Optimal steady-state voltage control using Gaussian process learning

Pareek, Parikshit; Yu, Weng; Nguyen, Hung Dinh

2020

Pareek, P., Yu, W. & Nguyen, H. D. (2020). Optimal steady-state voltage control using Gaussian process learning. IEEE Transactions On Industrial Informatics, 17(10), 7017-7027. https://dx.doi.org/10.1109/TII.2020.3047844

https://hdl.handle.net/10356/150722

https://doi.org/10.1109/TII.2020.3047844

Optimal Steady-state Voltage Control using Gaussian Process Learning

Parikshit Pareek†, *Student Member, IEEE,* Weng Yu†, *Student Member, IEEE,* and Hung Dinh Nguyen⋆, *Member, IEEE*

***Abstract*—In this paper, an optimal steady-state voltage control framework is developed based on a novel linear Voltage-Power dependence deducted from Gaussian Process (GP) learning. Different from other point-based linearization techniques, this GP-based linear relationship is valid over a subspace of operating points and thus suitable for a system with uncertainties such as those in power injections due to renewables. The proposed optimal voltage control algorithms, therefore, perform well over a wide range of operating conditions. Both centralized and distributed optimal control schemes are introduced in this framework. The least-squares estimation is employed to provide analytical forms of the optimal control which offer great computational benefits. Moreover, unlike many existing voltage control approaches deploying fixed voltage references, the proposed control schemes not only minimize the control efforts but also optimize the voltage reference setpoints that lead to the least voltage deviation errors with respect to such setpoints. The control algorithms are also extended to handle uncertain power injections with robust optimal solutions which guarantee compliance with the voltage regulation standards. As for the distributed control scheme, a new network partition problem is cast, based on the concept of Effective Voltage Control Source (EVCS), as an optimization problem which is further solved using convex relaxation. Various simulations on the IEEE 33-bus and 69-bus test feeders are presented to illustrate the performance of the proposed voltage control algorithms and EVCS-based network partition.**

***Index Terms*—Steady-state Voltage Control, GP Learning, Effective Voltage Control Source (EVCS)**

I. Introduction

THE modern distribution systems across the world are experiencing an ever-increasing level of renewable penetration and electric vehicle (EV) integration. The introduction of such distributed generations (DG) and EV fleets bring a great deal of uncertainties, typically in power injections, that further compromise the voltage compliance. Then the voltage regulation problem needs to be revisited in the presence of such uncertainty sources [1], [2].

In a power system, the objective of the steady-state voltage control problem is to maintain all nodal voltages at an acceptable level when the system is subject to disturbances such as load changes [3], [4]. This voltage control problem can be challenging to solve due to its large scale and complexity. A typical power system may consist of a large number of components, i.e., loads, generators, compensators, etc. The inherent nonlinear interaction among these components constitutes the complexity of many steady-state problems including the voltage control [5], [6]. Thus, designing an optimal voltage control framework that is able to regulate the voltage profile of the entire large-scale system effectively becomes crucial to the system operator [7]. Various voltage control approaches have been developed in the literature [8]. One important approach is to incorporate the voltage control objective into an optimal power flow (OPF) framework [9]. While recent years have witnessed tremendous success in applying OPF, such as those using convexification approaches, the high dimensional issue and the inherent nonlinear property remains the major hurdles in solving such optimization problem in practice. Thus, the OPF-based voltage control can be computationally expensive and time-consuming [10]. As a result, the remedy control actions, if are designed relying on such an OPF framework, cannot be timely implemented to control the system's voltage levels [4]. Another contemporary trend in voltage control research focuses on the uncertainties such as those brought by intermittent renewable sources and electrical vehicles [2], [11], [12]. As such uncertainties pose additional challenges in operation, most of the proposed voltage control algorithms, including decentralized controls under limited communication conditions, require certain assumptions/simplifications to perform well [8], [13]. Some important assumptions are negligible line losses [14], adjustable PQ injections located at all nodes [10], [15], or purely resistive network [16].

As dealing with the original nonlinear power systems is difficult, many existing control works rely on linearized systems. An important linearization is linearized power flow called *LinDistFlow* [17]. Authors in [10] utilizes the *LinDistFlow* model to provide a distributed feedback control algorithm for solving optimal voltage control problem and to identify the limitations of non-incremental control. Further, two incremental voltage control schemes have also been presented based on *LinDistFlow* model. The accuracy of *LinDistFlow* model is highly dependent upon the base-load point where the first-order approximation is done. Once we move away from the base-load, the error in voltage magnitude estimation increases [18]. To elevate this, authors in [19] presented the linear model of power flow is obtained based on a topology estimator. The method requires repeated voltage sensitivity computation and online parameter estimation. Recently, [20], which is a departure from the first-order approximation based linearization, uses Input Convex Neural Networks to obtain the convex power flow formulation for voltage control purposes. However, the convex formulation proposed in [20] is determin-

This work is supported by NTU SUG, AcRF TIER 1 2019-T1-001-119 (RG 79/19, EMA & NRF EMA-EP004-EKJGC-0003.

⋆Corresponding Author; † Contributed equally.

Authors are with School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore. *pare0001, yu006, hunghtd@ntu.edu.sg*

istic and may not be directly suitable for voltage control under load uncertainties. In short, the limitations of model-based linearizations under uncertain load motivates the development of a method which is more applicable under load variations.

In this paper, we develop an approximated linear relationships between the node voltage magnitudes and loads via GP learning. The proposed linearization is more accurate over load subspace (*subspace-wise* approximation) and can handle multiple types of uncertainties (*non-parametric* nature). These features mitigate the issue of error under load variation, present in model-based linearization like *LinDistFlow* [21].

Leveraging the resulting GPR-based *Voltage-Power* linear relationship, we then propose optimal voltage control algorithms which seek for the minimum control efforts, such as power injection changes, that sufficiently bring the system voltages to the acceptable level. The proposed optimal control framework does not rely on any strong assumptions on the networks as the above, and thus being applicable to more general power systems. Using GPR-based *Voltage-Power* linearization, the least squares technique can be deployed to provide explicit forms of the required optimal control efforts. These analytic forms offer a great computational benefit in voltage control and thus lending themselves to time-sensitive tasks, such as online voltage regulation as well as emergency controls. Moreover, GPR-based *Voltage-Power* linear relationship is valid over a subspace of the inputs, the proposed optimal control can work well in a wide range of power injections, instead of confining within a small neighbourhood of the linearized point as in other linearization-based control methods.

More importantly, unlike existing optimal voltage control algorithms with fixed voltage references, the proposed optimal framework can also optimize the voltage reference values that lead to the smallest voltage error deviations, from such reference values, with minimum control efforts. The proposed voltage control framework is then extended to handle injection uncertainties. Voltage variation bounds are derived to give an estimation of the resulting system voltages under uncertainties. Such variations bound will be used duly in the voltage control to ensure that all nodal voltage magnitudes will lie in the acceptable range while the power injections may vary within some forecasted sets.

For large-scale systems that may be equipped with limited communications, distributed voltage control alternative plays an important role. We propose a novel network partition technique that identifies the most effective local pilot buses as well as their neighbourhood members. The partition technique is developed based on the concept of Effective Voltage Control Source (EVCS) which are the most relevant, effective nodes for voltage control. More specifically, the nodes corresponding to which we have non-zero entries in optimal load vector, are considered for EVCS. These EVCS node indices are then used to perform network partition using clustering algorithm.

