Electric Power Systems Research 189 (2020) 106741

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Data-Driven Optimal Voltage Regulation Using Input Convex Neural Networks☆

Yize Chen*, Yuanyuan Shi, Baosen Zhang

Department of Electrical and Computer Engineering, University of Washington, Seattle, WA, USA

ARTICLE INFO

Keywords:
Distribution Network
Machine Learning
Neural Network
Reactive Power
Voltage Regulation

ABSTRACT

Fast time-scale voltage regulation is needed to enable high penetration of renewables in power distribution networks. A promising approach is to control the reactive power injections of inverters to maintain the voltages. However, existing voltage regulation algorithms require the exact knowledge of line parameters, which are not known for most distribution systems and are difficult to infer. In this work, by utilizing the convexity results of voltage regulation problem, we design an input convex neural network to learn the underlying mapping between the power injections and the voltage deviations. By using smart meter data, our proposed data-driven approach not only accurately fits the system behavior, but also provides a tractable and optimal way to find the reactive power injections. Various numerical simulations demonstrate the effectiveness of the proposed voltage control scheme.

1. Introduction

Voltage regulation in distribution networks has played an important role to maintain acceptable voltage magnitudes at all buses. The higher penetration of distributed energy resources (DERs), for example rooftop PV and electric vehicles, could lead to fast voltage fluctuations in distribution networks [1]. To complement slow time-scale control of discrete devices such as tap-changing transformers and switched capacitors, reactive power injections via the inverter-based distributed resources are often proposed for fast time-scale voltage regulations [2].

Research efforts on inverter-based voltage regulations have focused on approaching the control problem through an optimization framework [3–5]. By formulating it as an Optimal Power Flow (OPF) problem with either voltage deviations objectives [3,5,6] or constraints [7,8], the optimal reactive power injections can be solved via linearized power flow equations or convex relaxation. A significant amount of research efforts has been devoted to this problem and it has been shown to be convex for many systems. Recent advances also extended the centralized control strategies into distributed algorithms with convergence guarantees [6,7,9], and considered uncertainties brought by DERs [10].

However, a fundamental challenge is that distribution networks often suffer from a lack of observability, while system parameters are often unknown or hard to estimate in real systems [11,12]. Yet accuate topology and parameter information are necessities for almost all existing voltage regulation algorithms [13]. This requires additional topology and parameter learning steps before any control actions can be taken. Considering the fact that topology of distribution networks are changing through time, probing network topology and parameters usually require PMU data which are not available in most systems.

To mitigate such burdens on system identification, in [14], a controller is proposed by using linearized system model estimated from advanced metering data. While in [15], a reinforcement learning based controller is directly applied to learn the voltage regulation policy based on power system measurements. However, the optimality and feasibility of such controller are not discussed, and do not leverage the large body of literature on the convexity of the voltage regulation problem. Indeed, in the research community of machine learning and especially reinforcement learning, "learning to control" remains to be a challenging problem when considering physical model constraints and optimality guarantees [16,17].

In this work, we focus on the problem of optimal reactive power compensation for voltage regulation *of distribution networks with unkown topology and parameters*, and we are interested in how to use a *structured neural network* to realize voltage regulations with optimality guarantees. As the cornerstone for deep learning and artificial intelligence, neural networks have shown strong capabilities in function approximation, and have been applied in various tasks such as regression and

☆ The authors are partially supported by the NSF Grants ECCS-1807142 and CNS-1931718.
* Corresponding author. Tel.: 2062299256.
E-mail addresses: yizechen@uw.edu (Y. Chen), yyshi@uw.edu (Y. Shi), zhangbao@uw.edu (B. Zhang).

https://doi.org/10.1016/j.epsr.2020.106741
Received 4 October 2019; Received in revised form 20 April 2020; Accepted 2 August 2020
Available online 08 August 2020
0378-7796/

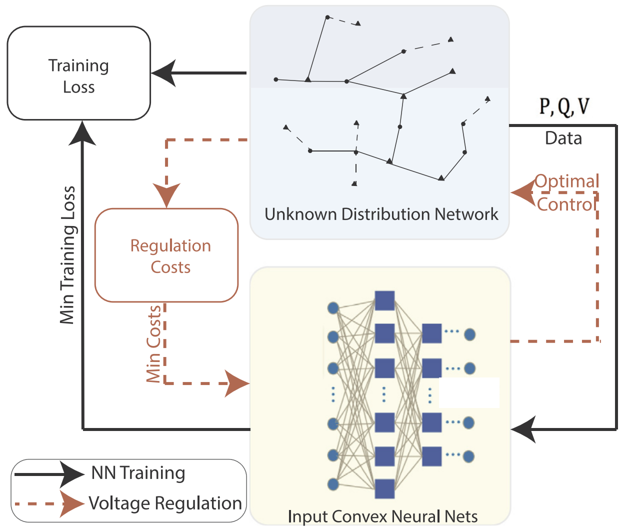

Fig. 1. Proposed data-driven method for distribution networks with unknown topology. An input convex neural network is fitted to learn the mapping from power injections to voltage deviations, and then an optimization problem is solved to find the best reactive injections.

classification [18]. Yet due to the composition of nonlinear functions, normal neural networks are hard to be directly applied for decision-making or solving optimization problem. We tackle the modeling accuracy and control tractability tradeoff by building on the *input convex neural networks (ICNN)* in [19,20] to both represent unknown distribution systems and to find optimal reaction power injections.

