A framework for analytical power flow solution using Gaussian process learning

Pareek, Parikshit; Nguyen, Hung D.

2021

Pareek, P. & Nguyen, H. D. (2021). A framework for analytical power flow solution using Gaussian process learning. IEEE Transactions On Sustainable Energy.
<https://dx.doi.org/10.1109/TSTE.2021.3116544>

<https://hdl.handle.net/10356/153083>

<https://doi.org/10.1109/TSTE.2021.3116544>

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at:
<https://doi.org/10.1109/TSTE.2021.3116544>.

Downloaded on 26 Apr 2023 21:06:33 SGT

A Framework for Analytical Power Flow Solution using Gaussian Process Learning

Parikshit Pareek, *Student Member, IEEE*, and Hung D. Nguyen* *Member, IEEE*

This paper has been accepted for publication in *IEEE Transactions on Sustainable Energy*

Link: <https://ieeexplore.ieee.org/document/9552521>

Abstract—This paper proposes a novel analytical solution framework for power flow (PF) solutions in active distribution networks under uncertainty. We use the Gaussian process (GP) regression to learn node voltage as a function of effective bus load or negative net-injection vector. The proposed approximation is valid over a subspace of load and provides an understanding of system behavior under uncertainty via GP interpretability. We interpret the relative variation extent of different node voltages using the *quality ratio* (QR) defined based on the *hyper-parameters* of GP. Further, the application of the proposed framework in calculation of voltage limit violation probability and dominant voltage influencer ranking has also been presented. Through test simulations for 33-bus and 56-bus systems, the proposed method achieves low mean absolute error (MAE) of order E-05 (pu) in voltage magnitude and E-04 (rad) in angle. The discussions on salient features of the proposed method and comparative analysis with large-scale Monte-Carlo simulations, and state-of-art methods is also presented for the proposed applications.

Index Terms—Analytical Power Flow Solution, Gaussian Process Learning, Voltage Violation Probability.

I. INTRODUCTION

THE task of calculating the steady-state of the grid through solving a set of nonlinear power flow (PF) equations is one of the classical problems in circuit and power system analysis. The set of nonlinear equations describing the alternate current power flow (ACPF) does not come with analytical or closed-form expression. Further, solving ACPF solution, in presence of injection uncertainties, is necessary for solving problems such as probabilistic power flow [1]–[3], renewable housing capacity [4], [5] and active distribution network (ADN) operation [6] etc.

Traditionally, this challenge in solving ACPF has led considerable efforts in developing approximated linear power flow methods, among which direct current power flow (DCPF) approximation is most notable [7]. DCPF has been utilised for numerous applications like optimal power flow (OPF) [8], market studies [9], and risk assessment [10] etc. The DCPF is mostly studied for transmission system as it relies on r/x -based approximation and does not model the voltage magnitude [11]. Recently various other linear power flow (LPF) methods have been proposed. In [12] linearization based on selection of different *independent variable* has been generalised with discussion on accuracy. Error bounds on the

performance of the general linear model have been presented [13], while details of various methods can be found in [14].

Recently, advancements in phasor measurement unit (PMU) deployment and applications have made data-driven methods an interesting choice for power system modeling [15]. In some initial efforts, data-driven PF approximations have been presented in [16], [17]. Methods such as partial least square (PLS) have been proposed to approximate linear PF relationships in [17]. In [18], authors extend their model to approximate PF in the presence of measurement noise. They formulate a linearly constrained quadratic programming-based algorithm to obtain the initial point and perform linearization. Recently, [19] attempts merging advantages of data-driven and model-based analysis to obtain linearized power flow. Most of the model-based approximations focus on linear form via decoupling the voltage-power relationship under r/x ratio assumptions. Further, data-based works explicitly focus on learning linear models and present testing results on a set of loading points. Developing general analytical and closed-form solution framework or model interpretability is not the objective in many of the works. Moreover, the different RES generation and EV loads have various types of uncertainty distributions. The information and parameters of these distributions are often not easy to obtain. Thus various existing methods suffer from accuracy limitations when load or injection probability distribution is not estimated accurately. These issues are the motivation behind developing the proposed analytical solution framework for power flow which directly handles interaction between ADNs and multiple types of renewable sources and uncertain demands.

In this paper, we propose a general analytical power flow solution framework which provides a closed-form expression for characterizing the power flow (PF) solution. Here, analytical solution mean a closed-form expression relating nodal voltage with power injections or net-demand. We define the closed-form PF (CFPF) as a mapping (function) between net bus injections and nodal voltage. The main focus of this work is to map the individual node voltage as a function of an effective bus load vector using the Gaussian process (GP) regression for learning. The closed-form expressions obtained can provide approximate voltage solution just by *plugging* the load vector into the expression. The present work builds the basis of an extensive framework where desired accuracy-complexity trade-off can be selected based on the application for which CFPF is needed. We develop various closed-form approximations based on different covariance functions.

*Corresponding Author

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

Further, we build two different applications, violation probability calculation [20] and dominant influencer of voltage fluctuation (DIVF) ranking [21], based on the proposed CFPF. These proposed CFPF expression provides distributional robustness to these applications. Also, it reduces computational complexity via avoiding re-training needs and faster functional evaluation compared to traditional ACPF solving methods. The additional motive of using GP as a learning tool is to provide interpretability to the proposed CFPF. The attempt is to explain *what is happening inside the learning model?* A quality ratio (QR) is defined based on covariance function *hyper-parameters* and discussed in detail to provide *qualitative* measure of the relative extent of variation, and comparative nonlinearity among different node voltage functions. The proposed work focuses directly on distribution networks as they are subjected to large scale penetration of renewable sources (RES) of different type as well as electric vehicle (EV) related demand uncertainties. These uncertainties are preset at large number of ADN nodes, as penetration of RES and EVs increases in future. The proposed CFPF and applications based on it can work with multiple number and type of uncertain net-loads (demand minus generation) simultaneously.

The main contributions of this work can be summarized as:

- Introducing a novel analytical solution framework for power flow in active distribution networks interacting with sustainable sources and varying demands. The framework facilitates the accuracy-complexity analysis for various voltage approximation based applications.
- Developing methods to calculate violation probability and dominant influencer of voltage fluctuation (DIVF) ranking in non-parametric and computationally light manner.
- Demonstrating the PF learning model's interpretability to infer details about the accuracy of learning. A quality ratio (QR) is defined to provide a *qualitative* measure of voltage function variability and nonlinearity.

The idea of obtaining the analytical solution of power flow lies in developing a closed-form approximation of equation relating node voltage with net-load vector. The closed-form approximation means that we can directly obtain node voltage solution by putting numerical values of injection vector into the function. Further, unlike linearization works [17]–[19], the present develops a framework to obtain a general, nonlinear closed-form analytical solution of power flow. Moreover, the proposed CFPF framework provides unique features namely : *Subspace-wise Approximation*, *Non-parametric Nature*, *Inherent Accuracy Indicator*, and *Interpretability*. These features provide advantages to proposed CFPF, like probabilistic information (available without additional computational cost) is proposed to be used as an inherent error indicator. The work also utilizes the analysis capabilities of the CFPF by proposing the novel parameter for voltage nonlinearity analysis as the quality ratio (QR). Such features and analysis have not been explored in existing linearization works [17]–[19]. Further, [22], [23] are time series forecasting works which is the first frontier where GP was applied in power system being a supervised learning tool. The works [24]–[27] are the ones where GP is directly used to learn a particular

input-output relationship as a black-box, without exploring the major possibilities of analysis, except in [24] where attempts have been made to exploit interpretability of GP and provide understanding about results. Further, the work [24] uses GP to perform uncertainty quantification (UQ) for ACOPF. That work's objective is limited to exploit the non-parametric nature of GP for efficient UQ. The work [24] does not intend to provide a closed-form-based analysis of ACOPF or ACPF problems.

The work [28] is focused on presenting probabilistic robust stability certificate constructed using GP upper confidence bound (GP-UCB) algorithm in Bayesian optimization (BO) framework, which is a different problem. The biggest shortcoming of BO is that it is only useful with low dimensional input space and does not scale well for medium to large dimensions. We develop a UQ application, violation probability calculation, only to show the effectiveness of the proposed CFPF, and the use of the proposed framework is not limited to UQ only. The closed-form expressions developed are differentiable, and thus these can be utilized to perform various tasks of control and optimization under uncertainty.

