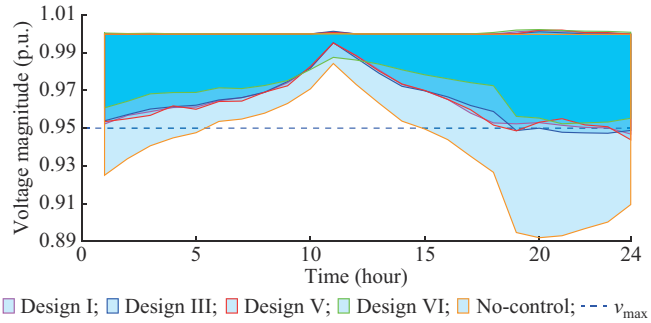

Fig. 7. Network-wide voltage profile for Designs I, III, V, VI, and no-control for test case B using actual test data (July 31) and upon solving nonlinear power flows with *Z*-bus method.

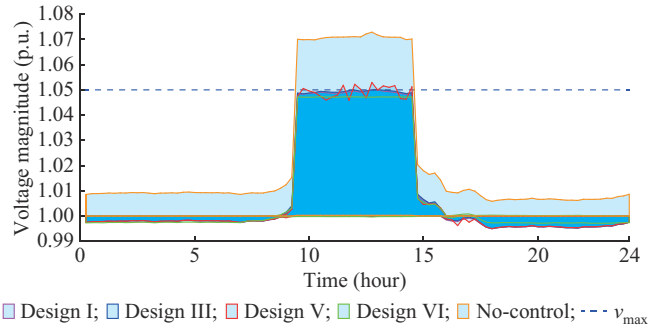

Fig. 8. Network-wide voltage profile for Designs I, III, V, VI, and no-control for test case C using actual test data (July 31) and upon solving nonlinear power flows with *Z*-bus method.

D. Percentage Improvement in Thermal Losses

Table III lists the percent improvement (or reduction) of average thermal losses under the considered control designs versus the case where no reactive power control is applied. The regression-based controllers (Designs III and IV) achieve the best percent improvement compared with the other designs. The ANN controllers (Designs I and II) and case-based learning (Design V) show similar improvement, only slightly less than the regression designs. Interestingly, the affine controller (Design VI) performs the worst in terms of thermal loss improvement, whereby losses have actually increased for test case D.

TABLE III
PERCENTAGE IMPROVEMENT OF AVERAGE THERMAL LOSSES

Test case	Percentage improvement (%)					
	I	II	III	IV	V	VI
A	8.60	8.70	9.96	9.93	8.10	3.88
B	8.21	8.23	9.96	9.92	8.10	0.83
C	13.65	13.84	13.87	14.04	13.50	11.90
D	2.23	2.26	2.51	2.52	1.66	–0.53

Overall, while the regression-based controllers (Designs III and IV) show good thermal loss improvement, they perform poorly for under-voltage scenarios, i. e., test cases A and B. The affine controller (Design VI), on the other hand, outperforms all the other designs concerning voltage regulation. However, this superior performance comes at the cost of the poor improvement in average thermal losses. ANN with ReLU (Design II) results in slightly better improvement to thermal losses compared with ANN with tangent-sigmoid (Design I), but exhibits voltage violation for test case B. The ANN with tangent-sigmoid controller (Design I) provides a good middle ground, with low probability of voltage violations in both under- and over-voltage scenarios, and simultaneously achieves large improvement in terms of thermal losses.

V. CONCLUSION

This paper develops a data-driven control based on ANNs to compute the reactive power setpoints utilizing conservative convex approximations of chance constraints. The controllers can be implemented in a decentralized fashion, without the need for monitoring and communication infrastructure. The developed ANN controllers are compared with regression-based ones, as well as optimization approaches featuring affine feedback rules and a case-based learning approach. ANN controllers turn out to be robust to uncertainties for voltage regulation when compared with other control polices. In future research, we will focus on extending this approach considering more types of possible DERs such as battery energy storage systems and electric vehicles. It is also worth investigating the effect of coordinating different distribution system assets (e.g., step-voltage regulators and shunt capacitors) utilizing a stochastic optimization framework combined with data-driven control. Future research will also look into real-time implementation of the proposed ANN controllers using hardware-in-loop.