ICNN bridges the strong learning and modeling capabilities of deep neural networks with physical insights of voltage regulation by solving a constrained convex optimization problem. With specifically designed neural network architecture, the output of ICNN is guaranteed to be convex with respect to the neural network's inputs, which is a good fit in solving convex optimization problem. As shown in Fig. 1, in the training stage, ICNN is making use of system operating states measurements to learn the unknown, nonlinear mappings from active and reactive power injections to nodal voltage deviations; in the voltage regulation stage, ICNN serves as the model to be optimized over, and is guaranteed to find the optimal reactive power injections by design. By leveraging the convexity results [3] of voltage regulation problem, the proposed method can be used as a plug-and-play component for distribution system operations, which does not require explicit topology learning nor parameter inference techniques, while always satisfying reactive power capacity constraints.

The remainder of the paper is organized as follows. In Section 2 we introduce the power flow model and the voltage regulation problem formulation; in Section 3 we present the setup for our specifically designed ICNN, and describe how to apply such neural networks into learning and voltage control; simulations on IEEE 13-bus and 123-bus distribution illustrate the performance of proposed scheme; concluding remarks are presented in Section 5.

2. Modeling and Problem Formulation

2.1. Power System Model and Voltage Regulation Problem

Consider a power distribution network consisting of a set $\mathcal{N} = \{1, ..., N\}$ of buses and a set $\mathcal{E} \in \mathcal{N} \times \mathcal{N}$ of distribution lines connecting buses. For each bus i, denote V_i as the voltage magnitude and θ_i as the voltage phase angle; let p_i and q_i denote the active and reactive power injections; let $s_i = p_i + jq_i$ be the complex power injection at bus i. The corresponding active and reactive power injection vectors are denoted as $\mathbf{p} = [p_1 \ \ p_2 \ \ \cdots \ \ p_N]^T$, $\mathbf{q} = [q_1 \ \ q_2 \ \ \cdots \ \ q_N]^T$. For each line $(i, k) \in \mathcal{E}$, denote line admittance $y_{ik} = g_{ik} - jb_{ik}$ with $b_{ik} > 0, g_{ik} > 0$ as the real and imaginary parts of the admittance matrix element Y_{ik}. For each bus $i \in \mathcal{N}$, its power injection is governed by

$$p_i = \sum_{k=1}^{N} V_i V_k(-g_{ik}\cos(\theta_i - \theta_k) + b_{ik}\sin(\theta_i - \theta_k)) \tag{1a}$$

$$q_i = \sum_{k=1}^{N} V_i V_k(g_{ik}\sin(\theta_i - \theta_k) + b_{ik}\cos(\theta_i - \theta_k)). \tag{1b}$$

The focus of this paper is to address the voltage fluctuations due to higher penetration level of DERs in distribution networks. By making use of the power electronics interfaces of DERs such as PV inverters, the reactive power injections q_i can be controlled within certain limits. By changing reactive power injections, the goal is to maintain voltage magnitude V_i within a small distance from the nominal value $V_{i,0}$ for all buses (e.g., plus/minus 5%). Formally, we can cast voltage regulation as the following optimization problem:

$$\min_{\mathbf{q}} \quad \sum_{i=1}^{N} \alpha_i |V_i - V_{i,0}| \tag{2a}$$

$$\text{s.t.} \quad \underline{\mathbf{q}} \le \mathbf{q} \le \bar{\mathbf{q}} \tag{2b}$$

$$\text{Power Flow Equations (1)} \tag{2c}$$

where α_i is a weighted parameter which can be adjusted by the system operator. The constraints in (2b) capture the hard constraints on available reactive power injections on each bus i. The constraints in (2c) capture the power flow models. The active power $\mathbf{p}$ is considered as an exogenous input vector which is not controlled.

Even when the distribution network is known exactly, e.g., the network topology and $\{g_{ik}, b_{ik}\}$ values are available, directly solving (2) is still not trivial, because the problem is not convex due to the nonlinear relationship between bus voltage magnitudes and powers. In the next subsection, we will briefly review previous results showing that it is possible to reformulate voltage regulation as a convex optimization problem. However, practical concerns on the knwoledge of distribution networks have impeded their applications.

2.2. Convexification and Practical Challenges

For each line $(i, k) \in \mathcal{E}$, denote $s_{ik} = p_{ik} + jq_{ik}$ as the complex power flow and $z_{ik} = r_{ik} + jx_{ik}$ as the line impedance. The DistFlow equations [21] model the distribution network flow as

$$-p_k = p_{ik} - r_{ik} l_{ik} - \sum_{l:(k,l)\in\mathcal{E}} p_{kl} \tag{3a}$$

$$-q_k = p_{ik} - x_{ik} l_{ik} - \sum_{l:(k,l)\in\mathcal{E}} q_{kl} \tag{3b}$$

$$V_k^2 = V_i^2 - 2(r_{ik}p_{ik} + x_{ik}q_{ik}) + (r_{ik}^2 + x_{ik}^2) l_{ik} \tag{3c}$$

$$l_{ik} = \frac{p_{ik}^2 + q_{ik}^2}{V_i^2}. \tag{3d}$$

By further making the following relaxation for (3d)

$$l_{ik} \ge \frac{p_{ik}^2 + q_{ik}^2}{V_i^2} \tag{4}$$

which can be written as a second-order cone constraint, the relaxed constratins together with the voltage regulation objective is then a Second Order Cone Program (SOCP) [3].