The main contributions of the paper can be summarised as:

- Employing a novel linear *Voltage-Power* relationship for distribution system voltage control that is built using GPR and applicable to a wide range of operating points.
- Proposing optimal voltage control algorithms for centralized and distributed schemes with closed-form solution under different controllable conditions. Both control efforts and voltage references are optimized and provided with analytical solutions.
- Proposing robust optimal control algorithm for load uncertainties. Possible voltage variations are bounded, guaranteeing compliance with the voltage regulation limits.
- Formulating an optimization problem and its relaxation for obtaining optimal Effective Voltage Control Source node set and network partition.

II. Linear Voltage-Power Relationship

In this section, we present the Gaussian process learning-based linear approximation of *Voltage-Power* relationship. The Taylor series-based linearization methods attempt to obtain the tangent plane of the *Voltage-Power* manifold at a solution point. This leads to a higher error with variation in the loading point. However, the loading variation is not a limitation in the proposed linearization using GP. At first, we briefly describe the GP regression [22] as a learning tool, and then we present the linearization of *Voltage-Power* relationship using GP.

A. Gaussian Process Regression (GPR)

The GPR is a Bayesian framework based modeling technique that allows us to incorporate a prior understanding of data in the models [22]. The GPR is interpretable, meaning that the user can explain the learning level's causality and effects up to a certain extent. These two features, prior and interpretability, make GPR very useful for learning the physical models such as power flow relationships. In the power system applications, GPR is used for wind power forecast [23], small-signal stability [24], demand forecast [25], and probabilistic optimal power flow [26].

A general GPR— over training data-set $\{\mathbf{x}^i, \hat{y}(\mathbf{x}^i) | i = 1 \dots \Omega\}$ where, $\hat{y}(\mathbf{x}^i)$ is observed function value at input $\mathbf{x}^i \in \mathbb{R}^d$— is given as [22]:

$$\hat{y}(\mathbf{x}^i) = y(\mathbf{x}^i) + \varepsilon^i, \quad i = 1 \dots N. \tag{1}$$

In the present work, vector $\mathbf{x}$ is the bus load vector, while $\hat{y}$ is the node voltage magnitude of any node. In (1), ε^i is independent and identically distributed noise variable with zero mean and σ_ε^2 variance. Details of GP basics and explicit expressions of means and variance of predicted posterior distribution can be found in [22]. Importantly, GPR approximates the input-output relationship via numerical observations. The noise variable can also be interpreted as combined error due to measurement noises, system parameter uncertainties, state estimation errors, etc. This gives the GPR some degrees of robustness against the error in training data.

To find the mapping between input and output, GPR relies on covariance or kernel function $k(\mathbf{x}^i, \mathbf{x}^j)$. The covariance function controls the accuracy and complexity of the resulting approximation of a function. For this work, we use linear covariance function $k_{LN}(\cdot, \cdot)$ as our target is to obtain linear *Voltage-Power* relationship for steady-state voltage control. The linear kernel allows us to obtain the least complexity relationship, suitable for finding optimal injections or control rules under load uncertainty. There are multiple methods which use the *kernel trick* [22] to perform the regression in finite dimension. Below we briefly discuss the difference between kernel regression (KR) and GPR.

The KR generally refers to the methods that use kernel to learn input-output relationship. One of the prominent KR method is kernel ridge regression [27]. Both GPR and KR use the kernel (or called covariance function generally in case of GP [22]) to establish the relationship between different input data-points. However, there are some fundamental differences in the way of learning the function for regression. The GPR learns the kernel *hyperparameters* via optimizing the *maximum a posteriori* (MAP) estimation. The MAP is mean of the posterior probability distribution of *hyperparameters* for a given training input set [22]. Contrary to this, the KR uses grid search on a mean-squared error loss function. Furthermore, the GPR is based on Bayesian theory and a generative model while KR is not. This means that GPR can generate posterior samples providing ability to predict the output at different input points not included in the dataset. Moreover, generalised regularization procedure can also be compared with GPR. The main difference is that GP provides uncertainty information. It calculates marginal likelihood which is useful for *hyperparameter* calculation, and direct model comparison. Interested readers can refer to chapter 6 of [22] for comparative analysis between GP and other methods.

In the following subsection, we build a linear relationship using the GPR. We first describe the required training set and then present linear expression representing the *Voltage-Power* relationship based on the predicted mean expression [22].

B. Linear Voltage-Power Relationship

In GP learning, the relationship between input and output is learned through covariance function. The complexity and accuracy of approximated function directly depends upon the selection of covariance function. As discussed above, we use linear covariance function $k_{LN}(\cdot,\cdot)$ to obtain linear function approximation. First, we obtain a training set $\{s^i\,,\hat{V}_j^i\,|\,i=1\ldots\Omega\}$ for an individual node j which belongs to the node set $\mathcal{N}$. Here, vector $s=[p\,;\,q]\in\mathbb{R}^{2n}$ is collection of real p_j and reactive power q_j load at all the nodes, with n being equal to number of load buses. The $\hat{V}_j^i$ is voltage magnitude (referred as voltage hereafter) solution obtained from ACPF [28] at load s^i. The collection of all load vectors for training is *design matrix* $S=[s^1\ldots s^\Omega]$ of size $2n\times\Omega$. The idea is to learn the voltage relationship in a load subspace. Formally, we define the $2n$-dimensional load subspace $\mathcal{L}=\big\{[p_j;q_j]\big|\,p_j\in[p_j^o\pm\delta p_j^o];\,q_j\in[q_j^o\pm\delta q_j^o];\forall j=1\ldots n\big\}$, with δ being level of uncertainty. The value $\delta=0.5$ means that load subspace is constructed with $\pm 50\%$ uncertainty in base-load of each node. Here, we consider that power flow Jacobian is invertible (well-conditioned) at all the points within the considered load variation range $\mathcal{L}$. Furthermore, for this steady-state study, we assume the network topology and line parameters are fixed. Thus, no sudden disruption in connections is considered.

Now, we employ GP to perform a linear approximation of the individual node voltage as a function of the bus load vector. Using the log-marginal likelihood estimator with GP, we obtain the mean prediction of voltage as [22]:

$$V_j(s)=k_{LN}(S,s)^T\alpha_j \tag{2}$$

$$\text{where}\quad \begin{aligned} k_{LN}(S,s)&=\tau^2\left(\gamma\mathbf{1}+S^Ts/l^2\right),\\ \alpha_j&=[k_{LN}(S,S)+\sigma_\epsilon^2 I]^{-1}\hat{V}_j. \end{aligned} \tag{3}$$

Here τ,γ,l are hyper-parameters. The voltage linearly depends on variable load vector s in (2), as the α_j is constant vector upon learning and depends on training set only. The relationship (2) can be simplified as:

$$V_j(s)=m_js+c_j'. \tag{4}$$

Here, the intercept $c_j'=\tau^2\gamma[\alpha_j^T\mathbf{1}]$ and voltage sensitivity row vector $m_j=\{\tau^2\alpha_j^TS^T\}/l^2\in\mathbb{R}^{1\times 2n}$. Now, collecting all n linear functions of voltage vector $V(s)=V$, the linear system of equation relating voltage vector with load vector is

$$V=M\,s+C. \tag{5}$$

The matrix $M\in\mathbb{R}^{n\times 2n}$ can be interpreted as the slope of voltage plane, which is valid for the learning subspace $\mathcal{L}$. Now, let s_o be the base load and V_o be the corresponding node voltage vector. Then, the voltage change ΔV can be derived as a function of load change $\Delta s=s-s_o$ as

$$\begin{aligned} \Delta V&=V-V_o\\ &=M\Delta s \end{aligned} \tag{6}$$

The proposed GP learning method relies on learning the *Voltage-Power* relationship only based on the training set $\{s^i\,,\hat{V}_j^i\,|\,i=1\ldots\Omega\}$. This training data set does not have explicit information about the network graph or the parameters. Thus, the proposed method can be employed to learn the *Voltage-Power* relationship using real state observations of an unknown network graph.