In following sections, we first define the basic premises of PF and describe GP regression. Then, we build the general closed-form power flow (CFPF) approximation framework. We provide various closed-form expressions of voltage magnitude function using different covariance functions with GP in matrix form. We also present details about unique features of the proposed framework. We propose two different applications of CFPF to understand power system behaviour under uncertainty. Particularly, the applications are focused on ADNs as they are experiencing combined effect of load and RES uncertainty and need special methods. The applications developed based on CFPF are non-parametric and thus can handle the uncertainty present in net-load of ADNs jointly due to RES and load variations. The conclusion then follows the results and discussion.

II. BASICS

In this section, we revisit the power flow from function point of view and describe details of the GP regression. Before moving to subsections, we describe the distribution system model and notations. Consider a single-phase balanced representation of the network with only one voltage-controlled slack bus. The set $\mathcal{B}$ contains all other buses, which are load buses (PQ bus type). The node index is j , and the cardinality of the PQ bus set is defined as $|\mathcal{B}|$. Bus and node are used interchangeably. We follow the polar form and define complex voltage at a particular node as $v_j = V_j \angle \theta_j$ with V_j as magnitude and θ_j as the angle. The voltage vector and angle vector are indicated as $\mathbf{V}$ and $\boldsymbol{\theta}$. We use voltage to refer to the magnitude of complex node voltage hereafter. We use the modeling convention wherein the load is considered as positive, and any distributed generation (DG) at that particular bus is modeled as a negative load. This effective load is the same as the negative of net-injection. The load hereafter means effective or net load only. Real and reactive loads are defined as p_j and q_j respectively, while a joint column vector of load

is $\mathbf{s} = [\mathbf{p}; \mathbf{q}]$. All bold symbols represent the column vector of the appropriate dimension unless stated otherwise. Now we define the functional form of power flow.

A. Power Flow in Functional Form

The power flow equation set represents the relationship between bus load and node voltage phasor through nonlinear trigonometric equations [29]. In compact form

$$\mathbf{z} = F(\mathbf{s}), \quad (1)$$

here, $F(\cdot)$ is a family of functions and $\mathbf{z} = [\boldsymbol{\theta}; \mathbf{V}]$ represents the node voltage angle and magnitude vector. There is no closed-form relationship between load and voltage where one can obtain a numerical voltage solution by putting numerical load values. Therefore, the set of equations (1) is solved using iterative numerical methods like the Newton-Raphson load flow (NRLF) [29]. We propose to learn the function family $F(\cdot)$, relating the node voltage solution $\mathbf{z}$ with bus load $\mathbf{s}$.

An important point to discuss is that the nonlinear power flow problem may have multiple voltage solutions [30]–[32]. However, normal operating conditions may admit only one high voltage solution. Most other low voltage solutions are not viable and easily distinguishable from the high voltage point [33], [34]. While we do not attempt to give a condition for the uniqueness of the high voltage solution in this work, the training set is restricted to contain only high voltage points. In other words, we learn only the relationship between the high voltage solution and the power injections. At the same time, we do not rule out the possibility that the learning can mix the high voltage solution with a low voltage solution when the two solutions are close to each other as the operating point approaches the tip of the nose curve. In this scenario, the Jacobian of the power flow becomes ill-conditioned and eventually violates the assumption on the non-singularity of the Jacobian which is the underlying condition of solving power flow using NRLF [29].

In this work, our purpose is to obtain an approximate CFPF: $\mathbf{z} = F(\mathbf{s})$, where $\mathbf{s} \in \mathcal{L}$. We define the load vector set $\mathcal{L}$ in more generic way compared to conventional load set descriptions based on load scaling and power factor fixation. We use a box type description of individual bus real and reactive power load. Typical bus load is given as $p_j \in [p_j^o \pm \delta p_j^o]$ and $q_j \in [q_j^o \pm \delta q_j^o]$ where random variable $\delta \in (0, 1)$ is in fraction constructing the region around the base-load p_j^o and q_j^o . Importantly, we do not fix the power factor and the set $\mathcal{L}$ is not described in terms of total load variations but by varying the individual bus's real and reactive power load. Formally, the D -dimensional load subspace $\mathcal{L} = \{[p_j; q_j] | p_j \in [p_j^o \pm \delta p_j^o]; q_j \in [q_j^o \pm \delta q_j^o]; \delta \in (0, 1) \forall j = 1 \dots |\mathcal{B}|\}$. Below we introduce details of GP based method for learning the approximated PF function family $F(\mathbf{s})$.

Note: The proposed CFPF framework is independent of the load flow method used to solve the power flow equations for training and testing data. The discussion related to function family of power flow $\mathbf{z} = F(\mathbf{s})$ in section II-A remains valid with any power flow method such as [35], [36]. This work employs MATPOWER's 'runpf' function [37] with NRLF to

obtain training and testing data as it is a well-established tool. However, any other load flow solver can also be used without requiring any modifications in the proposed framework.

B. Gaussian Process Learning

The Gaussian process (GP) regression is a non-parametric technique that obtains a distribution over the probable functions, having consistency with the observed data [38]. It affords some level of interpretability of models and allows us to infuse a prior understanding of the possible functions like other Bayesian methods [38]. This prior knowledge infusion allows us to model physical properties, and gives a choice on the desired complexity of the model. The GP is non-parametric in nature, meaning it does not require any particular uncertainty description of input. This property also gives GP regression greater flexibility to fit model over given data [26].

In this work, we employ GP regression to learn the function family of PF $F(\mathbf{s})$. In other words, we want to model the relationship between load vector $\mathbf{s}$, and individual node V_j and θ_j . Some recent works have applied GP to various power system applications like various forecasting problems [22], [23], probabilistic optimal power flow and economic dispatch [24], [25] and probabilistic PF [26], [27], etc. A recent work also attempts to learn the critical Eigenvalue function utilizing GP [28]. The problem of learning or model fitting with GP regression means finding suitable properties of the covariance function. The covariance or kernel function defines the relationship between the input and target output. The kernel function properties are defined mathematically using some parameters called *hyper-parameters* indicated as set ϕ . This set ϕ gives GP its distinct feature: *Interpretability*.

The GP interpretability means that upon learning, the ϕ contains subtle details about the function, which is inferred by fitting a model over given data. These details can then be used to explore the function behavior, dependencies on different dimensions of input, and estimating the extent of variation of the function. It allows us to know what is happening inside the learning algorithm. Below we present the general GP regression from a function learning point of view.

At first, we construct a training set which involves input vector and corresponding target function value. We define an input training matrix, called *design matrix*, $X = [\mathbf{x}^1 \dots \mathbf{x}^i \dots \mathbf{x}^N]$ with $\mathbf{x}^i$ being i^{th} input vector. We also define $\mathbf{y} = [y^1 \dots y^i \dots y^N]$ as target vector which gives us the learning data pair of total N samples as $\{\mathbf{x}^i, y^i | i = 1 \dots N\}$. Now using this training data, a GP model is fit to obtain a regression model as

$$y = f(\mathbf{x}) + \epsilon \quad (2)$$

where, $\mathbf{x}$ is the input vector, $f(\cdot)$ is function value, and y is the target value. Here, (2) includes noise to model the situation when target y does not exactly follow the function $f(\mathbf{x})$. The assumption over the noise is that it follows an independent, identically distributed Gaussian distribution given as $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ with zero mean and σ_ϵ^2 variance. Now, we

define a zero mean GP function-space view of GP [38], using a general form of the covariance function $k(\mathbf{x}^i, \mathbf{x}^j)$ as

$$f(\mathbf{x}^i) \sim \mathcal{GP}(0, k(\mathbf{x}^i, \mathbf{x}^j)). \quad (3)$$

Now, employing the maximum-likelihood estimator, mean prediction using noisy observations of the function can be given as (chapter 2, function-space view [38])

$$\bar{f}(\mathbf{x}^*) = \mathbf{k}^T \boldsymbol{\alpha} \quad (4)$$

Here, $\boldsymbol{\alpha} = (K + \sigma_\epsilon^2 I)^{-1} \mathbf{y}$, while $K = k(X, X)$ is kernel matrix of training samples constructed using the covariance function $k(\mathbf{x}^i, \mathbf{x}^j)$. The kernel vector between X and $\mathbf{x}^*$ is given as $\mathbf{k} = k(X, \mathbf{x}^*)$. The details of kernel matrix, vector and predictive variance function are given in Appendix A.