Remark 2.1. The relaxed voltage control problem with regulation objective (2a), reactive power injection constraint (2b), power flow constraints (3a)-(3c) and (4) is convex. Under many circumstances, this relaxation is tight, see [3,4].

In order to further simplify the analysis of the original voltage control problem (2), many linearized power flow models are adopted, while Simplified Distflow model is widely used [5,22], which sets l_{ik} to be zeros, and approximates $V_i^2 - V_k^2$ by $2(V_i - V_k)$:

$$-p_k = p_{ik} - \sum_{l:(k,l)\in\mathcal{E}} p_{kl} \tag{5a}$$

$$-q_k = q_{ik} - \sum_{l:(k,l)\in\mathcal{E}} q_{kl} \tag{5b}$$

$$V_i - V_k = r_{ik} p_{ik} + x_{ik} q_{ik}. \tag{5c}$$

Similarly, by replacing (2c) with the linearized version (5), we are also able to solve the voltage regulation as a convex optimization problem. However, considering the increasing variability of load and generation in distribution networks, voltage regulation based on linearized approximation model (5) may not be accurate enough to represent the true distribution network models, while the resulting control signals of reactive power injections may not be optimal when applied in the real distribution networks.

In summary, to solve voltage regulation as a convex optimization, all these aforementioned approaches require the exact information on line parameters (e.g., line impedances) and network topology. Unfortunately, due to the lack of observability of distribution systems, directly learning the topology is hard without PMU data [11,12].

2.3. Practical Design Requirements

Given the practical challenges of voltage regulation, we want to design an optimal controller that satisfies following requirements:

- The controller must learn an accurate representation of the power injections to nodal voltage magnitudes;
- Such representation is easy to be integrated into the optimization framework.

Intuitively, we are trying to design and find functions $|V_i - V_{i,0}| = f_i(\mathbf{p}, \mathbf{q})$, $i = 1, ..., N$, which could accurately represent the relationship from active and reactive power injections to nodal voltage magnitude deviation. By leveraging historical smart meter data to fit f_i, we want to see if the fitted model could represent the underlying grid. More importantly, if f_i is a convex function from $\mathbf{p}, \mathbf{q}$ to $|V_i - V_{i,0}|$, then the following problem

$$\min_{\mathbf{q}} \sum_{i=1}^{N} \alpha_i |V_i - V_{i,0}| \tag{6a}$$

$$\text{s.t.} \quad \underline{\mathbf{q}} \le \mathbf{q} \le \bar{\mathbf{q}} \tag{6b}$$

$$|V_i - V_{i,0}| = f_i(\mathbf{p}, \mathbf{q}) \tag{6c}$$

is still a tractable convex optimization problem. Note that we integrate voltage magnitude deviations constraint (6c) into the voltage regulation framework, which is a general formulation to make sure once f_i is convex, (6) is a convex optimization problem. Such formulation is comparable to previous formulations by either treating voltage magnitude deviations as the optimization objective [4] or as box constraints [7,10].

Because of the convexity results [3], restricting f_i to be convex leads to optimal solutions. In the next section, we will describe how we design the learning model f_i based on neural networks, which not only efficiently learns the mapping from active and reactive power injections to the voltage magnitude deviations more powerfully than a linear model, but is also guaranteed to be a convex function.

3. Input Convex Neural Networks

In this section, starting from the standard neural networks architecture, we illustrate how to construct a neural network whose outputs are convex with respect to inputs. We then show how to apply such input convex neural networks (ICNN) in the task of voltage regulation in distribution networks, and describe a practical algorithm to find optimal reactive power injections under reactive power capacity constraints.

3.1. Neural Networks for Function Fitting

For a standard setup of neural networks (NN) model, the multi-layer network is composed of an input layer x, m hidden layers z_l, $l = 1, ..., m$ with parameters $\theta := \{W_l, b_l\}$ $i = 1, ..., m$, and an output y. For notation simplicity, we use $h_\theta(x)$ to denote the neural networks with input x and parameters θ. For the computation at layer l, in addition to the matrix multiplication using W_l and b_l, activation function $g_i(\cdot)$ has also been widely adopted to increase the nonlinearity from input to output. For instance, rectified linear unit (ReLU) is a popular choice with $g(x) = \max(0, x)$. Given input x, a neural network is implementing the following computation

$$z_1 = g_1(W_1 x + b_1);$$
$$z_l = g_{l+1}(W_l z_{l-1} + b_l), \quad l = 2, ..., m$$

and the neural network output is the value of the last layer z_m. In the task of supervised learning, back-propagation algorithms based on gradient descent are used to train a group of $\{W_l, b_l\}$ that minimize the training loss defined as $L(y, h_\theta(x))$ [23][1]. Starting from this basic building blocks of neural networks, many major advancements have been made in learning complicated, nonlinear functions such as image classification [18], load forecasts [24] and model-free controller design [15]. We are interested to see if we could borrow such strong modeling power into modeling the underlying unknown distribution grids. Essentially, we want to design specific neural networks such that the output of $h_\theta(x)$ is convex with respect to x, which could be then representing f_i in (6c) and be used for solving (6) once the model is trained.