III. Optimal Voltage Control Algorithms

This section presents a number of optimal voltage control algorithms for both centralized and distributed control schemes. These control algorithms are extended to handle net load or injection uncertainties due to DGs and EVs.

A. Centralized voltage control

We introduce a voltage control algorithm that helps— to regulate the entire network's voltages— to achieve a reference voltage profile. We formulate the voltage control problem as:

$$\begin{aligned} \min_{V}\quad & \|V-V_0\|^2\\ \text{subject to}\quad & s=s(\theta,V) \end{aligned} \tag{7}$$

In this optimization framework, the control objective is to minimize the voltage deviation from the nominal voltage V_0. The constraint is the power flow relation to which the network respects. In this work, we limit ourselves to the voltage derivation measured using the Euclidean norm. So the objective is to minimize $\|V-V_0\|_2^2$. For simplicity, we omit the subscript 2 hereafter.

The problem (7) is widely used in literature [29]. One challenge associated with this optimization is the nonlinear and non-convex power flow constraints that makes solving this control problem inefficient. Therefore, the linear relationship (5) is employed to approximate the nonlinear power flow constraints in this work. Moreover, a common practice is to implement the voltage control from a base steady-state solution or typically the current operating point $\{V_\star,\, s_\star\}$. Let, $\Delta s = s - s_\star$ as the control efforts used to regulate the network voltage. The linearized power flow starting with the current operating point $\{V_\star,\, s_\star\}$ can be rewritten as follow:

$$\begin{aligned} V &= M(s_\star + \Delta s) + C \\ &= M\Delta s + V_\star. \end{aligned} \tag{8}$$

If the matrix M is square and invertible, solving (8) for the control effort Δs that leads to some desired voltage V becomes simple. However, the linearized matrix M might not be necessarily square and invertible in general. For a rectangular matrix M, the system can be either underdetermined or overdetermined, depending on the size of the matrix M. For the underdetermined case, there are fewer rows than columns, so one has fewer equations than unknowns. As a result, there are infinite solutions satisfying the system of equations. Rouché–Capelli theorem presented in [30] formally describes this case as the rank of the coefficient matrix is less than the rank of its augmented matrix. On the other hand, the system becomes overdetermined when there are more rows than columns or one has more equations than unknowns. The exact solution to an overdetermined system might not exist, so the common practice is to find the best approximation, typically using least-squares estimation.

1) Overdetermined system with minimum control efforts: In practice, the voltage control problem may exhibit redundant control freedom when multiple voltage control sources, such as distributed generators, are connected to the same voltage bus [31]. Then all such control sources will control the same nodal voltage. Without a proper coordination, the control efforts from such sources may cancel out each other, for example, one source generates power and the others consume power locally. This situation leads to a waste of using control efforts. Moreover, the nodal voltage is determined by the effective or net combination of all injections at the node, not by the individual injection of each control source. For the same net injection value, there are many combinations of the individual injections. As a result, there exit many solutions to the voltage control problem that leads to the same voltage control effect.

From a mathematical perspective, this multiple-solution phenomenon happens when the system is under-determined, which means there are fewer equations than variables that need to find. Rouché–Capelli Theorem presented in [30] formally describes this situation as follow. For an under-determined system, the rank of coefficient matrix of is less than the rank of its augmented matrix, so there exist infinite solutions. Due to the existence of multiple solutions to the voltage control, the optimal voltage control will need to extend to minimize not only the voltage deviation errors ΔV but also the control efforts $\|\Delta s\|$. The resulting optimal control solution corresponds to the least control effort among all multiple solutions that have the same voltage regulation effect. Also, it is worth noting that the linearized matrix M might not be necessarily square and invertible in general.

Particularly, the underdetermined system, as discussed above, will arise when matrix M is a rectangular matrix since the number of equations is less than the number of variables when redundant control freedom exists. In this work, the size matrix M is $n \times 2n$ where n is the number of bus nodes, which equals to the number of equations, and the variables matrix containing active power and reactive power of each node with size $2n \times 1$. Under this condition, matrix M is not invertible, and infinite solutions will be there for the optimization problem (7).

However, in under-determined system, one scenario that may show up is that the matrix M is not full column rank, which means the columns of matrix M is not fully independent, so that $M^T M$ is also not invertible. Usually, in power system, this scenario represents the condition that the performance of some control units can be fully replaced by the other units.

Thus, with admitting $M^T M$ may also not invertible, to achieve minimum voltage deviation $\min \Delta V$ with the minimal control effort $\min \Delta s$, combining equation (8), the original optimization problem (7) now becomes:

$$\begin{aligned} &\min_{\Delta s} \quad \|M\Delta s + V_\star - V_0\|^2 \ \&\ \min \|\Delta s\|^2 \\ &\text{subject to} \quad V = M\Delta s + V_\star \end{aligned} \tag{9}$$

This double-aim optimization problem can be practically achieved by using pseudo-inverse matrix, which also overcome the un-invertible problem of $M^T M$. Let M^+ be the pseudo-inverse of M. Then, the optimal solution that minimizes both voltage deviation ΔV and control effort Δs is given below:

$$\Delta s_{opt} = M^+(V_0 - V_\star). \tag{10}$$

The mathematical proof is presented in reference [32]. Note that this minimum control effort $\Delta s_{opt} = M^+(V_0 - V_\star)$ is the unique vector with smallest Euclidean norm that minimizes the voltage deviation $\|V - V_0\|$.

2) Underdetermined system with analytical form of optimal control effort: One simple practical solution to the redundant control freedom problem is removing unnecessary voltage control sources, such as remote or expensive devices. As a result, one can make the matrix M a full column rank matrix, so the matrix $M^T M$ will be invertible. Thus, the well-known least square estimation can be applied to obtain an explicit optimal control solution:

$$\Delta s_{opt} = (M^T M)^{-1} M^T\,(V_0 - V_\star). \tag{11}$$

This close and concise form (11) of optimal solution offers a great benefit in designing the optimal power injections, without a need for searching for such optimal values numerically. With less computational intensity, the proposed optimal voltage control lends itself to time-sensitive tasks, such as online regulation and emergency controls.

Note that the explicit form (11) is also the optimal control solution to an overdetermined system. Opposite to the under-determined case discussed in section III-A1, an over-determined system has more equations than unknowns. This scenario mostly happens to a distribution system with insufficient voltage control equipment, typically reactive power compensators and generators. This situation can be caused by the presence of too many uncontrollable power injections from renewable sources and EVs.

3) Optimal voltage reference: Another important factor in optimal voltage controls is how to set the voltage reference. In most cases, one often assumes that the voltage reference is given so the main task is to find the optimal power injections to minimize the voltage mismatch. However, the resulting voltage deviations, due to the control action and the control efforts, depend significantly on the voltage reference. More specifically, based on the current voltage profile, a proper voltage reference can help reduce both the control effort and the voltage deviation. This section presents a systematic approach to optimize such a voltage reference.