REFERENCES

[1] M. Farivar, C. R. Clarke, S. H. Low *et al.*, "Inverter var control for distribution systems with renewables," in *Proceedings of IEEE International Conference on Smart Grid Communications*, Brussels, Belgium, pp. 457-462, Oct. 2011.
[2] J. W. Smith, W. Sunderman, R. Dugan *et al.*, "Smart inverter volt/var

control functions for high penetration of PV on distribution systems," in *Proceedings of IEEE/PES Power System Conference and Exposition*, Phoenix, USA, pp. 1-6, Mar. 2011.
[3] M. Farivar, X. Zho, and L. Che, "Local voltage control in distribution systems: an incremental control algorithm," in *Proceedings of IEEE International Conference on Smart Grid Communications*, Miami, USA, pp. 732-737, Nov. 2015.
[4] D. K. Molzahn and I. A. Hiskens, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1-221, Feb. 2019.
[5] K. Turitsyn, P. Sulc, S. Backhaus *et al.*, "Options for control of reactive power by distributed photovoltaic generators," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1063-1073, Jun. 2011.
[6] B. Zhang, A. D. Domínguez-García, and D. Tse, "A local control approach to voltage regulation in distribution networks," in *Proceedings of North American Power Symposium*, Manhattan, USA, pp. 1-6, Sept. 2013.
[7] N. Li, G. Qu, and M. Dahleh, "Real-time decentralized voltage control in distribution networks," in *Proceedings of the Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, USA, pp. 582-588, Sept. 2014.
[8] K. S. Ayyagari, N. Gatsis, and A. F. Taha, "Chance constrained optimization of distributed energy resources via affine policies," in *Proceedings of the IEEE Global Conference on Signal and Information*, Montréal, Canada, Nov. 2017, pp. 1050-1054.
[9] R. A. Jabr, "Linear decision rules for control of reactive power by distributed photovoltaic generators," *IEEE Transactions on Power Systems,* vol. 33, no. 2, pp. 2165-2174, Mar. 2018.
[10] W. Lin and E. Bitar, "Decentralized stochastic control of distributed energy resources," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 888-900, Jan. 2018.
[11] B. Zhang, A. Y. S. Lam, A. D. Domínguez-García *et al*., "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1714-1726, Jul. 2015.
[12] 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.
[13] F. McLoughlin, A. Duffy, and M. Conlon, "A clustering approach to domestic electricity load profile characterisation using smart metering data," *Applied Energy*, vol. 141, no. 1, pp. 190-199, Mar. 2015.
[14] F. Hasan, A. Kargarian, and A. Mohammadi, "A survey on applications of machine learning for optimal power flow," in *Proceedings of the IEEE Texas Power and Energy Conference (TPEC)*, College Station, USA, pp. 1-6, Feb. 2020.
[15] R. Mieth and Y. Dvorkin, "Data-driven distributionally robust optimal power flow for distribution systems," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 363-368, Jul. 2018.
[16] R. Dobbe, O. Sondermeijer, D. Fridovich-Keil *et al*., "Towards distributed energy services: decentralizing optimal power flow with machine learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1296-1306, Mar. 2020.
[17] 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," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6461-6471, Nov. 2019.
[18] M. Jalali, V. Kekatos, N. Gatsis *et al*., "Designing reactive power control rules for smart inverters using support vector machines," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1759-1770, Mar. 2020.
[19] K. S. Ayyagari, R. Gonzalez, Y. Jin *et al*., "Artificial neural network-based adaptive voltage regulation in distribution systems using data-driven stochastic optimization," in *Proceedings of the IEEE Energy Conversion Congress and Exposition (ECCE)*, Baltimore, USA, Sept. 2019, pp. 5840-5847.
[20] Q. Yang, G. Wang, A. Sadeghi *et al*., "Two-timescale voltage control in distribution grids using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2313-2323, May 2020.
[21] M. Bazrafshan and N. Gatsis, "Voltage regulation in electricity distribution networks using the conditional value-at-risk," in *Proceedings of the IEEE Global Conference on Signal and Information Processing (Global SIP)*, Atlanta, USA, Dec. 2014, pp. 909-913.