3.2. ICNN Architecture Design

We adapt the neural networks design from our previous work [20] and the original input convex neural networks (ICNN) proposed in [19] to the setting of multi-inputs (e.g., $\{\mathbf{p}, \mathbf{q}\}$) and multi-outputs (e.g., $|V_i - V_{i,0}|, \forall i$). Specifically, by convexity we mean each dimension of ICNN's output is convex with respect to all the dimension of inputs.

The following proposition summarizes the major adaptations we make to the standard neural networks that guarantee modeling convexity:

Proposition 3.1. *The network shown in Fig. 2(a) is a convex function from inputs to outputs provided that all $W_{2:m}$ are non-negative, and all g_l are convex and non-decreasing functions.*

The convexity of the proposed neural network directly follows from the composition rule of convex functions [25], which states that the composition of an inner convex function and an outer convex, non-decreasing function is convex. The restriction on the g_l function to be convex and non-decreasing function is actually not a strong restriction. Popular activation functions like ReLU function shown in Fig. 2(c) already satisfies such restriction. Thus with the nonnegative constraints on $W_{2:m}$ and with the choice of activation functions, we are already constructing a neural network whose output is convex with respect to inputs.

To compensate the neural network representability loss due to the constraints of $W_l \ge 0$, $l = 2, ..., m$, we add direct passthrough layers

[1] The choice of training loss can be task specific, e.g., cross-entropy loss, mean squared error.

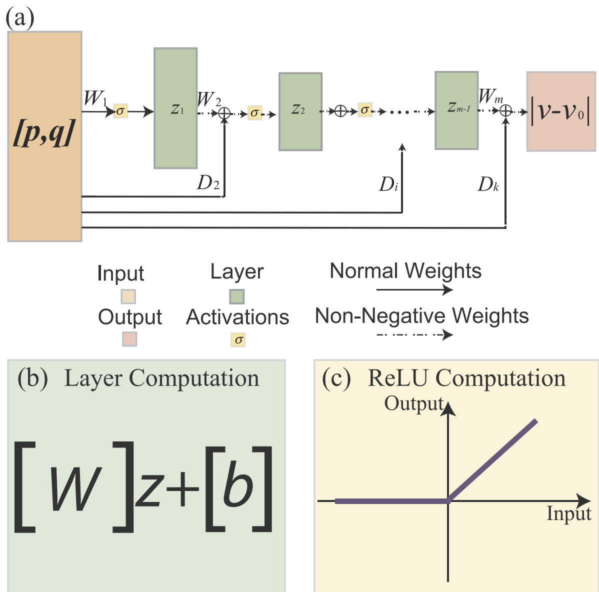

Fig. 2. (a). The input convex neural networks (ICNN) architecture design; (b) the layer computation for ICNN, where we constrain $W_{2:\,m}$ to be non-negative; (c). a ReLU nonlinear activation function.

D_l, $l = 2, \ldots m$ from input to subsequent layers, and there is no constraint on the weights of these links. Such direct links have also been widely used in the design of deeper neural networks, which have achieved better performance in various learning tasks [26]. Combined with layer bias b_l, layer i passes its value through ReLU activation function and goes to next layer $i + 1$.

Mathematically, for each layer $l = 1, \ldots, m$, the layerwise computations are modified as follows

$$z_1 = g_1(W_1 x + b_1); \tag{8a}$$

$$z_l = g_l(W_l z_{l-1} + D_l x + b_l), \quad l = 2, \ldots, m \tag{8b}$$

We note that our neural networks design naturally extends to the scenario when input x and output y are high-dimensional vectors. Meanwhile, each dimension of output is convex with respect to all the inputs. Such property also makes it possible to fit a single ICNN h_θ to model multiple convex functions.

A natural question is related to the function approximation capabilitiy of the proposed ICNN. For instance, once ICNN is used in (6) for voltage regulation, it should be able to fit the underlying convex functions from $\mathbf{p}, \mathbf{q}$ to $|V_i - V_{i,0}|$ as accurate as possible. The following Lemma from [20] guarantees that ICNN is able to fit any convex functions, which opens the door to integrate ICNN in convex optimization:

Lemma 3.2. *For any Lipschitz convex function over a compact domain, there exists a neural network with nonnegative weights and ReLU activation functions that approximates it arbitrary closely.*

We refer interested readers to [20] for detailed proof. The sketch of proof follows from the fact that any continuous, Lipschitz convex function can be approximated by the maximum of a finite number of affine functions [27]. Then it is sufficient to show that ICNN can implement any max of affine functions.

3.3. Neural Networks and Optimization

When the underlying topology and the line parameters are unknown, we propose to first learn a convex mapping from $\{\mathbf{p}, \mathbf{q}\}$ to voltage magnitude deviations using an ICNN. Once fitted using collected observations, we are able to use the same ICNN, and integrate it to (6) to find optimal reactive power injections.