Assume matrix M is full column rank and has the size of $m \times n$, $m > n$. For this rectangular matrix M, let B be an orthogonal matrix such that $M = BR$ where $R = [R_1; 0]_{m\times n}$. R_1 is a upper triangular and non-singular matrix. Since B is an orthogonal matrix, we have that

$$\begin{aligned} \|M\Delta s - \Delta V_\star\|^2 &= \|B^T(M\Delta s - \Delta V_\star)\|^2 \\ &= \|R\Delta s - B^T\Delta V_\star\|^2. \end{aligned} \tag{12}$$

Furthermore, we split the matrix $B^T = [B^{Tu}; B^{Tl}]$ where B^{Tu} consists of the first n rows of B^T and B^{Tl} consists of the rest $m-n$ rows of B^T. Then, $B^T\Delta V_\star = [B^{Tu}\Delta V_\star; B^{Tl}\Delta V_\star] = [\Delta V_\star^u; \Delta V_\star^l]$ where the upper part $\Delta V_\star^u$ has the size $n \times 1$ and the lower part $\Delta V_\star^l$ has the size $(m-n)\times 1$. The objective function (12) becomes

$$\begin{aligned} \|M\Delta s - \Delta V_\star\|^2 &= \|R\Delta s - B^T\Delta V_\star\|^2 \\ &= \|R_1\Delta s - \Delta V_\star^u\|^2 + \|\Delta V_\star^l\|^2. \end{aligned} \tag{13}$$

Now, (13) provides insights into the least square solution and helps improve such an estimation. The term $\|R_1\Delta s - \Delta V_\star^u\|^2$ can be minimized easily as R_1 is non-singular and upper triangular. The corresponding solution Δs is the least square estimation (11). The objective function can be further minimized by controlling the last term $\|\Delta V_\star^l\|^2 = \|B^{Tl}(V_0 - V_\star)\|^2$ with the voltage reference V_0. An optimization formula for finding the voltage reference which minimizes last term thus further minimizing the objective (12) is

$$\begin{aligned} \min_{V_0} \quad & \|B^{Tl}(V_0 - V_\star)\|^2 \\ \text{subject to} \quad & V_m \le V_0 \le V_M. \end{aligned} \tag{14}$$

where V_m and V_M are the lower bound and upper bound of the feasible voltage level. The optimization (14) can be solved using quadratic programming. A simplification can be made by constructing the voltage reference V_0 based on the current voltage $V_\star$, in particular, by keeping the acceptable nodal voltages and only optimizing the new reference values of the non-compliant buses.

B. Voltage Control with Uncertainties

In the presence of uncontrollable power injections, such as those of renewable sources or electric vehicles, the above optimal voltage optimization needs to be revised and the explicit form of the optimal solution (11) can be used directly. In this case, the power injection vector can be written as $\Delta s = [\Delta s_c; \Delta s_{uc}]$ where Δs_c denotes the injection variations on the controllable nodes and Δs_{uc} denotes the uncertain injection variations. The linearized power flow equation (16) can be reformulated as:

$$V = M_c\Delta s_c + M_{uc}\Delta s_{uc} + V_\star \tag{15}$$

where $M_c = [M_{cc}; M_{uc}]$, $M_{uc} = [M_{uc}; M_{uu}]$ are the corresponding linearized coefficients. The uncontrollable power injection Δs_{uc}, if without proper prior knowledge, may pose additional challenge to the voltage control problem.

1) Optimal controllable values: In this work, we consider the uncertain power injection variations Δs_{uc} that fluctuate around zero, which for simplicity is equal to the expectation value, with a variation ϵ bounded by a constant $\alpha > 0$. With such assumptions on the uncontrollable injections, the optimal voltage control problem can be formulated as:

$$\begin{aligned} \min_{V} \quad & \Delta V = \|V - V_0\| \\ \text{subject to} \quad & V = M_c\Delta s_c + M_{uc}\Delta s_{uc} + V_\star \\ & \|\Delta s_{uc}\| \le \alpha \end{aligned} \tag{16}$$

Taking expectation of the voltage deviation within the objective function, yields:

$$\begin{aligned} &\mathbb{E}[M_c\Delta s_c + M_{uc}\Delta s_{uc} + V_\star - V_0] \\ &= M_c\,\mathbb{E}[\Delta s_c] + M_{uc}\,\mathbb{E}[\Delta s_{uc}] + V_\star - V_0 \\ &= M_c\,\mathbb{E}[\Delta s_c] + V_\star - V_0. \end{aligned} \tag{17}$$

Assume M_c is full column rank, as discussed above, using the least square estimation method again, the optimal controllable power injection can be given as:

$$\Delta s_{c,\,opt} = (M_c^T M_c)^{-1} M_c^T\,(V_0 - V_\star). \tag{18}$$

However, if M_c is not full column rank, the optimal control effort can be given as:

$$\Delta s_{c,\,opt} = M_c^{+}(V_0 - V_\star). \tag{19}$$

2) Bounding the voltage variation: With the calculated optimal values of controllable inputs, in order to make sure a certain range of uncertainties will not cause the voltages violate the reference values, the voltages vary range need to be considered.

Below the bound of voltage variation when uncertainties exist is presented. The power injections with the uncertainties can be expressed as:

$$s_{uc} = \mathbb{E}[s_{uc}] + \Delta s_{uc} = (1+\epsilon)\mathbb{E}[s_{uc}] \tag{20}$$

The uncontrollable power variation $\Delta s_{uc} = \epsilon\mathbb{E}[s_{uc}]$ will lead to the following voltage deviation

$$\Delta V_u = \epsilon M_{uc}\mathbb{E}[s_{uc}] \tag{21}$$

Note that this voltage deviation ΔV_u is compared with the case when the uncontrollable power injection is fixed at $\mathbb{E}[s_{uc}]$ and the voltage V_e. This can be bounded as follow.

$$\|\Delta V_u\| \le \epsilon\|M_{uc}\|\|\mathbb{E}[s_{uc}]\| \tag{22}$$

Comparing equation (21) with the expected voltage, the relative variation range is:

$$B_{V_{eu}} = \frac{\|\Delta V_u\|}{\|V_e\|} = \frac{\|\epsilon M_{uc}\mathbb{E}[s_{uc}]\|}{\|M_c s_c + M_{uc}\mathbb{E}[s_{uc}] + C\|} \tag{23}$$

With indexes (22) and (23), we can update the reference value accordingly during the calculation of optimal inputs, so that even there have voltage variations, no node voltage falling below the requirement value can be promised.

The bound prediction with indexes (22) and (23) is relatively tight and can be used directly. On one hand, for a tighter bound, more details on the uncertainty properties can be included. On the other hand, these indexes are also influenced by the error bound of the linearized matrix M. In the case that has high voltage requirement resistant to the uncertainties, the bound prediction can be magnitude by considering M estimation bound or bigger uncertainty range ϵ in order to get extra control margin.

We propose the implementation procedures for the optimal control with uncertainties are given below:

Step 1: Check for node voltage limit violations.
Step 2: Acquire information of uncertainties, including location, type, expected values $\mathbb{E}[s_{uc}]$ and variation ϵ. Calculate voltage variation bound by (22) or (23).
Step 3: Calculate the optimal voltage reference V_0 from (14) with expectation values $\mathbb{E}[s_{uc}]$ of uncertainties.
Step 4: Update voltage reference V_0 in equation (18) or (19) according to calculated bound in earlier step.
Step 5: Calculate optimal controllable power injections $\Delta s_{c,\,opt}$ with equation (18) or (19) using $\mathbb{E}[s_{uc}]$.
Step 6: Plug in $\Delta s_{c,\,opt}$ and uncertainties in current operating point. Run power flow.
Step 7: Check if all nodes voltage compliant with the reference value.