It is necessary to highlight that in (4), $\boldsymbol{\alpha}$ is constant upon learning, and thus the function depends on $\mathbf{x}^*$ through the kernel vector $\mathbf{k}$. Here, if the $\mathbf{x}^*$ is a variable vector, we obtain the mean prediction in the functional form. Interested readers can find details of GP regression in [38]. Now, before explaining the proposed framework, we briefly describe the motivation behind using GP as a regression tool.

Implementing the proposed CFPF framework's idea requires a nonlinear regression method that provides closed-form expression, is non-parametric, and has a generative model to obtain output distributions upon learning. Further, we want interpretability and uncertainty handling capacity associated with the learning method to understand the CFPF model better. Among supervised learning methods, the kernel methods (Support Vector Machine (SVM), GP, etc.) directly satisfy the main criteria by providing nonlinear close-form relating input with output. However, being a generative model and having uncertainty handling capabilities, GP is a more suitable choice due to the features like [38]: i) kernel parameters learned from data with high flexibility, thus providing nonlinear forms; ii) interpretable training and prediction uncertainty information available without extra computational cost to test the learning quality, iii) no cross-validation needed for regularization parameter learning, resulting in computational time saving etc. Therefore, this work uses GP as a regression tool. However, any method that can fully fill the framework's essential requirements can be used as a learning tool.

In this work, a significant GP implementation modification is that a modified way to generate training samples is used to avoid tractability issues of exponential sample demand in high-dimensional grid sampling and lower subspace coverage of random sampling. Rather than directly generating random input samples, we generate random grid indices for a given resolution. First, we obtain ' r ' equally spaced points, including limits, for each input dimension. Then, we generate N random grid indices via generating D random integers between 0 to ' r ', total N times. Importantly, in this manner, these random grid indices of two different input dimensions are independent. The effect of this can be seen directly on the Euclidean distance metric of training inputs. Compared with random sampling, the proposed random index sampling has a 17.62% (4.23%) higher minimum (maximum) value and 6.38% higher variance in Euclidean distance between 200 samples with $r = 20$ sections for each 64 input dimensions. These higher

values indicate that proposed random index sampling avoids having nearby training samples and provides a better spread overload subspace. Further, one can modify the coverage and spacing by changing the number of sections ' r '.

III. ANALYTICAL SOLUTION OF POWER FLOW PROBLEM

In this section, we describe the method to obtain an analytical solution of power flow problem via closed-form approximation of F . At first, we explain the approximation with generic covariance function $k(\cdot, \cdot)$ and then we present different covariance function-based closed-forms. The description in the section is presented in terms of voltage magnitude function approximation, and angle function approximation can be derived similarly. Further, the mean prediction of function with noisy observations (4) is used to derive CFPF. In the second half of section, we explain the features of the proposed approximation for analytical solution of power flow.

First, we obtain a training set for an individual node voltage as $\{\mathbf{s}^i, \hat{V}_j^i \mid i = 1 \dots N\}$ where $\hat{V}_j$ is ACPF voltage magnitude solution [37] at j^{th} bus with Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. We define the *design matrix* $S = [\mathbf{s}^1 \dots \mathbf{s}^N]$ as a $D \times N$ matrix containing load vectors while $\hat{\mathbf{V}}_j = [\hat{V}_j^1 \dots \hat{V}_j^N]$. In this work, we consider the network topology to be fixed. This means that network admittance matrix is constant.

Now, a GP model is fitted on this training data to obtain the analytical solution voltage V_j , as a function of variable load vector $\mathbf{s}$ as $V_j = f(\mathbf{s})$. The ' $\mathbf{s}$ ', without superscript index, indicate a variable vector here, and $f(\cdot)$ is PF mapping function. Using the mean prediction expression (4) with covariance function $k(\cdot, \cdot)$ we obtain

$$V_j(\mathbf{s}) = \mathbf{k}^T(S, \mathbf{s})(k(S, S) + \sigma_\epsilon^2 I)^{-1} \hat{\mathbf{V}}_j$$

Or $V_j(\mathbf{s}) = \mathbf{k}^T \boldsymbol{\alpha}_j$ (5)

An interesting interpretation of (5) is in terms of linear combination of $\mathbf{k}$ and $\boldsymbol{\alpha}_j$ elements. The form where elements of $\boldsymbol{\alpha}_j$ act as optimal weights over variable vector $\mathbf{k}$ is

$$V_j(\mathbf{s}) = \sum_i^N \alpha_j(i) k(\mathbf{s}^i, \mathbf{s}). \quad (6)$$

Note that for each function $V_j(\mathbf{s})$ we obtain a set of *hyper-parameters* ϕ , which determines the covariance function upon learning. Interestingly, the inclusion of σ_ϵ in $\boldsymbol{\alpha}_j$ allows GP to learn from the noisy data, obtained from PMU measurements. The analytical expressions (5) and (6), represents voltage as a function variable load vector $\mathbf{s}$, while $\mathbf{s}^i$ is known numerical training samples, and $\boldsymbol{\alpha}_j = (k(S, S) + \sigma_\epsilon^2 I)^{-1} \hat{\mathbf{V}}_j$ is constant upon learning. Similarly, the approximated function expressing relationship of θ_j with ' $\mathbf{s}$ ' can be obtained by learning using a training data set $\{\mathbf{s}^i, \hat{\theta}_j^i \mid i = 1 \dots N\}$. Following this mechanism, we get function $V_j(\mathbf{s})$ and $\theta_j(\mathbf{s})$ for each node.

From (5), and (6) with definitions of covariance vector $\mathbf{k}$ and constant $\boldsymbol{\alpha}_j$ upon training (as explained with (4)), it is clear that approximated closed-form function of voltage $V_j(\mathbf{s})$ is directly dependent on covariance function, selected for the GP regression. The known or presumed properties of $V_j(\mathbf{s})$ drive the selection of the covariance function. Another criterion is complexity in the closed-form expression

obtained as a result of using any covariance function. This criterion restricts us from using covariance functions such as *Matérn Class* of covariance function and non-stationary covariance function in general. We examine the performance of three different types of covariance functions: *Polynomial* k_{P_n} , *Squared Exponential* (SE) k_{SE} and *Rational Quadratic* (RQ) k_{RQ} [38]. We restrict ourselves up to polynomial of degree two (Quadratic) k_{QD} to limit the complexity in CFPF expression of $V_j(\mathbf{s})$. These functions are stationary, meaning that Euclidean distance governs the covariance between any two points. Another important aspect of consideration is the smoothness of the target function. As PF is defined inside the solvable region, the voltage function is considered smooth without discontinuities.

Below we list these different covariance functions and associated voltage closed-form expressions following (5).

- Polynomial Kernel (degree n): $\phi_P = \{\tau, c, l\}$

$$k_{P_n}(S, \mathbf{s}) = \tau^2 \left(c\mathbf{1} + \frac{S^T \mathbf{s}}{l^2} \right)^n; \quad (7)$$

$$\implies V_j(\mathbf{s}) = [k_{P_n}(S, \mathbf{s})]^T \boldsymbol{\alpha}_j$$

- Square Exponential Kernel: $\phi_{SE} = \{\tau, l\}$

$$k_{SE}(\mathbf{s}^i, \mathbf{s}) = \tau^2 \exp \left(-\frac{\|\mathbf{s}^i - \mathbf{s}\|^2}{2l^2} \right); \quad (8)$$

$$\implies V_j(\mathbf{s}) = [k_{SE}(\mathbf{s}^1, \mathbf{s}) \quad \dots \quad k_{SE}(\mathbf{s}^N, \mathbf{s})] \boldsymbol{\alpha}_j$$

- Rational Quadratic Kernel: $\phi_{RQ} = \{\tau, \beta, l\}$

$$k_{RQ}(\mathbf{s}^i, \mathbf{s}) = \tau^2 \left(1 + \frac{\|\mathbf{s}^i - \mathbf{s}\|^2}{2\beta l^2} \right)^{-\beta}; \quad (9)$$

$$\implies V_j(\mathbf{s}) = [k_{RQ}(\mathbf{s}^1, \mathbf{s}) \quad \dots \quad k_{RQ}(\mathbf{s}^N, \mathbf{s})] \boldsymbol{\alpha}_j.$$

In the following, we discuss unique features associated with the proposed approximation framework. These features differentiate the proposed framework from the existing models and data-based power flow approximation methods.