In order to train ICNN and learn its parameters h_θ, we need to minimize the supervised traning loss. For the kth training instance, it is defined as the mean square error between the ground truth voltage magnitude deviation vector $\mathbf{V}_{target}$: $=\{|V_i^k - V_{i,0}^k|\}$, $i = 1, \ldots, N$ and the ICNN output:

$$L(\mathbf{V}_{target}, h_\theta(\mathbf{p}, \mathbf{q})) = \frac{1}{N} \|\mathbf{V}_{target} - h_\theta(\mathbf{p}, \mathbf{q})\|_2^2, \tag{9}$$

and the update of h_θ is based on gradient descent algorithm. In addition, to take the constraints of $W_{2:\,m} \geq 0$ into account, we need to make sure the gradient descent update always falls into the feasible regions (e.g., nonnegative weights). Hence we use a projected gradient algorithm to guarantee the constraint holds [25].

Definition 3.3. The projection of a point y, onto a set X is defined as

$$\Pi_X(y) = \underset{x \in X}{\operatorname{argmin}} \frac{1}{2} \|x - y\|_2^2 \tag{10}$$

Given a starting point $x^{(0)} \in X$ and step-size $\gamma > 0$, projected gradient descent (PGD) extends the standard gradient descent settings with the projection step onto the feasible sets of feasible reactive power. At iteration t, the algorithm takes the following PGD step:

$$x^{(t+1)} = x^{(t)} - \gamma \Pi_X(x^{(t)} - \nabla f(x^{(t)})), \ \forall\ t \geq 1 \tag{11}$$

which is implemented iteratively until a certain stoppping criterion (e.g., fixed number of iterations or gradient value is smaller than predefined ϵ) is satisfied. The ICNN weights are then updated as follows

$$h_\theta = h_\theta - \gamma \Pi_{W_{2:m} \geq 0}(h_\theta - \nabla_{h_\theta}(L(\mathbf{V}_{target}), h_\theta(\mathbf{p}, \mathbf{q}))). \tag{12}$$

In practical implementations where there are large groups of measurements $\{\mathbf{p}^k, \mathbf{q}^k, \mathbf{V}^k\}$ with k standing for the index for measurement index, it is possible to use small batch of training data to do PGD steps (12). Such practical algorithms, e.g., stochastic gradient descent, can accelerate training convergence [23]. The training procedure is summarized in Algorithm 1 by using collected training data and stochastic gradient descent training algorithm.

Once the ICNN training process is finished, we fix model parameters h_θ, and use it as a proxy model for the unknown distribution networks model f_i, $i = 1, \ldots, N$ in (6c). Since h_θ represents the convex mappings from $\mathbf{q}$ to $|V_i - V_{i,0}|$, $\forall\ i$, we are now ready to solve (6) computationally. In the similar spirit of ICNN training, where we optimize over network weights using gradient descent to minimize training loss, in the voltage regulation setting, we optimize over ICNN inputs $\mathbf{q}$ to minimize the optimization objective $\sum_{i=1}^{N} \alpha_i |V_i - V_{i,0}|$ using gradient descent. Again, to take the constraints of reactive power injection range into account, we need to make sure that the gradient descent update always falls into the feasible reactive power injection regions. By adapting PGD to the trained ICNN, starting from uncontrolled reactive power $\mathbf{q}^{(0)} = \mathbf{q}$, we take iterative PGD steps on the voltage regulation objective (6a) until gradient convergence. PGD steps also guarantee the convergence to optimal solution under the convex settings. The overall algorithm for ICNN training and finding optimal reactive power injections are described in Algorithm 1.

4. Numerical Results

In this section, we evaluate the proposed voltage regulation scheme on standard distribution networks. Our simulation focus on the IEEE 13-bus and IEEE 123-bus test systems. Linear models, standard neural networks and the optimal SOCP formulations are used for comparison.

4.1. Simulation Setup

For both 13-bus and 123-bus system, we use AC power flow

Require: Learning rate η, Step size γ, Batch size T, Training iterations $n_{training}$, Optimization stopping criterion ϵ
Ensure: Training dataset $\{\mathbf{p}, \mathbf{q}, \mathbf{V}\}$
Ensure: Initial model h_θ
  # ICNN Training
  for $iter = 0, ..., n_{training}$ **do**
    # Update parameters for ICNN
    Sample batch from historical data:
    $\{\mathbf{p}^k, \mathbf{q}^k, \mathbf{V}^k\}_{k=1}^T \sim \mathbb{P}_x$
    Update h_θ using stochastic gradient descent:
    $\mathbf{V}_{target}^k = \{|V_i^k - V_{i,0}^k|\}, i = 1, ..., N$
    $h_\theta = h_\theta - \eta \Pi_{W_{2:m} \geq 0}\left(h_\theta - \nabla_{h_\theta}(L(\mathbf{V}_{target}), h_\theta(\mathbf{p}, \mathbf{q}))\right)$
  end for
  Fix ICNN parameters h_θ
  # Voltage Regulation via ICNN
  Get measurements $\{\mathbf{p}, \mathbf{q}\}$, $t = 0$, $\mathbf{q}^{(t)} = \mathbf{q}$
  while $\nabla_{\mathbf{q}^{(t)}}(\sum_{i=1}^N h_\theta(\mathbf{p}, \mathbf{q}^{(t)})) > \epsilon$ **do**
    $\mathbf{q}^{(t+1)} = \mathbf{q}^{(t)} - \gamma \Pi_{\mathbf{q}}(\mathbf{q}^{(t)} - \nabla_{\mathbf{q}^{(t)}}(\sum_{i=1}^N h_\theta(\mathbf{p}, \mathbf{q}^{(t)})))$
    $t \leftarrow t + 1$
  end while
  Optimal reactive power injection: $\mathbf{q}^* = \mathbf{q}^{(t)}$

