A Rapid Method for Impact Analysis of Grid-edge Technologies on Power Distribution Networks

Feng Li, *Student Member, IEEE*, Ilhan Kocar, *Senior Member, IEEE*, and Antoine Lesage-Landry, *Member, IEEE*

***Abstract*—This paper presents a novel rapid estimation method (REM) to perform stochastic impact analysis of grid-edge technologies (GETs) to the power distribution networks. The evolution of network states' probability density functions (PDFs) in terms of GET penetration levels are characterized by the Fokker-Planck equation (FPE). The FPE is numerically solved to compute the PDFs of network states, and a calibration process is also proposed such that the accuracy of the REM is maintained for large-scale distribution networks. The approach is illustrated on a large-scale realistic distribution network using a modified version of the IEEE 8500 feeder, where electric vehicles (EVs) or photovoltaic systems (PVs) are installed at various penetration rates. It is demonstrated from quantitative analyses that the results from our proposed approach have negligible errors comparing with those obtained from Monte Carlo simulations.**

***Index Terms*—Grid-edge, electric vehicles, power distribution networks, stochastic analysis, Fokker-Planck equation, probability density function, Monte Carlo**

I. Introduction

GRID-edge technologies (GETs), e.g., plug-in electric vehicles (EVs), rooftop photovoltaic systems (PVs), demand response (DR) programs, etc., are being installed at customers' sites (edges of the grid) from which they are connected to the power distribution networks. With the on-going energy transition, the number of GETs is rapidly growing. It becomes critical to assess their influence on the distribution network in order to maintain system reliability and power quality, and to avoid service interruptions. As changes to customers' power demands (*load offset*) due to the use of GETs are behind-the-meter and are stochastic in terms of customers' usage behaviors, it is difficult to evaluate their impacts especially at high penetration rates. For example, EV charging may introduce undesired impacts to power networks such as overloading of key equipment, severe voltage variations, phase unbalancing, harmonic distortions, etc. [1], [2]. Various uncertain factors such as charging locations of EVs, charging time and duration, charging power, and battery capacities, must be properly modeled to perform an impact analysis to evaluate the network states with different penetration levels of EVs installed. The analysis should be extended to other GETs than EVs, and combined impacts should be studied when multiple types of GETs are installed to the same power network.

In the literature, deterministic methods have been proposed to study the impacts of EVs [3]–[6] and PVs [7]–[9], and their combined impacts [10] to the power distribution networks. These methods normally evaluate network states by generating typical or worst-case scenarios from the parameters with uncertainties, e.g., power flow analysis is usually used to study the impacts in each scenario. While these deterministic simulation models are easier to implement, uncertainty is often ignored when looking at selected scenarios. Analysis results may be consequently inaccurate and cannot provide meaningful insights for power system plannings. It is important to model the uncertainty in the input variables such as the locations and the charging/generation profiles of EVs/PVs, etc., and use stochastic approaches to compute the network states and their associated probability density functions (PDFs).

Such stochastic approaches can be categorized into numerical, analytical, and approximation approaches. Due to the simplicity in the implementation, several Monte Carlo simulation-based studies have been performed in the context of GET integration. Monte Carlo simulations can provide accurate results to complicated and/or non-linear systems with many random variables involved. For example, in [11], impacts of voltage drops and loading conditions of lines due to EV charging are studied, and in [12], network losses, voltage variations and transformer loadings are studied at various penetration levels of EVs. In [13], [14], the harmonic impacts due to EV charging are analyzed, where not only locations of EVs and charging patterns are randomized, but also the operating states of the residential household loads. Finally, impacts of multiple technologies including EVs and PVs to the distribution network in terms of abnormal voltages and transformer loading levels are studied in [15].

While the Monte Carlo simulation approach is a straightforward numerical method for the stochastic analysis and can provide accurate results if all uncertain variables are covered, it has a very slow rate of convergence. The construction of a large number of samples is time-consuming, and performing a power flow analysis to each sample is computationally expensive. Analytical approaches are, therefore, an alternative to derive PDFs of output variables from PDFs of input variables, such as in probabilistic power flow analysis. For example, the PDF of an output random variable can be computed from the convolution of input random variables which are independent. However, the convolution operation requires extensive computation, even with the discrete Fourier trans-

This work was funded by Eaton Corporation and the Institute for Data Valorization (IVADO).

Feng Li is with the department of Electrical Engineering, Polytechnique Montreal and with GERAD & Mila, Montreal, Quebec, Canada, H3T 1J4. He is also with Eaton's CYME International T&D, Brossard, Quebec, Canada J4Z 0N5 (e-mail: feng.li@polymtl.ca).

Ilhan Kocar is with the department of Electrical Engineering of The Hong Kong Polytechnic University, Hong Kong SAR, China (e-mail: ilhan.kocar@polyu.edu.hk).