C. Distributed Voltage Control

The proposed method can be localized and applied on distributed network. Different criterion can be utilized to cluster the system nodes such as dividing by area or by hierarchy. The network partition method is presented later in section IV.

Different from the frequency, the voltages manifest themselves as local variables, i.e., the local change in the injection will mostly result in the variation in the neighbourhood voltages. This effect will attenuate quickly when propagating through the system to further areas away from the local change. Thus, a distributed control scheme makes a perfect sense for the voltages. Also, the radial structure of the distribution network with low connectivity allows for clustering the entire network into several sub-systems where the voltage control can be done locally. If a disturbance happens in one sub-system, most likely the voltage variation can be fixed by just regulating the power injections within such a sub-system.

Assume disturbance happens at area i, based on the properties of the linearized matrix M, the subsystem can easily be decoupled from the whole system. Reformulate the system voltage control problem with block matrices as below.

$$\begin{bmatrix} V_I \\ V_{II} \\ \vdots \\ V_K \end{bmatrix} = \begin{bmatrix} M_{II} & M_{I\,II} & \cdots & M_{IK} \\ M_{II\,I} & M_{II\,II} & \cdots & M_{IIK} \\ \vdots & \vdots & \cdots & \vdots \\ M_{KI} & M_{KII} & \cdots & M_{KK} \end{bmatrix} \begin{bmatrix} s_I \\ s_{II} \\ \vdots \\ s_K \end{bmatrix} \tag{24}$$

here, the Roman numerals subscripts from I to K represent sub-areas of the distributed network. Take out all the variables and corresponding parameters in area i and the areas that influenced by the disturbance, rearrange the matrix and calculate the optimization problem following the solution in (10)-(11).

Different from the centralized voltage control scheme, the distributed counterpart may involve multiple trial and error. The injection computed using LSE within the sub-system will be applied to the network, which will result in a new voltage profile. This voltage profile may not be close enough to the reference, mostly due to lack of control degree and incomplete information of the network, so a new power injection will be produced. This iteration will continue till the voltage profile is accepted. Thus, the procedures in for distributed localization optimal control are given below

Step 1: Check whether any node voltage goes below reference value.
Step 2: Get information of system clustering states. Identify the node set $\mathcal{B}_k$ and make sure how many subsystem has voltage violations.
Step 3: Take out corresponding variables and parameters and acquire new matrix equation following (6) for each node set $\mathcal{B}_k$ that contains voltage violations.
Step 4: Calculate the optimal voltage reference V_0 through equation (14) for each subsystem that contains voltage violations.
Step 5: Calculate optimal power injections Δs_{opt} with equation (18) or (19) for these node set $\mathcal{B}_k$.
Step 6: Plug in optimal power injections Δs_{opt} and uncertainties on current operating point. Run power flow for whole system.

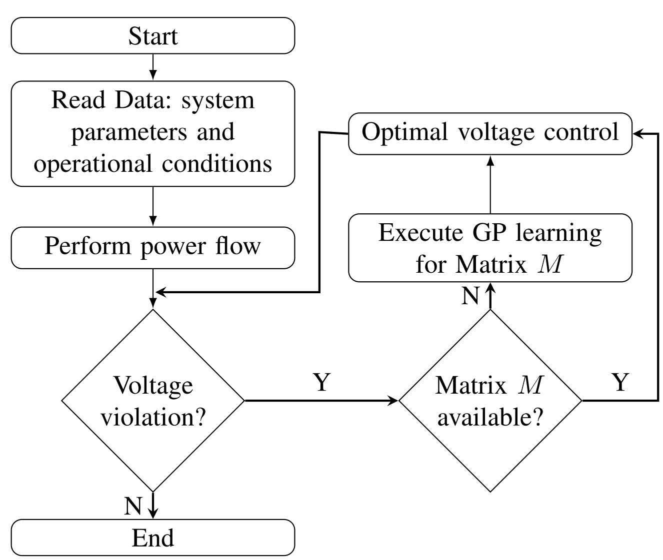

Fig. 1. Schematic flowchart of optimal control algorithm

Step 7: Check whether any node voltage goes below reference value. Repeat the iterations if there still has voltage violation.

To summarize, the proposed approach can be applied to multiple scenarios with exploiting the GP-acquired sensitivity matrix M. The schematic flowchart is shown in Fig. 1 and corresponding optimal control scheme under different conditions follows the steps described in III-A, III-C, and III-B. In the flowchart 1, the loop will stop when all the system voltages are compliant to the voltage regulation standards. If multiple optimal voltage control attempts are made but the voltage violation cannot be corrected, one can consider to recompute the matrix M to improve the approximation. A time window-based continuous/online control can be developed where one can set the time duration wherein the voltage check can be carried out. Optimal voltage control will be implemented within each time window.

IV. EVCS Ranking and Network Partition

In transmission systems, multiple voltage-controlled buses, called PV buses, are present as generator buses. These PV buses give the system operator multiple options to control the voltage. Unlike this, a low voltage distribution network generally has one PV or Slack bus— at point of coupling with medium or high voltage network. As we move away from the slack bus in radially arranged networks, the slack bus's ability to control system state decreases. To overcome this, the farthest bus— from slack bus— is selected as the voltage controlled bus as it directly shifts the voltage curve of network up-words, and reduce violation [33]. Nevertheless, the selection is criterion is not defined and analyzed in detail. Further, it is not always straightforward to define the electrically farthest bus, particularly with local DGs.

In the following, we propose an optimization framework for finding the most relevant power injections EVCS for voltage control purposes. Now, the linear Voltage-Power relationship in (6) relates the change in the voltage ΔV with the change in the power injection or load Δs. The idea is to place the voltage controller on a node which has *maximum capacity of control*. The control capacity of a node injection means that by how much it can shift the voltage vector. Thus, if the injection change at a particular node set leads to a higher value of ΔV, then those nodes have a higher capacity to keep the node voltage vector within limits. Let's assume we want to install three voltage controlling sources. This means that we need to find three nodes— as a set— which maximizes voltage control capacity. This can be done by finding the nodes which maximize the norm of voltage deviation vector. The matrix M appearing the linear relationship $\Delta V = M\Delta s$ can be considered as the sensitivity matrix between the voltage change and the power injection change. Therefore, we formulate the problem below with the objective of maximization of the voltage deviation while keeping the number of EVCS less than a fixed number using cardinality constraint $Card \leq t$.

$$\begin{aligned} \max_{\Delta s} \quad & \Delta V^T \Delta V = \Delta s^T M^T M \Delta s \\ \text{subject to} \quad & \|\Delta s\| = 1 \\ & \text{Card}(\Delta s) \leq t. \end{aligned} \tag{25}$$

As discussed above, the problem (25) searches for the optimal power injection that leads to the maximum load voltage changes. The length of this power injection change is normalized to the unit. The second constraint is to bound the maximum number of non-zero elements in the power injection vector, denoted by Card$(\cdot)$, by a given integer $t > 0$. This integer represents the number of considered relevant injections.