First we discuss the subspace-wise approximation feature. The proposed framework obtains the analytical power flow solution over the load subspace $\mathcal{L} = \{[p_j; q_j] | p_j \in [p_j^o \pm \delta p_j^o]; q_j \in [q_j^o \pm \delta q_j^o]; \delta \in (0, 1) \forall j = 1 \dots |\mathcal{B}|\}$. As we do not rely on slope calculation or Taylor series expansion at any base-load or voltage point for approximation like model-based methods [39], the accuracy is not a function of distance from base point. The *hyper-parameters* are obtained by log-marginal likelihood maximization over a training set, which is distributed over the subspace $\mathcal{L}$ [38]. Thus, the closed-form can be utilized for control and operational applications over entire load subspace.

A unique feature of the proposed CFPF is its non-parametric nature. Unlike parametric models of learning, the GP has no limitation related to the type of distribution function assumed to model the load inputs [26]. While learning CFPF, we have made no assumption of distribution of load vector and use ‘s’ selected from grid over the subspace $\mathcal{L}$ for training. This feature is particularly important while developing probabilistic assessment methods based on closed-form approximations. In short, we can obtain output distribution for all types of input

distributions, upon learning the closed-form once [26]. This non-parametric nature is, however, different from assumption taken for noise ϵ to be Gaussian. Nevertheless, the limitation of noise distribution being Gaussian will also be dealt with in future works using methods such as [40].

We now discuss about an inherent accuracy indicator provided by GP, and its importance in ML-based models for power system. The data-based approximations rely on error indicators such as mean absolute error (MAE), calculated over a vast number of testing samples, to evaluate the performance [17], [18]. However, solution quality at any new test point is seldom indicated directly. The GP provides predicted variance $\sigma^2[f(\mathbf{s})]$ (16) for every test point which is indicative of learning quality. In a way, it contains error information. Also, a confidence region for predictions can be constructed using standard deviation calculated from (16) given in Appendix A. With an increase in predicted variance, the error in voltage estimation tends to increase. Therefore, variance in the proposed framework is an inherent accuracy indicator, estimating the approximation quality without repeated comparison with true ACPF solution.

The proposed GP-based CFPF approximation supports model interpretability. A recent review [41] identifies interpretability as one of the big challenges for ML models applied to power system. Below we discuss the interpretability of the proposed analytical solution method and develop a *qualitative* measure of nonlinearity.

The GP facilitates the interpretability through analyzing *hyper-parameters*. Here, interpretability means understanding the reason behind certain voltage function’s are better approximated in closed-form and understanding the internal mechanism of approximation method, which affects the accuracy of approximation. A higher value of *characteristic-length* l_j in (8) means that particular voltage varies slower (or less) with respect to load changes. A lower τ_j suggests the $V_j(\mathbf{s})$ does not have large magnitude variations over the load subspace $\mathcal{L}$. Their ratio is important to establish a *qualitative measure* of nonlinearity and causality of accuracy levels in learning different node voltage and angle functions. We define l_j/τ_j as quality ratio (QR). The lower QR for $V_j(\mathbf{s})$ means large and fast variations in function $V_j(\mathbf{s})$ leading to poor approximation.

IV. APPLICATIONS OF CLOSED-FORM POWER FLOW

In this section, we present applications of the proposed analytical power flow solution method. We employ the CFPF for calculating the probability of operational voltage limit violation when ADN operation is subjected to uncertainties [20]. As the proposed method is developed for uncertainty in net-load, proposed VP calculation method is applicable when both load and distributed RESs are uncertain. Further, we also present the application of the analytical *voltage-power* function for identification of dominant influencer of voltage fluctuation (DIVF) at each node of ADN. The focus is on ADNs as they are experiencing the uncertain load and injection at large number of loads with increase in number of EVs and renewable sources. This means that different sources of uncertainty combines together and require a non-parametric treatment as done by proposed method. In future,

the framework will be extended to other types of networks and their applications.

A. Voltage Violation Probability

The probability of voltage violation is defined as chances of voltage magnitude being outside the operational limits, in the presence of uncertainties in net-load [4], [20], [42]. We define the voltage violation probability (VP) at j -th node as

$$q_j = 1 - \mathbb{P}(\underline{V} \leq V_j(\mathbf{s}) \leq \bar{V} | \mathbf{s} \in \mathcal{L}). \quad (10)$$

The efficient and accurate calculation of violation probability (VP) (10) is challenging as it requires accurate estimation of the probability distribution function (PDF) of the output. Thus, the VP estimation accuracy directly depends upon the quality of output PDF estimation by any method. The challenge increases manifold when input uncertainty distribution is not known prior or does not follow a standard form. The MCS of power flow is a promising approach for VP calculation as the numerical simulation-based methods can work with any uncertainty distribution. Yet, they are time consuming as the accuracy directly depends upon the number of samples used MCS and every sample requires a multi-iteration NRLEF. The methods to calculate VP, other than MCS, are built with the information of uncertainty distribution type and parameters [43]. Obtaining the uncertainty information is becoming increasingly challenging due to overlapping of EVs and RESs uncertainties in ADN net-load. In view of these obstacles in VP calculation, we present below the VP calculation method via non-parametric CFPF solution method.

The proposed CFPF approximation can work with any distribution like MCS and have reduced computational burden like polynomial chaos expansion [43]. The analytical function (7)-(9) can estimate the VP for various input distribution without retraining of the model. Further, once trained, the CFPF relationship works as an analytical function and point-wise calculation is performed for each input sample. The calculation method is similar to well established NRLEF based MCS method. In CFPF-based VP calculation, voltage function evaluations are done for large number of load samples. The computational complexity of the function evaluation does not change with the system size and time taken to obtain voltage for very large number of input load samples is negligible compared to solving NRLEF. Therefore, the proposed CFPF is an efficient method to obtain q_j for different input PDFs.

B. Dominant Influencer of Voltage Fluctuation (DIVF)

In ADNs, the increase in distributed RES and load fluctuations introduce large variations in the node voltage. To design control mechanism of such fluctuations, it is highly recommended to identify the particular node which is the biggest influencer of a node voltage variation [21]. This can help reduce the node voltage variations by controlling the dominant factor i.e. injection at DIVF node. Yet, identification of the dominant factor which is causing voltage variation at a node is not straightforward due to nonlinearity of power flow, network graph influence as well as variations in the type and size of uncertain net injections of DGs or load variations.

Fig. 1. Flowchart depicting process of CFPF model construction and using voltage function for different applications with its advantages.

The MCS for DIVF is a preferred choice as it works for all different type of distributions and preserves nonlinearity of power flow. Nevertheless, MCS is time consuming. The proposed framework represents the power flow in more efficient analytical functional form (5) to calculate the voltage. This means that we can obtain the numerical value of node voltage for large number of load inputs by just evaluating the analytical CFPF function. Further, the analytical expression (5) is valid for a subspace, eliminating the need of retraining for calculating the DIVF. Therefore, the use of proposed analytical PF formula will eliminate the parametric limitation i.e. specific uncertainty distribution information not required, and will be computationally light as only function evaluation is needed.

Similar to the true DIVF definition given in [21], we define a node j_d as dominant influencer (DI) for the node voltage V_j if injection at j_d causes highest variance change in the voltage. Mathematically,

$$DI = \arg \max_{j_d} \{ \sigma^2(V_j) - \sigma^2(V_j^{-j_d}) \} \quad (11)$$

here, $\sigma^2(V_j)$ is variance in j -th node voltage due to all uncertain loads and $\sigma^2(V_j^{-j_d})$ is variance due to all uncertain loads, except j_d node. Similarly, the DI can be defined based on voltage change variance as well. Further, we define relative influence (RI) of DI nodes based on normalization by most influencing node variance change as:

$$RI_j = \frac{\{ \sigma^2(V_j) - \sigma^2(V_j^{-j}) \}}{\max_{j_d} \{ \sigma^2(V_j) - \sigma^2(V_j^{-j_d}) \}} \quad (12)$$

It is clear from (12) that the highest DI rank node will have $RI = 1$. The RI is important to understand the dominance of a particular node over other nodes in influencing voltage change. In case there is not a very large difference in RI values of rank one and lower rank DI nodes, we can provide voltage control using any of the DI nodes. Otherwise, in situations of large RI variation among DIVF ranks, it is preferable to place the

control source on the rank-one DI. Therefore, the DI ranking must be interpreted along with RI values. The Fig. 1 provides details of using the proposed CFPF for both applications and the advantages associated.