Antoine Lesage-Landry is with the department of Electrical Engineering, Polytechnique Montreal and with GERAD & Mila, Montreal, Quebec, Canada, H3T 1J4 (e-mail: antoine.lesage-landry@polymtl.ca).

form (DFT) applied to reduce the computational burden [16]. Approximation methods for probabilistic power flow analysis are developed to improve the computational efficiency [17]–[23]. In [17], PDFs of power flows in transmission lines are approximated in terms of *cumulants*, which are an alternative set of quantitative measures to *moments* related to the shape of PDFs with the computational time significantly reduced. Another approximation method is the point estimation method which approximates the PDF of an output variable from its first and second moments as in [18]. In recent years, polynomial chaos expansion (PCE)-based methods have shown promising results in quantifying uncertainties in power systems; however, limitations still exist which make them less applicable to stochastic analysis of large-scale distribution systems. For example, it suffers from the "curse of dimensionality" [19] when the number of random input variables is large. Hence, the computation for the polynomial coefficients of the surrogate model is challenging even if the expansion degree is limited to 2 [20]. Also, separate approaches are required to accompany the PCE method such as using the Copula theory in [21], [22] when random variables are correlated and certain constraints are to be considered, and using the Stieltjes procedure to construct the orthogonal basis for PCE such as in [23] when random variables have arbitrary probability distributions. Although such advancements have been made to characterize uncertainties in power systems, a common drawback of the analytical and approximated methods in the literature is that the models are developed for a fixed penetration rate. To consider multiple penetration rates of GETs, models will have to be recomputed, which is inefficient.

To bridge the gap, our objective is to develop a model which can characterize the evolution of network state PDFs in terms of the uncertainties of GETs while the penetration rate is taken as a parameter. Such a model can be derived from the Fokker-Planck Equation (FPE). In [24], an advection equation which is a simplified FPE is used to describe the evolution of temperature PDFs for a group of thermostatically controlled loads (TCLs). In [25], the FPE is modeled and is solved to describe the probability of power network stability over time under perturbations.

In this paper, we propose a rapid estimation method (denoted as REM hereinafter) based on FPE to perform a stochastic analysis on the impact of EVs and PVs to power distribution networks. The REM approach provides a rapid sweep of the network states at a wide range of penetration levels. It also indicates concentrated areas where GETs are installed on the network, and network sections susceptible to abnormal conditions, e.g., equipment overloading and under/over-voltages, at any given penetration level. Results from the REM can also identify the maximum penetration level possible that can be hosted by each network section for the network to operate securely. Conversely, the REM approach permits to determine the penetration level at which network expansion and/or mitigation actions are required.

The rest of the paper is organized as follows: in Section II the REM approach to perform a stochastic impact analysis based on FPE is presented. The solution to the FPE model characterizes the evolution of equipment loading levels and voltage levels on the network with the GET penetration rate. In Section III, a numerical method to solve the FPE model is introduced, and a calibration method to improve the accuracy of the solution is proposed. Section IV demonstrates the performance of the REM by evaluating the impacts of EVs on the modified IEEE-8500 test feeder. Finally the conclusion is made in Section V along with some future work directions pointed out.

II. Stochastic impact analysis model

In this section, the REM model for the stochastic impact analysis of EVs/PVs is proposed.

A. Assumptions

We assume that the following information is provided. This information can often be obtained from statistical surveys or socio-economic analysis [26]–[28].

- Quantity probability $\Pr^{i}_{\text{num},k}(n)$ which specifies the probabilities of any customer $i = 1, 2, \cdots$ of having $n = 0, 1, \cdots$ GET k devices where $k \in \{\text{EV, PV}\}$.
- A set of load offset profiles $\mathcal{L}_k$ for GET k, $\mathcal{L}_k = \{l^j_k(t), j = 1, 2, \cdots\}$, where $l^j_k(t)$ is an EV charging or a PV generation profile, both at time t. The diversity of $\mathcal{L}_k$ characterizes different usage patterns of EVs and PVs.
- Adoption probability $\Pr^{i}_{\mathcal{L},k}(j),\ j = 1, 2, \cdots$. for customer i to adopt a load offset profile j in $\mathcal{L}_k$.

Based on the given data, the REM approach is formulated in the next section which provides PDFs of network states (equipment loading level and node/bus voltage level) at any penetration level of EVs and PVs without performing repetitive power flow analysis to the network.

In some work such as [29], the penetration rate is defined as the ratio of the number of GET devices in service to the total number of GET devices on the network. This definition is not applicable to our analysis because we do not know the total number of GET devices on the network. However, we do know the total number of customers/meters on the network. Hence in this paper, the penetration rate is defined as the ratio of the number of GET devices to the total number of customers/meters on the distribution network.

Definition 2.1 (Penetration rate): The penetration rate p_k for GET k is defined as

$$p_k \triangleq \frac{T_k}{N_m} = \frac{n_k + n^0_k}{N_m}, \tag{1}$$

where N_m is the total number of customers/meters on the network, T_k is the number of GET k devices, n^0_k is the number of GET k devices already existing on the network, and n_k is the number of devices that are required to add in order to reach p_k. The initial penetration rate of GET k of the network is therefore $p^0_k = n^0_k/N_m$. Note that the definition is somewhat similar to that adopted in [30], but we extend it to consider the case in which some GET k devices may already exist on the network (i.e., $p^0_k > 0$). When simultaneous impacts of multiple GET types are studied, a separate penetration rate should be defined for each GET type. By this definition, p_k can exceed 100%, which indicates that *on average* each customer has more than one device of GET type k.