The problem (25) is non-convex. Using the lifting procedure for semidefinite relaxation, we have

$$\begin{aligned} \max_{X} \quad & \text{Tr}(MX) \\ \text{subject to} \quad & \text{Tr}(X) = 1 \\ & \text{Card}(X) \leq t \\ & X \succeq 0;\ \text{Rank}(X) = 1 \end{aligned} \tag{26}$$

Here, $X = \Delta s \Delta s^T$. The cardinality constraint can be replaced by a weaker but convex $\mathbf{1}^T|X|\mathbf{1} \leq t$ with $|X|$ contains all absolute values of X's entries. Now, by ignoring the rank-one constraint, resultant convex formulation is

$$\begin{aligned} \max_{X} \quad & \text{Tr}(MX) \\ \text{subject to} \quad & \text{Tr}(X) = 1 \\ & \mathbf{1}^T|X|\mathbf{1} \leq t;\ X \succeq 0 \end{aligned} \tag{27}$$

The optimal Δs consists of at most t power injections that are the most relevant and effective control efforts in the network for voltage regulation purposes. This t-length set of injections will be used for network partition.

A. Network Partition

The task of network clustering has been proposed in the recent past using various methods [34]. Our idea here is to introduce the network partitioning based on the slope of voltage plane M, estimated via GP learning. In particular, we propose to use the sensitivity of voltage with respect to the EVCS bus for clustering. This is different than general idea of breaking system into parts using complete system information. The proposed partition allows us to identify the nodes where voltage can be controlled by high priority EVCS. The main idea of using the voltage slope or sensitivity for clustering is to identify the node voltages having similar sensitivities, with respect to EVCS node injection change. The similarity is measured with respect to EVCS node's own sensitivity of voltage change. To perform the clustering based on sensitivity, we have used $k - means$ with centroids fixed at the EVCS nodes, obtained by solving (27).

Let, we obtain optimal Δs, having t non-zero entries, by solving (27). Now, as described above, we will use the sub-matrix $A \in \mathbb{R}^{n \times t}$, having t columns of M corresponding to the non-zero entries in Δs. Now our target is to divide the network into $t+1$ sub-parts such that each part is effectively controlled by t EVCS buses and one slack bus. Below we present the step-by-step method to perform the network partition.

Step 1: Solve optimization problem (27) and obtain optimal Δs having t non-zero entries.

Step 2: Construct a sub-matrix A having t columns of M related to non-zero entries in Δs.

Step 3: Run clustering algorithm to cluster A into $t+1$ parts, and assign the corresponding EVCS bus to each cluster.

Upon completion of this procedure, we obtain $t+1$ index sets as $\mathcal{B}_k$ where $k = 1 \dots t+1$. Here, in the $t+1$-th set are the slack bus controlled nodes. In the next section, we present centralized and distributed voltage control mechanisms based on GP based linearization. It is important to note here that we do not focus on any specific type of clustering algorithm. Our idea is based on using EVCS bus related sensitivities of voltage to find clusters.

In the present work, we have employed $k - means$ as it is a well known standardised clustering method with fixed centroids. The idea is general and open to employ any state-of-art clustering algorithms such as Gaussian mixture model (GMM) [35]. The effect of different clustering algorithms on network partitions will be analyzed in future works.

V. Numerical Results and Discussions

The simulations have been presented on the IEEE 33-Bus radial distribution system [28] and a modified IEEE 69-Bus system [36] with all nodes having non-zero load. For the GP learning, we use the well established GPML Toolbox [37] with MATLAB 2020a. The power flow calculations for learning data-set for Voltage-Power relationship and voltage control problem are performed using MATPOWER *runpf* [28]. We use the Newton-Rapson power flow option in *runpf* as it is well established method and used by various works. We first show the accuracy of the linearization, particularly when loading point is away from base-load. Further, we show the network partition and associated pilot bus. At the end we present the centralised and decentralised voltage control results.

A. Linearization Performance

The GP-based linear *Voltage-Power* relationship is valid for load subspace $\mathcal{L}$ as discussed in II-B. The model-based linearization methods perform relatively poorly when we move away from the base-point [18]. The difference in performance is captured in Fig. 2. The results show that at load 1.5 times of base load, the proposed method outperforms other linearization methods proposed in [18] and [21] significantly. This fact, that linearization is loading dependent, has also been reported and analyzed by authors in [18] . With a training set of 300 samples, $N = 300$, we perform learning for different size of load uncertainty set. Further, we compare the linearization (5) with ACPF solution [28] for 10^4 samples. The maximum mean absolute error (MAE), among all node voltage magnitudes, is recorded as 2.36×10^{-6}, 5.95×10^{-6} and 5.65×10^{-5} for uncertain spaces of δ being 10%, 20% and 50% respectively.

Fig. 3 shows the comparative results of maximum approximation error for *LinDistFlow* [21], method proposed in [18] and purposed GP-Based approximation for IEEE 33-Bus system. The histograms clearly indicate that proposed method is approximately 100 times more accurate than other two methods. In case of the model-based linearization methods, the error histograms in Fig. 3 resemble the normal distribution because error is a function of load in these methods. The

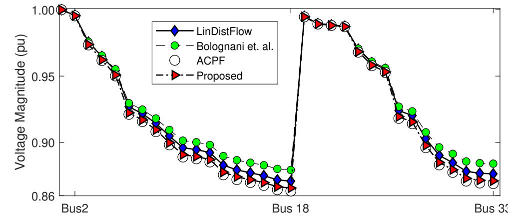

Fig. 2. Voltage comparison between ACPF, *LinDistFlow* [21], model-based linearization Bolognani et.al. [18] and proposed linearization at $1.5\times$ base-load in 33-Bus system.

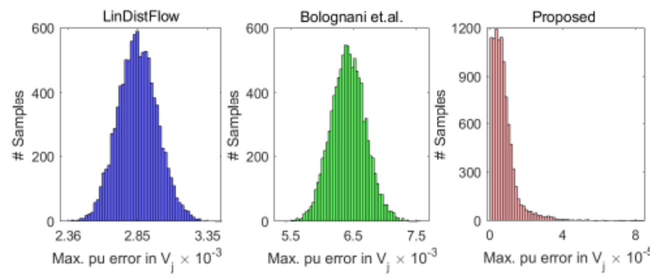

Fig. 3. Maximum absolute voltage magnitude error with *LinDistFlow* [21], Bolognani et.al. [18] and proposed linearization for IEEE 33-Bus system [28] using 10^4 samples. Bus loads follow normal distribution with base-load (p^o, q^o) as mean and standard deviation of 10% of base-load. GP Training data-set has 350 samples. The loading dependence of approximation error is clearly visible with normal distribution of error with [21] and [18].

proposed method's accuracy is not a function of system loading, therefore we obtain a one sided error histogram.

To show that proposed method works for different size systems, and effect of load on performance we present maximum error results for 69-Bus system [36], [38] in Fig.4. The error in *LinDistFlow* and method in [18] is one order of magnitude higher than that with the purposed method and load dependent as well. Another clear observation is that error has positive slope with loading of the system. This means that a voltage control method, for higher loading will not be as accurate as it will be for the lower. These results indicate that the GP-based linearization is suitable for obtaining linear *Voltage-Power* relationship, particularly under load uncertainty.

B. EVCS Ranking and Network Partition

As discussed in section IV, the controlled load variations at EVCS bus can greatly help to control the network voltage. In this part, we present the partitions of the 33-Bus network obtained via the step-by-step procedure explained in section IV along with associated EVCS bus. When we solve (27) with $t = 2$, we obtain optimal Δs with nonzero entries corresponding to node 18. Fig. 5 shows the results of partition when one EVCS source is placed at 18-th node. Further, using the partition procedure, we obtain one cluster where node 9 to node 18 are clubbed together for possible distributed control. The other bigger cluster is with the slack bus.