Remark. A significant benefit of the space-wise approximation is that the resulting closed-form expression (5) is valid in a region of load/injection, which is less conservative compared to the neighborhood region where conventional linearization method works. The explicit form allows for fast computation of the voltage in terms of power injections, so solving nonlinear power flow problem is avoided. Moreover, for a trained analytic solution within a loading subspace, we have more freedom to allow different combinations of uncertain load injections. This flexibility makes the analytic solution appealing when multiple scenarios, such as those arising due to renewable injection types and locations, are considered. This implies all the variance calculations in (11) can be performed based on only single model, without retraining. This will save a lot of computational efforts compared to the MCS while preserving the accuracy and nonlinearity capturing advantages of MCS. Further, the non-parametric nature of the proposed approximation means that we can study the effect of different types of DG distributions on DIVF.

V. RESULTS AND DISCUSSION

This section contains simulation results and discussions. In addition to approximation error, discussions have been done on features of the proposed CFPF method such as interpretability, predicted standard deviation as inherent error indicator, etc. First, we start by describing the performance evaluation of CFPF when compared with NRLF. Later we also show the performance of CFPF on applications of VP and DIVF against multiple different distributions of input uncertainty.

The simulations have been presented on two different radial distribution test systems under various conditions of active operation. One is the 33-bus system [37], and another is the 56-bus system, which is modified from the IEEE 123-bus system [39]. The 56-bus has two variants 1) 56-bus OH with overhead lines, and 2) 56-bus UG with underground cables [39]. They differ in r and x of network branches, thus providing different r/x ratio cases with other properties being constant.

A. Performance Evaluation

Table I contains the worst-node mean absolute error (MAE) for voltage and angle under various cases. The MAE in voltage at j^{th} bus is $\xi_V(j) = \left(\sum_{i=1}^m |\hat{V}_j(i) - V_j(i)| \right) / m$ using m testing samples. Similarly, MAE in angle $\xi_\theta(j)$ is also defined. The worst-node MAE is the maximum MAE value among all the node measurements of voltage and angle.

As expected, in a larger subspace, with increasing δ , closed-form approximations have higher error. Nevertheless, even with $\delta = 0.3$, the MAE values are minimal and well within the tolerable range. The worst-node MAE values for the 56-bus OH system are significantly lower than MAE values associated with 33-bus and 56-bus UG system. The average r/x ratio of 56-bus OH is much lower (0.6345) than the other two systems

Fig. 2. Strong proportional relationship between Mean Absolute Error (MAE) with k_{LN} and voltage variance at a particular node. The graph is drawn with the same results as shown in Table I for 33-Bus system and $\pm 10\%$ uncertainty.

(1.443 for 33-bus and 2.099 for 56-bus UG). This difference in network parameters is one of the possible reasons affecting the approximation quality as voltage nonlinearity depends on the r/x ratio.

A deeper analysis in MAE results in Table I helps to identify the limitations and features of different forms of CFPF approximation. A significant observation is that in most of the simulation instances, worst-node MAE is at the node which has minimum voltage magnitude value i.e., farthest from the source or voltage-controlled slack bus (#18 in 33-bus and #32 in 56-bus system). We further look into the reason of having different nodes with maximum error with various covariance functions. The linear kernel k_{LN} approximates power flow (PF) as a D -dimensional plane in load subspace. However, due to PF nonlinearity, there is significant error in such linear forms. Further, assessing the extent of nonlinearity of power flow is challenging, particularly in high dimensions. Nevertheless, variance can be indicative of voltage function variability, within a load subspace. With Fig. 2 clearly, MAE in node voltage approximation, with k_{LN} , is strongly correlated with variance at that particular node voltage, within the subspace. This is because higher variance means there will be more voltage solutions at far distance from a linear approximation e.g. mean plane. The SE covariance function preserves the PF nonlinearity in CFPF. Therefore, the MAE in voltage approximation with k_{SE} does not show a correlated relationship with variance in Fig. 2. However, as $V_j(s)$ is a multi-dimensional nonlinear function with s , it is hard to approximate the function accurately. This nonlinearity effect gets aggravated at the edge nodes of radial networks as they are sensitive to load changes at all nodes. Thus, we get maximum MAE at tail nodes in radial network with k_{SE} and nonlinear kernel functions. This shows that nodes far from the source have higher nonlinearity and thus inferior approximation.

An important variation in Table I is type of covariance function. The minimum worst-case MAE is obtained with k_{RQ} or k_{QD} in case of voltage. For angle approximation, different covariance functions provide a minimum worst-node error. Importantly, various covariance functions (k_{QD} , k_{RQ} , and k_{SE}) lead to almost comparable worst-node MAE value for angle approximation. This similarity indicates that increased complexity will not lead to a better approximation of angle function. We below compare the performance of linear covariance function k_{LN} (k_p with $n = 1$) for approximating

TABLE I
WORST-NODE MEAN ABSOLUTE ERROR (MAE) IN VOLTAGE ξ_V AND ANGLE ξ_θ

System	% δ	$\max_j(\xi_V(j)) \times 10^{-6}$ (pu)				$\max_j(\xi_\theta(j)) \times 10^{-4}$ (rad)			
		k_{LN}	k_{QD}	k_{RQ}	k_{SE}	k_{LN}	k_{QD}	k_{RQ}	k_{SE}
33-bus	10	2.36(33)	1.28(18)	1.15(18)	1.24(18)	0.69(33)	0.32(18)	0.34(18)	1.25(16)
	20	5.95(33)	4.58(18)	4.37(18)	4.38(18)	3.52(33)	1.30(18)	1.39(18)	2.05(18)
	30	13.5(33)	10.4(18)	8.94(18)	10.4(18)	8.39(33)	3.14(18)	3.36(18)	2.84(17)
56-bus OH	10	1.02(32)	0.66(32)	0.62(32)	0.74(32)	0.46(32)	0.27(32)	0.27(32)	0.51(32)
	20	1.61(32)	1.11(32)	1.05(32)	1.06(32)	0.52(32)	0.36(32)	0.38(32)	0.35(32)
	30	3.71(32)	2.50(32)	2.45(32)	2.55(31)	1.24(32)	0.79(32)	0.88(32)	0.77(32)
56-bus UG	10	5.16(32)	2.81(32)	2.99(32)	3.32(32)	1.01(32)	0.93(32)	0.88(32)	0.80(32)
	20	18.3(32)	10.8(32)	11.1(32)	11.9(32)	3.79(32)	3.09(32)	2.86(26)	2.83(32)
	30	37.5(32)	26.6(32)	27.1(32)	29.7(32)	7.98(32)	6.99(32)	6.67(32)	6.93(32)

Learning with $N = 300$ and testing with $m = 5 \times 10^4$ in all the cases for uniform comparison basis.

Bold represent lowest worst-node MAE for a system with a given δ ; worst error node in brackets

Fig. 3. Relationship between average value of predicted STD $\sqrt{\sigma^2[V_j(s)]}$ (16) and voltage approximation MAE in 56-bus OH system

both voltage and angle.

Predicted Variance as Inherent Accuracy Indicator: In association with feature discussed in section III, Fig. 3 shows that the average value of predicted standard deviation (STD) $\sqrt{\sigma^2[V_j(s)]}$ (16) increases with an increase in the error for the 56-bus OH system. The trend is consistent over the other subspace sizes, and a similar trend is observed for other systems. Thus, we can evaluate the performance of approximation without requiring a vast number of ACPF calculations, by predicted STD value. This is more significant when the testing point is away from the base-load and eliminates the requirement of repeated brute-force calculation. In a nutshell, higher STD $\sigma[V_j(s)]$ means higher approximation error.