Algorithm 1. ICNN for Voltage Regulation

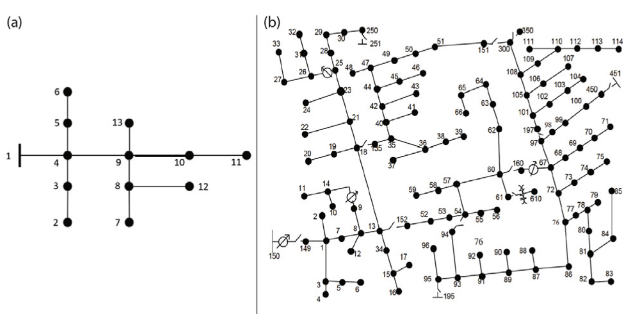

Fig. 3. Schematic diagram of (a). IEEE 13-bus test feeder and (b). IEEE 123-bus test feeder. Reference buses: 1 and 149.

model (1) to generate 10,000 instances of simulation data composed of $\{\mathbf{p}, \mathbf{q}, \mathbf{V}\}$. We assume both the distribution network topology and line parameters are not revealed to the optimization algorithm, except when the optimal SOCP is used as a baseline. We allow plus/minus 20% of reactive power injections at each node as control inputs. We develop three algorithms and compare their performances for two test feeders shown in Fig. 3:

- *Linear Model:* We consider using a linear model to fit the unknown dynamics from active and reactive power to the deviations between nodal voltage and the nominal voltage. Such linearized models have been widely used in power systems literature [5,10];
- *Neural Networks Model:* We construct standard three-layer and four-layer neural networks for the 13-bus and 123-bus cases, respectively. We tune the parameters of neural networks (e.g., number of neurons, learning rate) and stop the training process once the fitting performance on validation data converges;
- *Input Convex Neural Networks:* We keep the number of layers and matrices W_i, $i = 1, ..., k$ the same dimension as those of neural networks models, but add direct layers D_i, $i = 2, ..., k$ correspondingly. We constrain network weights $W_{2:\ k}$ to be non-negative during training.

To fit the parameters of neural network models, we use mean squared error as the loss function during training. To solve the voltage regualtion problem (2), we set α_i, $i = 1, ..., k$ in (6) to be 1 in our simulation cases. When α is not equal to 1, we could adapt the optimization problem using weighted sum of voltage deviation correspondingly. Note that we could also flexibly add reactive power costs to the objective function (2a), as long as they are convex functions over reactive power injections. All the implementations are conducted on a MacBook Pro with 2.4GHz Intel Quad Core i5.

To benchmark the performance of the proposed algorithms under unknown topology and parameters, we also follow [3] to relax $l_{ij} \geq \frac{P_{ij}^2 + Q_{ij}^2}{V_i^2}$ in the Dist-flow equations, and use the same validation datasets to solve the resulting convex SOCP. We calculate the optimal reactive power injections along with the resulting voltage profiles. We use CVX to solve the SOCP and linearized models [28], and use Tensorflow to set up and solve NN and ICNN models [29].

4.2. Estimation Accuracy

We firstly validate that ICNN can be used as a proxy for power flow equations, and predict the nodal voltage magnitude deviations. By

Table 1
Comparison between SOCP, Linear model, Neural Networks model, and Input Convex Neural Networks model for IEEE 13-bus and IEEE 123-bus systems.

SImulation Network	IEEE 13-Bus				IEEE 123-Bus		
Model	SOCP	Linear	NN	ICNN	Linear	NN	ICNN
Model Fitting MAE	-	9.93%	3.45%	3.86%	12.98%	3.56%	4.25%
Regulated voltage out of 3% tolerance	3.46%	8.65%	7.88%	4.71%	21.46%	14.04%	7.51%
Regulated voltage out of 5% tolerance	0.47%	7.89%	6.86%	1.05%	19.19%	9.65%	1.64%
Computation Time (per instance/s)	0.9684	0.2022	0.3137	0.2512	0.2712	0.6297	0.4302

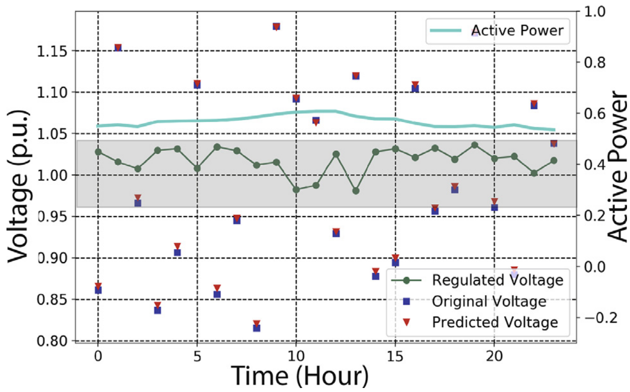

Fig. 4. Example of voltage regulation over a daily variation for the 13-bus test feeder. The voltage of bus 4 is shown. With ICNN accurately predicting voltages (red triangle), it could regulate voltage within 4% of nominal values (grey box) under varying load level throughout the day.

using 8,000 training instances, the ICNN can predict the voltage deviations on the validation instances accurately. As shown in Table 1, the mean absolute error (MAE) of ICNN fitting are smaller than 4.3% in both test systems, which are comparable to 3.45% and 3.56% by using neural networks. This is also illustrated in Fig. 4, where under different load levels throughout 24 hours, the ICNN can predict all the nodal voltages accurately. More importantly, linear model's fitting performances are over 2 times worse than the neural networks counterparts. We later show such fitting errors would also impact the controller performances.