[22] E. Dall'Anese, K. Baker, and T. Summers, "Optimal power flow for distribution systems under uncertain forecasts," in *Proceedings of the IEEE Conference on Decision and Control*, Las Vegas, USA, Dec. 2016, pp. 7502-7507.
[23] K. S. Ayyagari, N. Gatsis, A. F. Taha *et al*., "On static and adaptive policies for chance-constrained voltage regulation," in *Proceedings of the Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, USA, Oct. 2018, pp. 1858-1862.
[24] S. Toma, T. Senjyu, Y. Miyazato *et al*., "Decentralized voltage control in distribution system using neural network," in *Proceedings of the IEEE International Power and Energy Conference*, Johor Bahru, Malaysia, Dec. 2008, pp. 1557-1562.
[25] N. A. Zambri, A. Mohamed, and M. Z. C. Wanik, "Performance comparison of neural networks for intelligent management of distributed generators in a distribution system," *International Journal of Electrical Power & Energy Systems*, vol. 67, pp. 179-190, May 2015.
[26] M. Pertl, P. J. Douglass, K. Heussen *et al*., "Validation of a robust neural real-time voltage estimator for active distribution grids on field data," *Electric Power Systems Research*, vol. 154, pp. 182-192, Jan. 2018.
[27] S. Gupta, V. Kekatos, and M. Jin, "Controlling smart inverters using proxies: a chance-constrained DNN-based approach," *IEEE Transactions on Smart Grid*, vol. 13, no. 2, pp. 1310-1321, Mar. 2022.
[28] S. Lin, S. Liu, and H. Zhu, "Risk-aware learning for scalable voltage optimization in distribution grids," *Electric Power Systems Research*, vol. 212, p. 108605, Nov. 2022.
[29] O. Sondermeijer, "Regression-based inverter control for power flow and voltage regulation," Master's thesis, Delft University of Technology, The Netherlands, 2015.
[30] R. Platon, V. R. Dehkordi, and J. Martel, "Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis," *Energy and Buildings*, vol. 92, no. 1, pp. 10-18, Apr. 2015.
[31] V. Kekatos, L. Zhang, G. B. Giannakis *et al*., "Voltage regulation algorithms for multiphase power distribution grids," *IEEE Transactions on Power Systems,* vol. 31, no. 5, pp. 3913-3923, Sept. 2016.
[32] *Electric Power Systems and Equipment Voltage Ratings (60 Herz)*, ANSI Standard C84.1-1995, 2011.
[33] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, no. 3, pp. 21-41, Apr. 2000.
[34] K. Baker, A. Bernstein, E. Dall'Anese *et al*., "Network-cognizant voltage droop control for distribution grids," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2098-2108, Mar. 2018.
[35] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge: The MIT Press, 2016.
[36] The Mathworks Inc. (2018, Dec.). MATLAB version 9.1.0.441655 (R2016b). [Online]. Available: https://www.mathworks.com
[37] D. J. MacKay, "Bayesian interpolation," *Neural Computations*, vol. 4, no. 3, pp. 415-447, May 1992.
[38] S. Montani and L. C. Jain, *Successful Case-based Reasoning Applications*. Berlin: Springer, 2010.
[39] J. Kolodner, *Case-based Reasoning*. San Mateo: Morgan Kaufmann, 2014.
[40] M. Bazrafshan and N. Gatsis, "Comprehensive modeling of three-phase distribution systems via the bus admittance matrix," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2015-2029, Mar. 2018.
[41] M. D. Sankur, R. Dobbe, E. Stewart *et al.* (2016, Jun.). A linearized power flow model for optimization in unbalanced distribution systems. [Online]. Available: https://arxiv.org/abs/1606.04492v2
[42] K. S. Ayyagari, S. A. Abraham, Y. Yao *et al*., "Assessing the optimality of LinDist3Flow for optimal tap selection of step voltage regulators in unbalanced distribution networks," in *Proceedings of the IEEE Conference on Decision and Control (CDC)*, Cancun, Mexico, Dec. 2022, pp. 1-6.
[43] Pecan Street. (2022, Dec.). Pecan street dataport. [Online]. Available: https://dataport.pecanstreet.org/
[44] M. Grant and S. Boyd. (2014, Mar.). CVX: MATLAB software for disciplined convex programming, version 2.1. [Online]. Available: http://cvxr.com/cvx