Performance with Mesh-Grid: The proposed CFPF framework does not distinguish between grid topology types (radial or mesh) or network parameters such as r/x ratio. The proposed CFPF approximation's quality for a mesh-grid network can be assessed by the MAE results (same as in Table I) presented in Table II. These results are obtained by training over 300 power flow samples, and testing is done with $m = 5 \times 10^4$ samples of MCS for IEEE 30-bus system [37]. All the PQ buses with non-zero loads are considered to have uncertain injections. The values are in order of 10^{-5} suggesting that the proposed method is highly accurate in approximating the CFPF for the mesh-grid network as well. Further, we have observed that node 26 has the highest approximation error. One possible reason is that voltage at node 26 has highly variable sensitivity regarding load change at the same node. A possible explanation is that node 26 is the only non-zero load node in the 30-bus system where only one line is connected.

TABLE II
MAE FOR FOR MESH-GRID 30-BUS SYSTEM WITH k_{SE}

Distribution	Normal-II	Beta	Laplace	Mixed
$\xi_V(j) \times 10^{-5}$	4.00 (26)	4.54 (26)	5.12 (26)	4.67 (26)

A detailed study on mesh-grids is out of scope for this work – as this paper focuses on radial topology features– and will be explored in future works.

B. Interpretability

The closed-form approximation's interpretability is one of the fundamental advantages of the proposed method. The *hyper-parameters* ϕ_j contains information about the function behaviour as discussed in section III. The MAE values in different node voltage predictions are plotted against the ratio of *characteristic-length* l_j and *scaling factor* τ_j from ϕ_j in Fig. 4. The ratio l_j/τ_j , defined as quality ratio (QR) in section III for (8), indicates quality of approximation of a function. The inverse relationship in Fig. 4 means that higher value of l_j/τ_j implies lower error. From properties of covariance function k_{SE} , we can infer that a function $V_j(s)$ with a higher value of l_j/τ_j moves 'slowly' and 'remains close' to its mean in comparison to the other functions. This low 'wiggly' behavior can be interpreted as an *qualitative indication* of a function being less nonlinear or smoother. Further, when the QR is studied with the node index, it is observed that QR decreases when moving away from the source. The lowest QR value is observed on nodes that are farthest from the source in radial systems, implying that node voltages far from the source have higher nonlinearity than those close to the source. This provides a basis to develop causality analysis for power flow explainability further and will be explored in future.

C. Violation Probability Calculation and Analysis

The violation probability (VP) calculation is performed for different types of injection distributions (details in appendix B) to show the non-parametric feature of the CFPF. The GP-based CFPF is obtained using k_{SE} (8) with $\% \delta = 50$ and then same model is used for all the calculations presented in this subsection. The VP is calculated with voltage limits of 0.92 and 1.04 pu. Table III present results of error with five different distributions of net-load uncertainty in both real and reactive

TABLE III
APPROXIMATION AND VP ESTIMATION ACCURACY WITH VARIOUS INPUT
DISTRIBUTIONS FOR 33-BUS SYSTEM WITH k_{SE}

Distribution	$\max_j (\xi_V(j))$ pu	$\max_j \varrho_j^m - \varrho_j $
Normal-I	2.15E-05(18)	0.0026 (29)
Normal-II	1.86E-05(18)	0.0018 (12)
Beta	2.39E-05(18)	0.0028 (29)
Laplace	2.29E-05(18)	0.0029 (29)
Mixed	1.74E-05(18)	0.0009 (11)

power load. We consider all PQ node to have uncertain net-load, i.e. 100% actor nodes, similar to concept of 50% actor nodes considered in [20]. The ϱ_j^m indicate the VP calculated using MCS method via 5×10^4 samples and ϱ_j indicate VP calculated via proposed method. The maximum mean absolute error (MAE) values (similar to the once presented in table I) presented in table III are still of order of 10^{-5} , although the uncertainty is much higher then cases presented in Table I. Most importantly, the error in table III does not depend on type of uncertain injection distribution, showing non-parametric nature of proposed analytical power flow solution.

Performance with Correlated Uncertain Injections: Till now, the case studies establish the proposed CFPF's non-parametric feature where the analytical solution can be obtained for various distributions without retraining and with high accuracy. Moving forward, we consider correlated uncertain injections at various nodes. Fig. 5 shows the absolute error observed in VP calculation via proposed CFPF when compared with MCS on 56-bus OH system for correlates uncertain loads. Similar to the 33-bus system, the error is in 10^{-3} order and does not exceed 3×10^{-3} for any node. Further, the figure presents the results with two different correlations set for real and reactive power uncertainties, following Normal-II distributions (details are in Appendix B). It is clear that the error does not increase with the increase in correlation between uncertain node loads. This shows that the proposed closed-form *voltage-power* relationship is accurate even when uncertain net-loads or injections are correlated.

These VP calculation results also establish that proposed CFPF obtains the output voltage distribution with higher accuracy near tail as accuracy in VP calculation is directly related to tail estimation accuracy. Therefore, it is straightforward to use proposed analytical power flow expressions (7)-(9) for uncertainty quantification in probabilistic power flow.

Fig. 4. Interpretability of MAE ξ_V , with quality ratio l/τ of different covariance functions in 33-bus system. Inverse relationship shows that variance $V_j(s)$ with higher value of QR varies less and is better approximated.

D. DIVF Ranking

This sub-section presents the DIVF ranking and RI results for different node voltages in 33-bus system and 56-bus OH system. The DIVF ranking is done considering total 50% renewable penetration compared to total base-load. All even number buses have DGs connected and are of same rated capacity. Thus, the rated DG capacity for any system is obtain by dividing 50% of total load with number of even index nodes in the system. Thus, for 33-Bus the rated DG capacity is 116.1kW and 16 nodes have DG connected. For 56-Bus system individual DG capacity is 64.3kW and there are total 28 RES in the system. The information of distributions is given in appendix B. For both the systems, we use same CFPF model used for VP calculation results in sub-section V-C i.e. 500 training samples and k_{SE} covariance function (8).

Firstly, we highlight the DIVF ranking dependency on the type of distribution in Table IV. It is evident that for both node 11 and node 32, proposed method has been able to provide DIVF rank accurately. Further, The DIVF rank depends on the type of distribution. This is due to the difference in loading conditions of the different nodes also as the PF nonlinearity is not same at all loading conditions. By analysing results, we also found that nodes which are far from the source or slack bus, have higher DIVF ranking in general. This phenomena is significant and indicate that voltage control is more effective when far end nodes are supplied with voltage control sources. The reason for this lies in the nature of radial systems where far end buses influence power losses and voltage drop significantly. In Table V we present the DI rank and RI values for DG nodes influencing V_5 in 33-bus system. The results signifies the need to RI based interpretation of DIVF rank as multiple nodes have similar RI values and rank is decided by small margins of RI. This means that controller at node 32 will be as effective (almost) as at node 26. This case study also highlights that proposed method may provide different ranking index when RI values are close to each other. Nevertheless, the proposed method is helpful in DI ranking interpretation with RI. Now, we discuss the comparative performance and computational results of the proposed method.

Comparison with Existing Techniques: In this part, the proposed work is compared with existing power flow solution approximation and VP calculation methods. In [17], the data-

Fig. 5. Error in VP calculation $|\varrho_j^m - \varrho_j| \times 10^{-3}$ for two different correlated Normal-II distribution for 56-bus OH system. The legend is correlation coefficient for real and reactive powers in bracket. The maximum MAE error in training for uncertain load subspace with $\delta = 50\%$ is $9.6E-06$ among 5×10^4 samples and using k_{SE} as in (8).

driven approximation of voltage-power relationship is made in load space of $\pm 20\%$ of total active load variation and constrained power factor. For the 33-bus system, [17] reported MAE in voltage and angle to be of 1.0×10^{-5} and 1.9×10^{-3} , respectively, over considered load subspace. Using k_{LN} and considering the generic construction of subspace $\mathcal{L}$, described in section II-A, the worst-node MAE for 33-bus system is obtained as 5.95×10^{-6} in voltage and 3.52×10^{-4} in angle with $\delta = 20\%$. Further, the model-based approximation presented in [39] report average absolute errors respectively of 0.0041 and 0.0097 in voltage and angle at nominal load (Table 1 [39]). These values are much higher than the error values obtained by the proposed method using linear covariance. Therefore, the proposed method can approximate linear CFPF with higher accuracy than data and model-based approximations in a more general subspace of load.