4.3. Voltage Regulation Performance

In Figure 4, we show the regulated voltage using ICNN in the IEEE 13-bus case. Under this day's load profile, we are able to regulate node 4's voltage magnitude within $\pm$ 4% per unit with constrained reactive power injections (Equation 6b). In Figure 5, we show that the mean and variance on each bus's voltage deviations using three models for the 13-bus feeder. On the one hand, with similar fitting performances, ICNN outperforms the standard neural network in regulating nodal voltages. This is due to the fact that neural networks may have many local minima, and the NN-based controller can not find the optimal reactive power injections. On the other hand, even though linear model provides a easier venue for solving optimization problem, it suffers from inaccurate modeling of the underlying distribution grids, and the regulated bus voltages have greater level of fluctuations. Similar observations also hold in the 123-bus test case, where in Fig. 6 we show the nodal voltage comparison using three models, and voltage regulated by

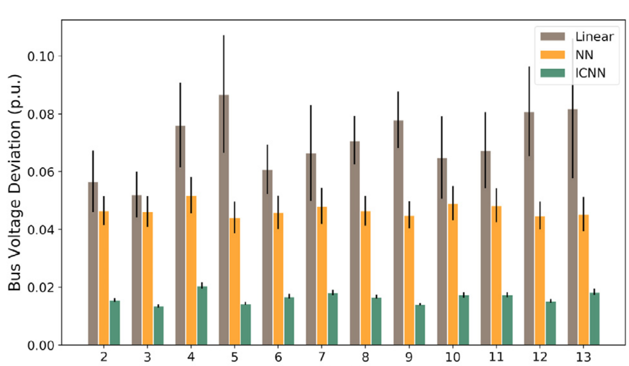

Fig. 5. Comparisons on nodal voltage deviation bar plots of linear-fitted model, neural network model and input convex neuural network model on IEEE 13-bus system. On average, the mean voltage deviation for ICNN is 4.3 times better than linear model, and 2.7 times better than standard NN model.

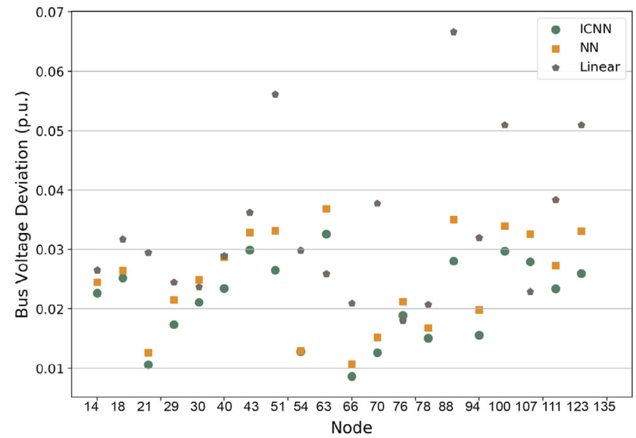

Fig. 6. Comparisons on 20 randomly selected buses' nodal voltage deviation plots of linear-fitted model, neural network model and input convex neuural network model on IEEE 123-bus system.

ICNN are constrained to be in a much narrower range. More results on voltage regulation performances are summarized in Table 1. Under varying load and power generation profiles, ICNN is able to maintain over 98.3% of nodal voltages within 5% deviations from nominal voltages, which are comparable to SOCP solutions. On the contrary, linear fitted models can not scale to larger system, and nearly 20% of voltages are out of 5% tolerance in the 123-bus case. We note that with more injections of reactive power, the proposed control scheme can make use of such injections to achieve better regulation performance.

We also give an analysis on the computation time for each algorithm. Compared to linear model, optimization based on ICNN generally takes longer time, but it is still able to find the optimal solutions within the acceptable time range. Note that we are solving the ICNN optimization problem using our own solver, while solving SOCP and linear model using the off-the-shelf CVX solver. More importantly, in the 13-bus case, ICNN-based optimization is faster than SOCP solver, and it scales to 123-bus case with moderate computation time increases, while SOCP solver is hard to scale to larger network. An interesting observation is that it takes longer for NN to find solutions compared to ICNN, partly due to the fact that gradient-based optimizer is stuck in some local minima. We will also discuss the performance of proposed method on distributed case in the future work [30].