C. Voltage Control Performance

In this section, the proposed voltage control performance when system voltage falls below the reference level on IEEE 33-Bus system and modified IEEE 69-Bus system [36]. The low voltage conditions are created by increasing the base load taken from the corresponding test case in MATPOWER [28].

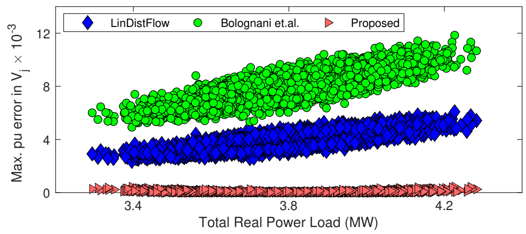

Fig. 4. Variation in the maximum absolute voltage magnitude error with load change for *LinDistFlow* [21], model-based linearization Bolognani et.al. [18] and proposed linearization for IEEE 69-Bus system [36], [38] using 10^4 samples. The learning and testing is performed with normal distribution having base-load (p^o, q^o) as mean and standard deviation of 10% of base-load, at all the nodes. Total 450 samples are used to learn the GP based linearization.

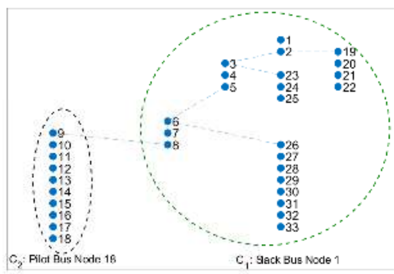

Fig. 5. IEEE 33-Bus system [28] partition with one EVCS at 18-th bus.

Table I shows the comparison with/ without full column rank matrix M under global centralized control. In case-I, there are total 10 violations, which means 10 node voltages go below 0.90 pu, and the minimum voltage is 0.879 pu in this case. In the case-II, there are total 13 node voltage violations, and the minimum voltage is 0.855 pu. In the table I, Min-V-P and Min-V-S denote the minimum node voltage after different control activities, Ratio-P gives the ratio with sum of the absolute values of power injection changes, which indicate the control efforts and the efficiency of the control activity. Further, Min-Npq is the minimum number of controllable components that system needs in order to bring the voltages back to the reference value, which also suggests the minimum column rank of matrix M, as discussed in III-A2.

Fig. 6 shows the voltage profiles with different voltage reference setpoints for IEEE 69-Bus system. It is clear that the proposed voltage control method is able to bring the system voltage back to the required level by adjusting the reference setpoints duly. In Fig. 6, the cyan curve ("Original Voltage" legend) depicts the uncontrolled, low voltage profile where several node voltages fall below 0.90 p.u. and thus violate the voltage regulation standards. The other three blue, orange, and red curves ("V-Ctr-Ref1", "V-Ctr-Ref2", "V-Ctr-Ref3" legends) plot the resulting system voltage profiles after implementing the optimal voltage control using three different voltage reference setpoints. Such setpoints are chosen to bring the system voltage levels above three lower bounds of 0.90 pu., 0.92 pu., and 0.94 pu. One can see that the proposed optimal control is successful in regulating the system voltages

TABLE I
CENTRALIZED CONTROL IN 33-BUS SYSTEM

	#Violated Bus	Min V	Min $V-P$	Min $V-S$	Ratio P	Min Npq
Case-I	9	0.879	0.909	0.906	1.225	11
Case-II	13	0.855	0.913	0.905	0.790	26

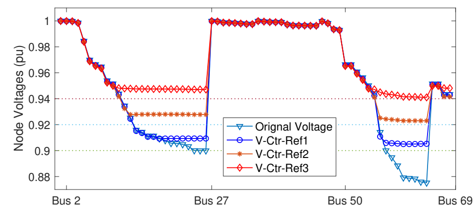

Fig. 6. System voltage profiles with different voltage reference setpoints in IEEE 69-Bus system

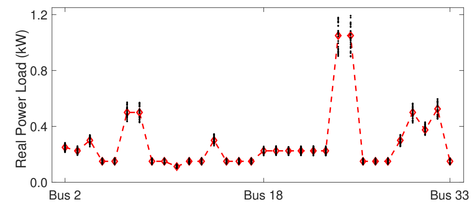

Fig. 7. Real power load with 30% uncertainties in 33-Bus system

following desired voltage requirements. Note that the control process of lifting the system voltages from the limit of 0.90 pu to the higher limit of 0.94 pu can be done in one single step of control or in several successive control steps. From the practical point of view, the latter fashion is recommended as the system voltage can be regulated gradually to avoid large jumps in the system operation. Another important point to note is that there are a number of good bus voltages, such as those around buses 26 to 55, are higher the voltage lower limits. The voltages at such buses almost remain the same during the voltage control process. One reason is that the voltage reference setpoints are selected wisely to focus on the low voltage buses and not to change the already good voltage levels. This helps reduce the unnecessary control actions. The localised voltage control effect is also manifested here as remote buses' voltages will not change much due to some local power injection controls [31].

The simulation results of the proposed method to handle uncertainties are shown in Fig. 7 and 8. Here, we assume the active power on all the load nodes contains a level of uncertainties and all the reactive power are controllable. This resembles condition where in distributed system, all nodes contain renewable generation or intermittent loads like EVs. Here the variation of the uncertainty is 30% showing in figure 7. According to the proposed indexes (22) and (23) in section III-B, the voltage reference needs to lift up 0.022 pu in order to tolerant the uncertainties. Fig. 8 compares the voltage curves with and without optimal control. These simulation results show that after deploying the proposed methods, the voltage profile has been taken back to above 0.90 pu. The vertical dots in Fig. 8, at each node, show the voltage variation range for 30% uncertainty. The maximum voltage variation is 0.016 pu, which stays within the calculated voltage bound index range.

Fig. 9 shows the voltage control results of distributed areas. The performances are compared with the global control. The 33-Bus system is partitioned following the Fig. 5 and following the nodes numbers into three sub-system. From the figure we can see the voltage profiles of distributed control and centralized global control are almost overlapped, which

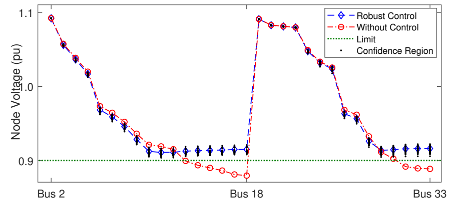

Fig. 8. Voltage control profiles with 30% uncertainties in 33-Bus system

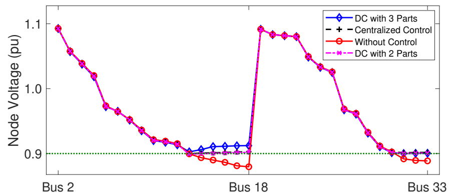

Fig. 9. Distributed control voltage profiles in 33-Bus system

means with the proposed clustering and distributed control can achieve almost the same performance with the global control. The distributed control for three subsystems achieves a higher value with the need of an extra iteration as one node fail to recover with only one round of control.

Now, we compare the computational burden between global control and distributed control. As the size of matrix M decreases, the time spent with distributed control is 43% less than the global control. And if the bad voltage nodes only exist in one sub-area, the time for distributed control only takes 12% of the time that global optimal voltage control needs. As the system size grows, this will be a big benefit for calculation.