In comparison to Latin Hypercube Sampling (LHS) based MCS, we observe that the proposed method is more accurate in VP estimation as shown in Fig. 6. The main reason is that VP calculation accuracy relies on the method's ability to approximate the output voltage distribution's center and tail ends. The reduced number of samples in LHS hampers the quality of tail approximation in MCS considerably. Therefore, the proposed CFPF is more suitable to perform complete uncertainty quantification on power flow. Further, the maximum absolute difference in MCS and proposed methods VP estimation is less than 0.0030 in all the cases presented in Table III. The work [20] record error of 0.016 With 50% actor nodes. Whereas, with all the load buses have uncertain load i.e. 100% actor nodes, the proposed method obtain one order of magnitude less violation probability error as reported in in Table III.

Computational Performance: In Table VI we report the computational time taken in various tasks by proposed method and MCS. The time reported here is of execution of un-optimized codes on MATLAB 2020b running on a PC with Intel Xeon E5-1630v4@3.70 GHz, 16 GB RAM. It is clear from results in Table VI that proposed method is much faster than MCS based conventional method. Further, with detailed bifurcation of proposed method execution time, the estimation time for 0.5 million samples less than 1 sec. for both 33-bus and 56-bus OH system. There is no effect of system size as the voltage calculation via CFPF is just a function evaluation, where dimension of input does not alter time significantly. Also, this means that due to non-parametric and subspace-wise approximation features, a trained CFPF model can calculate

TABLE V
DI AND RI FOR V_5 IN 33-BUS SYSTEM WITH *Beta* DISTRIBUTION

DI Rank	MCS		Proposed	
	Node	RI_j	Node	RI_j
1	26	1.0000	26	1.0000
2	16	0.9764	32	0.9728
3	32	0.9758	28	0.9625
4	28	0.9658	16	0.9620
5	6	0.9557	6	0.9534
6	8	0.9372	8	0.9333

Fig. 6. Error in VP calculation $|\varrho_j^m - \varrho_j| \times 10^{-3}$ for Normal-II case on 33-Bus system. It shows error using proposed CFPF and Latin Hypercube Sampling (LHS) against MCS (5×10^4 samples). GP model training settings are same as in Table I with k_{SE} in (8).

voltage for 0.5 million uncertain load inputs of any distribution in less than a second. The DIVF calculations times are also listed in the table, with 10^4 load samples used for calculating DI of each input node thus total $D \times 10^4$ iterations. The increase in the t_e from DIVF for 56-bus is due to higher number of random DG injections ($D=28$) and thus requiring 28×10^4 iterations instead of 16×10^4 as in case of 33-bus system. Yet it is clear that for both of the applications that proposed method is computationally light.

VI. CONCLUSION

The paper proposes a general closed-form approximation method for obtaining voltage as a function of effective load vector. The proposed mapping between net-load vector and node voltage is defined in analytical function form using GP to model the relationship with different covariance functions. The paper qualitatively analyzes the relative nonlinearity and variation of node voltage functions. A set of features of the proposed method has been discussed and shown to be promising through simulations. We present two applications of the proposed analytical power flow solution in calculating violation probability calculation and dominant influencer of voltage fluctuation with relative importance index. The methods to solve both of these problems are shown to be non-parametric and computationally light. These methods consider simultaneous presence of multiple uncertainties in ADNs with RES and EV integration. The work also attempts to understand the internal working of the learning mechanism. This provides the path for improvement in learning along with other possibilities, like developing interpretations of mesh-grid systems, and multi-phase distribution systems. The potential applications of proposed approximation include voltage control, state estimation, and bad data identification, to name a

TABLE IV
DIVF WITH DIFFERENT DG DISTRIBUTIONS IN 56-BUS OH SYSTEM

DI Rank	Node 11				Node 31			
	Beta		Normal		Beta		Normal	
	M	P	M	P	M	P	M	P
1	30	30	32	32	30	30	32	32
2	34	34	24	24	32	32	30	30
3	16	16	20	20	28	28	28	28
4	28	28	18	18	20	20	24	24
5	38	14	28	16	22	22	20	20

Training with $N = 500$; M: MCS; P: Proposed

TABLE VI
COMPUTATION TIME FOR VARIOUS CASE STUDIES

System Description	Proposed (s) $\{t_s, t_t, t_e\}; total$	MCS (s)
Single node voltage distribution and VP calculation for 33-bus system (5×10^4 Samples)	$\{1.04, 1.66, 0.70\}; 3.4$	105.24
Single node voltage distribution and VP calculation for 56-bus system (5×10^4 Samples)	$\{1.37, 2.06, 0.72\}; 4.15$	140.88
DIVF for a node in 33-bus System ($D = 16$)	$\{1.12, 1.72, 2.4\}; 5.24$	359.11
DIVF for a node in 56-bus System ($D = 28$)	$\{1.42, 2.12, 4.7\}; 8.24$	715.76

Times t_s :PF solving ($N = 500$); t_t : Training; t_e : Estimation

few. These will be explored in future works. The proposed framework will also be modified to include network graph as input in future works.

ACKNOWLEDGEMENT

This work was supported in part by the Nanyang Technological University SUG, in part by the Academic Research Fund TIER 1 2019-T1-001-119 (RG 79/19), and in part by the Energy Market Authority (EMA) and National Research Foundation (NRF) Singapore under Grant EMA-EP004-EKJGC-000 and in part by NRF Distributed Energy Resource Management System for Energy Grid 2.0