Krishna Sandeep Ayyagari received the B.E. degree in electrical and electronics engineering from the Jawaharlal Nehru Technological University, Andhra Pradesh, India, in 2010. For the next six years, he pursued his career as Sr. Electrical Engineer (projects and maintenance) at Musimmas Oil Refinery, Nellore, India. He completed the M. S. degree in electrical and computer engineering from the University of Texas at San Antonio, San Antonio, USA, in 2018. Currently, he is working towards the Ph.D. degree in the Department of Electrical and Computer Engineering at the University of Texas at San Antonio, where he is a Graduate Research Assistant. He had a summer internship in 2021 at National Renewable Energy Laboratory (NREL) for the Power Systems Design and Planning (PSDP) group. His research interests include smart grids, control, planning, and optimization of

critical infrastructure including water distribution networks, adaptive protection, and machine learning.

Reynaldo S. Gonzalez is a Ph. D. student in the Department of Electrical and Computer Engineering at the University of Texas at San Antonio, San Antonio, USA. He graduated with his B.S. and M.S. degrees from University of Texas at San Antonio, in 2016 and 2018, respectively. He had a summer internship in 2019 at Ridley Engineering, working on DC-DC converters. He also interned at Southwest Research Institute (SWRI) from 2019 to 2021, for the power group in the Space Science and Engineering Division. His research interests include power electronics, high-frequency converter, GaN-based converter, inductive power transfer, real-time simulation, and applications of artificial intelligence.

Yufang Jin received her B. Sc. degree in automation from Zhengzhou University, Zhengzhou, China, in 1994, and her M.S. and Ph.D. degrees in electrical and computer engineering from the University of Central Florida, Orlando, USA, in 2002 and 2004, respectively. She is a Professor in the Department of Electrical and Computer Engineering at the University of Texas at San Antonio, San Antonio, USA. Her recent research interest includes modeling complex systems using both physical-driven and data-driven methods, especially the application and interpretation of different artificial intelligence algorithms.

Miltiadis Alamaniotis is an Assistant Professor in the Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, USA. Prior to joining University of Texas at San Antonio, he worked as a Research Assistant Professor at Purdue University, West Lafayette, USA. He received the Diploma in Electrical and Computer Engineering from the University of Thessaly, Thessaly, Greece, in 2005, and the M.S. and Ph. D. degrees in nuclear engineering emphasizing on applied artificial intelligence from Purdue University, in 2010 and 2012, respectively. He has been invited to serve as a Guest Editor in the International journal on Artificial Intelligence Tools, and as Program Chair in the IEEE International Tools with Artificial Intelligence 2018. His interdisciplinary research focuses on the development of artificial intelligence and machine learning approaches applied to intelligent energy systems, smart cities, and to nuclear security.

Sara Ahmed is an Associate Professor in Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, USA. She received her Bachelor's, Master's, and Ph.D. degrees all in electrical engineering with a power electronics concentration from Virginia Tech, Center for Power Electronics, in 2006, 2007, and 2011, respectively. She had five years of prior R&D industry experience as a Senior Scientist at ABB's U. S Corporate Research Center, Raleigh, USA. Her primary research interests include modeling, simulation and analysis of power electronics systems with a focus on control, stability, fault analysis, model prediction, integration of renewables and hardware-in-the-loop modeling and testing.

Nikolaos Gatsis received the Diploma (with Hons.) degree in electrical and computer engineering from the University of Patras, Patras, Greece, in 2005, the M. Sc. degree in electrical engineering, and the Ph. D. degree in electrical engineering with minor in mathematics from the University of Minnesota, Minneapolis, USA, in 2010 and 2012, respectively. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, USA. He was the recipient of the CAREER Award by the U.S. National science Foundation. He has also received the UTSA President's Award for Research Achievement. His research interests include optimal and secure operation of smart grids and other critical infrastructures including water distribution networks and global positioning system.