5. Conclusion

This paper describes an optimal voltage control algorithm for distribution networks with unknown topology and models. To close the gap between unknown models and optimal regulation decision-making, we proposed to construct a specifically designed neural network, the input convex neural network, which could be used for both distribution grid modeling and solving voltage regulation problem in a convex optimization framework. Numerical case studies validate that proposed scheme is able to achieve significantly better voltage regulation performances compared to current state-of-the-art model-based and learning-based techniques. In the future work, we would like to explore robust, data-driven voltage regulation framework with the existence of noises or adversarial data, and investiage more available control options such as mobilizing the active power resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P.M. Carvalho, P.F. Correia, L.A. Ferreira, Distributed reactive power generation control for voltage rise mitigation in distribution networks, IEEE transactions on Power Systems 23 (2) (2008) 766–772.
[2] P. Jahangiri, D.C. Aliprantis, Distributed volt/var control by pv inverters, IEEE Transactions on power systems 28 (3) (2013) 3429–3439.
[3] M. Farivar, C.R. Clarke, S.H. Low, K.M. Chandy, Inverter var control for distribution systems with renewables, 2011 IEEE international conference on smart grid communications (SmartGridComm), IEEE, 2011, pp. 457–462.
[4] M. Farivar, R. Neal, C. Clarke, S. Low, Optimal inverter var control in distribution systems with high pv penetration, 2012 IEEE Power and Energy Society general meeting, IEEE, 2012, pp. 1–7.
[5] H. Zhu, H.J. Liu, Fast local voltage control under limited reactive power: Optimality and stability analysis, IEEE Transactions on Power Systems 31 (5) (2015) 3794–3803.
[6] B. Zhang, A.Y. Lam, A.D. Domínguez-García, D. Tse, An optimal and distributed method for voltage regulation in power distribution systems, IEEE Transactions on Power Systems 30 (4) (2014) 1714–1726.
[7] G. Qu, N. Li, An optimal and distributed feedback voltage control under limited reactive power, 2018 Power Systems Computation Conference (PSCC), IEEE, 2018, pp. 1–7.
[8] S. Bolognani, R. Carli, G. Cavraro, S. Zampieri, Distributed reactive power feedback control for voltage regulation and loss minimization, IEEE Transactions on Automatic Control 60 (4) (2014) 966–981.
[9] S. Magnússon, G. Qu, C. Fischione, N. Li, Voltage control using limited communication, IEEE Transactions on Control of Network Systems (2019).
[10] P. Li, B. Jin, D. Wang, B. Zhang, Distribution system voltage control under uncertainties using tractable chance constraints, IEEE Transactions on Power Systems (2018).
[11] H. Li, Y. Weng, Y. Liao, B. Keel, K.E. Brown, Robust hidden topology identification in distribution systems, arXiv preprint arXiv:1902.01365(2019).
[12] D. Deka, S. Backhaus, M. Chertkov, Estimating distribution grid topologies: A graphical learning based approach, 2016 Power Systems Computation Conference (PSCC), IEEE, 2016, pp. 1–7.
[13] S.H. Low, Convex relaxation of optimal power flow-part i: Formulations and equivalence, IEEE Transactions on Control of Network Systems 1 (1) (2014) 15–27.
[14] H. Xu, A.D. Domínguez-García, P.W. Sauer, A data-driven voltage control framework for power distribution systems, 2018 IEEE Power & Energy Society General Meeting (PESGM), IEEE, 2018, pp. 1–5.
[15] Q. Yang, G. Wang, A. Sadeghi, G.B. Giannakis, J. Sun, Real-time voltage control using deep reinforcement learning, arXiv preprint arXiv:1904.09374(2019).
[16] R.S. Sutton, A.G. Barto, et al., Introduction to reinforcement learning, 2 MIT press Cambridge, 1998.
[17] S. Miryoosefi, K. Brantley, H. Daumé III, M. Dudik, R. Schapire, Reinforcement learning with convex constraints, Advances in Neural Information Processing Systems, (2019), pp. 1–9.
[18] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436.
[19] B. Amos, L. Xu, J.Z. Kolter, Input convex neural networks, International Conference on Machine Learning, (2017), pp. 146–155.
[20] Y. Chen, Y. Shi, B. Zhang, Optimal control via neural networks: A convex approach, International Conference on Learning Representations (ICLR), (2019), pp. 1–11.
[21] M.E. Baran, F.F. Wu, Optimal capacitor placement on radial distribution systems, IEEE Transactions on power Delivery 4 (1) (1989) 725–734.
[22] M.E. Baran, F.F. Wu, Network reconfiguration in distribution systems for loss reduction and load balancing, IEEE Transactions on Power delivery 4 (2) (1989) 1401–1407.
[23] L. Bottou, Large-scale machine learning with stochastic gradient descent, Proceedings of COMPSTAT'2010, Springer, 2010, pp. 177–186.
[24] Y. Chen, Y. Shi, B. Zhang, Modeling and optimization of complex building energy systems with deep neural networks, 2017 51st Asilomar Conference on Signals, Systems, and Computers, IEEE, 2017, pp. 1368–1373.
[25] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge university press, 2004.
[26] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, (2016), pp. 770–778.
[27] S. Wang, General constructive representations for continuous piecewise-linear functions, IEEE Transactions on Circuits and Systems I: Regular Papers 51 (9) (2004) 1889–1896.
[28] M. Grant, S. Boyd, Cvx: Matlab software for disciplined convex programming, version 2.1, 2014.
[29] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., Tensorflow: A system for large-scale machine learning, 12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16), (2016), pp. 265–283.
[30] Y. Chen, Y. Shi, B. Zhang, Input convex neural networks for optimal voltage regulation, arXiv preprint arXiv:2002.08684(2020).