REFERENCES

- [1] Y. Wang, N. Zhang, Q. Chen, J. Yang, C. Kang, and J. Huang, "Dependent discrete convolution based probabilistic load flow for the active distribution system," *IEEE Trans. on Sustainable Energy*, vol. 8, no. 3, pp. 1000–1009, 2017.
- [2] G. Grusso, R. S. Netto, L. Daniel, and P. Maffezzoni, "Joined probabilistic load flow and sensitivity analysis of distribution networks based on polynomial chaos method," *IEEE Trans. on Power Systems*, vol. 35, no. 1, pp. 618–627, 2020.
- [3] M. M. A. Abdelaziz, "Opencil-accelerated probabilistic power flow for active distribution networks," *IEEE Trans. on Sustainable Energy*, vol. 9, no. 3, pp. 1255–1264, 2018.
- [4] F. Ding and B. Mather, "On distributed pv hosting capacity estimation, sensitivity study, and improvement," *IEEE Trans. on Sustainable Energy*, vol. 8, no. 3, pp. 1010–1020, 2016.
- [5] M. S. S. Abad *et al.*, "Probabilistic assessment of hosting capacity in radial distribution systems," *IEEE Trans. on Sustainable Energy*, vol. 9, no. 4, pp. 1935–1947, 2018.
- [6] A. Valibeygi, S. A. R. Konakalla, and R. d. Callafon, "Predictive hierarchical control of power flow in large-scale pv microgrids with energy storage," *IEEE Trans. on Sustainable Energy*, vol. 12, no. 1, pp. 412–419, 2021.
- [7] B. Stott, J. Jardim, and O. Alsac, "Dc power flow revisited," *IEEE Trans. on Power Systems*, vol. 24, no. 3, pp. 1290–1300, 2009.
- [8] F. Li and R. Bo, "Dcopf-based lmp simulation: algorithm, comparison with acopf, and sensitivity," *IEEE Trans. on Power Systems*, vol. 22, no. 4, pp. 1475–1485, 2007.
- [9] T. J. Overbye, X. Cheng, and Y. Sun, "A comparison of the ac and dc power flow models for lmp calculations," in *37th Annual Hawaii International Conference on System Sciences*. IEEE, 2004.
- [10] M. Vaiman *et al.*, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Trans. on Power Systems*, vol. 27, no. 2, 2012.
- [11] J. W. Simpson-Porco, "Lossy dc power flow," *IEEE Trans. on Power Systems*, vol. 33, no. 3, pp. 2477–2485, 2017.
- [12] Z. Yang *et al.*, "A general formulation of linear power flow models: Basic theory and error analysis," *IEEE Trans. on Power Systems*, vol. 34, no. 2, pp. 1315–1324, 2019.
- [13] K. Dvijotham and D. K. Molzahn, "Error bounds on the dc power flow approximation: A convex relaxation approach," in *55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 2411–18.
- [14] D. K. Molzahn, I. A. Hiskens *et al.*, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends® in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.
- [15] H. Mohsenian-Rad *et al.*, "Guest editorial theory and application of pmus in power distribution systems," *IEEE Trans. on Smart Grid*, vol. 11, no. 1, pp. 723–725, 2020.
- [16] J. Yu, Y. Weng, and R. Rajagopal, "Robust mapping rule estimation for power flow analysis in distribution grids," in *NAPS*. IEEE, 2017.
- [17] Y. Liu, N. Zhang, Y. Wang, J. Yang, and C. Kang, "Data-driven power flow linearization: A regression approach," *IEEE Trans. on Smart Grid*, vol. 10, no. 3, pp. 2569–2580, 2018.
- [18] Y. Liu, Y. Wang, N. Zhang, D. Lu, and C. Kang, "A data-driven approach to linearize power flow equations considering measurement noise," *IEEE Trans. on Smart Grid*, 2019.
- [19] Y. Tan, Y. Chen, Y. Li, and Y. Cao, "Linearizing power flow model: A hybrid physical model-driven and data-driven approach," *IEEE Trans. on Power Systems*, vol. 35, no. 3, pp. 2475–2478, 2020.
- [20] S. Munikoti *et al.*, "Probabilistic voltage sensitivity analysis to quantify impact of high pv penetration on unbalanced distribution system," *IEEE Trans. on Power Systems*, pp. 1–1, 2021.
- [21] K. Jhala, B. Natarajan, and A. Pahwa, "The dominant influencer of voltage fluctuation (divf) for power distribution system," *IEEE Trans. on Power Systems*, vol. 34, no. 6, pp. 4847–4856, 2019.
- [22] H. Sheng, J. Xiao, Y. Cheng, Q. Ni, and S. Wang, "Short-term solar power forecasting based on weighted gaussian process regression," *IEEE Trans. on Industrial Electronics*, vol. 65, no. 1, pp. 300–308, 2017.
- [23] Y. Yang, S. Li, W. Li, and M. Qu, "Power load probability density forecasting using gaussian process quantile regression," *Applied Energy*, vol. 213, pp. 499–509, 2018.
- [24] P. Pareek and H. D. Nguyen, "Gaussian process learning-based probabilistic optimal power flow," *IEEE Transactions on Power Systems*, vol. 36, no. 1, pp. 541–544, 2021.
- [25] Z. Hu *et al.*, "A bayesian approach for estimating uncertainty in stochastic economic dispatch considering wind power penetration," *IEEE Trans. on Sustainable Energy*, vol. 12, no. 1, pp. 671–681, 2021.
- [26] Y. Xu, Z. Hu, L. Mili, M. Korkali, and X. Chen, "Probabilistic power-flow calculation based on a novel gaussian process emulator," *IEEE Trans. on Power Systems*, 2020.
- [27] Y. Xu *et al.*, "A data-driven nonparametric approach for probabilistic load-margin assessment considering wind power penetration," *IEEE Trans. on Power Systems*, 2020.
- [28] P. Pareek and H. D. Nguyen, "Probabilistic robust small-signal stability framework using gaussian process learning," *Electric Power Systems Research*, vol. 188, 2020.
- [29] J. J. Grainger and W. D. Stevenson, *Power System Analysis*, 2003.
- [30] D. Mehta, H. D. Nguyen, and K. Turitsyn, "Numerical polynomial homotopy continuation method to locate all the power flow solutions," *IET Gen., Trans. & Dist.*, vol. 10, no. 12, pp. 2972–2980, 2016.
- [31] D. K. Molzahn, "Computing the Feasible Spaces of Optimal Power Flow Problems," *IEEE TPWRS*, vol. 32, no. 6, pp. 4752–4763, Nove 2017.
- [32] B. Cui and X. A. Sun, "Solvability of power flow equations through existence and uniqueness of complex fixed point," *arXiv preprint arXiv:1904.08855*, 2019.
- [33] H. D. Nguyen and K. S. Turitsyn, "Appearance of multiple stable load flow solutions under power flow reversal conditions," in *2014 IEEE PESGM*. IEEE, 2014, pp. 1–5.
- [34] H. D. Nguyen and K. Turitsyn, "Voltage multi-stability in distribution grids with power flow reversal," in *2016 IEEE International Conference on Sustainable Energy Technologies (ICSET)*. IEEE, 2016, pp. 19–24.
- [35] Z. A. Memon, R. Trinchero, Y. Xie, F. G. Canavero, and I. S. Stievano, "An iterative scheme for the power-flow analysis of distribution networks based on decoupled circuit equivalents in the phasor domain," *Energies*, vol. 13, no. 2, p. 386, 2020.
- [36] P. Maffezzoni and G. Grusso, "Complex-array-operation newton solver for power grids simulations," *IEEE Access*, vol. 8, pp. 47984–47992, 2020.
- [37] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "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.
- [38] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

- [39] 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.
- [40] E. Snelson, Z. Ghahramani, and C. E. Rasmussen, "Warped gaussian processes," in *Advances in neural information processing systems*, 2004, pp. 337–344.
- [41] G. Ruan *et al.*, "Review of learning-assisted power system optimization," *arXiv preprint arXiv:2007.00210*, 2020.
- [42] L. Wang, M. Yuan, F. Zhang, X. Wang, L. Dai, and F. Zhao, "Risk assessment of distribution networks integrating large-scale distributed photovoltaics," *IEEE Access*, vol. 7, pp. 59 653–59 664, 2019.
- [43] G. Guosso, R. S. Netto, L. Daniel, and P. Maffezzoni, "Joined probabilistic load flow and sensitivity analysis of distribution networks based on polynomial chaos method," *IEEE Trans. on Power Systems*, vol. 35, no. 1, pp. 618–627, 2020.

- **Mixed:** Randomly mixed Beta and Laplace distributions in equal proportion of samples.

The uncertainty distributions used in results in table IV and V are constructed based on rated DG capacity for real power. The DG inverter is taken to be supplying reactive power at fixed power factor (0.95).

- **Normal:** 50% of rated DG capacity as Mean and one by sixth of rated DG capacity as standard deviation.
- **Beta:** Shape parameters (18,3) and interval 0.1 to 1 times of rated DG capacity.

APPENDIX A EXPRESSIONS OF KERNEL MATRIX, VECTOR AND PREDICTIVE VARIANCE

We define the *design matrix* $X \in \mathbb{R}^{D \times N}$ which contains N training input vectors of $D \times 1$ dimension each, as

$$X = [\mathbf{x}^1 \dots \mathbf{x}^i \dots \mathbf{x}^N] \quad (13)$$

The kernel matrix $K = K(X, X)$ is an $N \times N$ matrix showing the relationship between the training samples.

$$K(X, X) = \begin{bmatrix} k(\mathbf{x}^1, \mathbf{x}^1) & \dots & \dots & k(\mathbf{x}^1, \mathbf{x}^N) \\ \vdots & \ddots & \ddots & \vdots \\ k(\mathbf{x}^N, \mathbf{x}^1) & \dots & \dots & k(\mathbf{x}^N, \mathbf{x}^N) \end{bmatrix}. \quad (14)$$

The kernel vector $\mathbf{k} = k(X, \mathbf{x}^*)$ is an $N \times 1$ column vector and establishes the relationship between training matrix X and variable vector $\mathbf{x}^*$.

$$\mathbf{k}(X, \mathbf{x}^*) = [k(\mathbf{x}^1, \mathbf{x}^*) \dots k(\mathbf{x}^N, \mathbf{x}^*)]^T. \quad (15)$$

Similar to the predictive mean (4), GP regression also provides analytical expression for predictive variance for any variable vector $\mathbf{x}^*$ as (chapter 2, function-space view [38])

$$\sigma^2[f(\mathbf{x}^*)] = k(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{k}^T (K + \sigma_e^2 I)^{-1} \mathbf{k} \quad (16)$$

with definitions of kernel vector and matrix same as described before. The variance (16) depends on the kernel vector and does not depend on training target vector. Implies that variance in output is described based on input and *design matrix*.

APPENDIX B DETAILS OF UNCERTAINTY DISTRIBUTIONS

Here, we list the parameters of the uncertainty distributions constructed for simulation results for reproducibility of results.

All the uncertainty distributions used in results in Table III are uncorrelated and constructed for real and reactive power separately using same distribution parameters.

- **Normal-I:** Base-load as mean and 10% of base-load as standard deviation.
- **Normal-II:** Base-load as mean and 20% of base-load as standard deviation.
- **Beta:** Shape parameters (18,3) and interval 0.1 to 1.1 times of base-load.
- **Laplace:** Location parameter as base-load and scale parameter as 0.5% of base-load.