VI. Conclusion

In this work, we propose an optimal steady-state voltage control framework for distribution power systems. The proposed control algorithms are developed based on the novel Gaussian process regression-based linear *Voltage-Power* relationship that accurately approximates the voltage within a power injection subspace. The new control algorithms optimize not only the control efforts but also the voltage reference setpoints that further minimize the voltage deviation errors. Using the least squares estimation, analytical forms of the optimal control solutions are derived for both centralized and distributed control schemes. For uncertain injections, bounds on the system voltage variations are developed and used to design robust optimal control that guarantees the voltage compliance. A new network partition scheme for distributed voltage control is also presented, relying on the concept of Effective Voltage Control Sources and convex relaxation optimization. The simulation results on IEEE 33-Bus and IEEE 69-Bus systems illustrate that the proposed GP learning method and corresponding optimal voltage control algorithms perform well over a wide range of operating conditions. The future work will focus on an extension for handling uncertainties with considering time dependent parameters, such as the line impedances, and network topology changes.

References

[1] N. Hatziargyriou *et al.*, "Contribution to bulk system control and stability by distributed energy resources connected at distribution network," IEEE, Tech. Rep., 2017.

[2] M. N. Kabir *et al.*, "Coordinated control of grid-connected photovoltaic reactive power and battery energy storage systems to improve the voltage profile of a residential distribution feeder," *IEEE Trans. on industrial Informatics*, vol. 10, no. 2, pp. 967–977, 2014.

[3] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*. McGraw-hill New York, 1994, vol. 7.

[4] A. Zobian *et al.*, "A steady state voltage monitoring and control algorithm using localized least square minimization of load voltage deviations," *IEEE Trans. on Power Systems*, vol. 11, no. 2, pp. 929–938, 1996.

[5] M. D. Ilic *et al.*, "Improved secondary and new tertiary voltage control," *IEEE Trans. on power systems*, vol. 10, no. 4, pp. 1851–1862, 1995.

[6] H. D. Nguyen, K. Dvijotham, S. Yu, and K. Turitsyn, "A framework for robust long-term voltage stability of distribution systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 4827–4837, 2018.

[7] A. Conejo, T. Gomez, and J. De la Fuente, "Pilot-bus selection for secondary voltage control," *European Trans. on Electrical Power*, vol. 3, no. 5, pp. 359–366, 1993.

[8] K. E. A-Plytaria *et al.*, "Distributed and decentralized voltage control of smart distribution networks: Models, methods, and future research," *IEEE Trans. on smart grid*, vol. 8, no. 6, pp. 2999–3008, 2017.

[9] D. K. Molzahn *et al.*, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Trans. on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.

[10] S. Magnússon, G. Qu, and N. Li, "Distributed optimal voltage control with asynchronous and delayed communication," *IEEE Trans. on Smart Grid*, 2020.

[11] P. N. Vovos and Others, "Centralized and distributed voltage control: Impact on distributed generation penetration," *IEEE Trans. on Power Systems*, vol. 22, no. 1, pp. 476–483, 2007.

[12] R. Tonkoski, L. A. Lopes, and T. H. El-Fouly, "Coordinated active power curtailment of grid connected pv inverters for overvoltage prevention," *IEEE Trans. on sustainable energy*, vol. 2, no. 2, pp. 139–147, 2010.

[13] H. D. Nguyen and K. Turitsyn, "Robust stability assessment in the presence of load dynamics uncertainty," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1579–1594, 2015.

[14] H. J. Liu, W. Shi, and H. Zhu, "Decentralized dynamic optimization for power network voltage control," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 3, pp. 568–579, 2016.

[15] J. R. Castro *et al.*, "Optimal voltage control in distribution network in the presence of dgs," *International Journal of Electrical Power & Energy Systems*, vol. 78, pp. 239–247, 2016.

[16] J. Zhao and F. Dörfler, "Distributed control and optimization in dc microgrids," *Automatica*, vol. 61, pp. 18–26, 2015.

[17] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Power Engineering Review*, vol. 9, no. 4, pp. 101–102, 1989.

[18] S. Bolognani and S. Zampieri, "On the existence and linear approximation of the power flow solution in power distribution networks," *IEEE Trans. on Power Systems*, vol. 31, no. 1, pp. 163–172, 2015.

[19] H. Xu *et al.*, "Data-driven voltage regulation in radial power distribution systems," *IEEE Trans. on Power Systems*, vol. 35, no. 3, pp. 2133–2143, 2019.

[20] Y. Chen, Y. Shi, and B. Zhang, "Input convex neural networks for optimal voltage regulation," *arXiv preprint arXiv:2002.08684*, 2020.

[21] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Power Engineering Review*, vol. 9, no. 4, pp. 101–102, 1989.

[22] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

[23] J. Yan, K. Li, E.-W. Bai, J. Deng, and A. M. Foley, "Hybrid probabilistic wind power forecasting using temporally local gaussian process," *IEEE Trans. on Sustainable Energy*, vol. 7, no. 1, pp. 87–95, 2015.

[24] P. Pareek and H. D. Nguyen, "Probabilistic robust small-signal stability framework using gaussian process learning," *Electric Power Systems Research*, vol. 188, p. 106545, 2020.

[25] Y. Yang *et al.*, "Power load probability density forecasting using gaussian process quantile regression," *Applied Energy*, vol. 213, pp. 499–509, 2018.

[26] P. Pareek and H. D. Nguyen, "Gaussian process learning-based probabilistic optimal power flow," *IEEE Trans. on Power Systems*, 2020.

[27] M. Welling, "Kernel ridge regression," *Max Welling's Classnotes in Machine Learning*, pp. 1–3, 2013.

[28] R. D. Zimmerman *et al.*, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. on power systems*, vol. 26, no. 1, pp. 12–19, 2010.

[29] H. D. Nguyen, K. Dvijotham, and K. Turitsyn, "Constructing convex inner approximations of steady-state security regions," *IEEE Trans. on Power Systems*, vol. 34, no. 1, pp. 257–267, 2018.

[30] P. Suetin, A. I. Kostrikin, and Y. I. Manin, *Linear algebra and geometry*. CRC Press, 1989.

[31] M. Ilic-Spong, J. Christensen, and K. Eichorn, "Secondary voltage control using pilot point information," *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 660–668, 1988.

[32] M. Planitz, "Inconsistent systems of linear equations," *The Mathematical Gazette*, vol. 63, no. 425, pp. 181–185, 1979.

[33] D. Molzahn and L. A. Roald, "Grid-aware versus grid-agnostic distribution system control: A method for certifying engineering constraint satisfaction," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[34] Y. Chai *et al.*, "Network partition and voltage coordination control for distribution networks with high penetration of distributed pv units," *IEEE Trans. on Power Systems*, vol. 33, no. 3, pp. 3396–3407, 2018.

[35] X. He *et al.*, "Laplacian regularized gaussian mixture model for data clustering," *IEEE Trans. on Knowledge and Data Engineering*, vol. 23, no. 9, pp. 1406–1418, 2010.

[36] D. Das, "Optimal placement of capacitors in radial distribution system using a fuzzy-ga method," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6-7, pp. 361–367, 2008.

[37] C. E. Rasmussen and H. Nickisch, "Gaussian processes for machine learning (gpml) toolbox," *The Journal of Machine Learning Research*, vol. 11, pp. 3011–3015, 2010.

[38] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. on power Delivery*, vol. 4, no. 1, pp. 735–743, 1989.