For simplicity, when referring to one GET type, the subscript k in the penetration rate is dropped from hereinafter, unless stated otherwise.

B. Stochastic model

Let $x \in \mathbb{R}$ denote a network state of interests, whether the loading level of a network equipment or the voltage level of a network node/bus. The following general model is proposed here to describe the network state when GET k devices are installed to the network with respect to p:

$$\mathrm{d}x = u(x,p)\mathrm{d}p + \sigma(x,p)\mathrm{d}W_p, \tag{2}$$

where W_p is a Wiener process indexed by p. The term $u(x,p)$ is the drift velocity of the network state under the influence of EVs/PVs, and $\sigma(x,p) > 0$ denotes the magnitude of the additive disturbances.

Due to the presence of the Wiener process, the trajectory $x(p)$ does not have an analytical form, but given an initial network state $x(p^0)$, one can generate a single trajectory by numerical integration of (2) [31]. To study the stochastic behavior, one can further generate a large set of trajectories and construct a histogram of x at each p to approximate the PDF of x at p, denoted as $m(x,p)$. This approach is similar to Monte Carlo which is computationally heavy. Therefore, we derive a method that can directly describe the propagation of $m(x,p)$ with the penetration rate. To obtain such a method, the Fokker-Planck equation [32] can be used:

$$\frac{\partial m(x,p)}{\partial p} + \frac{\partial}{\partial x}\Big\{m(x,p)u(x,p)\Big\} = \frac{\partial^2}{\partial x^2}\Big\{m(x,p)d(x,p)\Big\}, \tag{3}$$

subject to $m_0 = m(x,p^0)$. Here, $d(x,p) = \sigma(x,p)^2/2$ is the diffusion velocity. It can be generally assumed that the magnitude of additive disturbances to the network states due to GETs is much smaller than the magnitude of the drift velocity, i.e., $\sigma(x,p) \ll |u(x,p)|$. Hence, we assume that the diffusion velocity is a small constant, i.e., $d(x,p) \triangleq d, \forall x,p$. We can then rewrite (3) as

$$\frac{\partial m(x,p)}{\partial p} + \frac{\partial}{\partial x}\{m(x,p)u(x,p)\} = d\frac{\partial^2 m(x,p)}{\partial x^2}. \tag{4}$$

To solve (4) for the propagation of $m(x,p)$, the drift term $u(x,p)$ must first be computed. The term depends on what network state x is describing, either the loading level of a device/equipment or the voltage level of a bus/node. In the following sections, we describe how $u(x,p)$ can be computed based on the stochastic data of customer's quantity probability $\Pr^i_{\mathrm{num},k}(n)$, load offset profiles $\mathcal{L}_k$, and customer's adoption probability $\Pr^i_{\mathcal{L},k}(j)$.

C. Drift velocity of loading level

Let $m(x,p)$ refer to the PDF of equipment loading level x at penetration rate p. For equipment on distribution networks, the loading level is computed by-phase, hence we denote $x = z^\phi_{e,k}(p)$ as the loading level of equipment e on phase $\phi \in \{A,B,C\}$ when devices of GET k are to be added to the network. The drift velocity, therefore, is the derivative of $z^\phi_{e,k}(p)$ with respect to p. Let $z^\phi_e(p^0)$ be the initial loading level of e without installing additional GET k devices other than those already existing on the network (which represent the penetration rate p^0). The value of $z^\phi_e(p^0)$ can be obtained from the power flow solution of the network. For a given technology k, let $g^\phi_{e,k}(p)$ denote the change of loading level on phase ϕ with respect to p. We can express $z^\phi_{e,k}(p)$ as devices of GET k are installed in the network by the following equation.

$$z^\phi_{e,k}(p) = z^\phi_e(p^0) + g^\phi_{e,k}(p). \tag{5}$$

The term $g^\phi_{e,k}(p)$ can be expressed as:

$$g^\phi_{e,k}(p) = \mathrm{sgn}(k)\frac{n_k(p)\Pr^\phi_{e,k}(p)\mathbb{E}[S^\phi_{e,k}] + \mathbb{E}[\Delta S^\phi_{e,\mathrm{loss}}(p)]}{S^\phi_e}. \tag{6}$$

Here, $\mathrm{sgn}(k)$ is a sign function where $\mathrm{sgn}(k) = 1$ if GET k consumes active power and $\mathrm{sgn}(k) = -1$ otherwise, $n_k(p) = N_m p - n^0_k$ by (1), $\Pr^\phi_{e,k}(p)$ is the probability that GET k devices are installed to sections downstream of e on the desired phase ϕ, $\mathbb{E}[S^\phi_{e,k}] \in \mathbb{R}$ is the expected apparent load offset (in kVA) of one GET k device installed downstream of e and on phase ϕ, $\mathbb{E}[\Delta S^\phi_{e,\mathrm{loss}}(p)] \in \mathbb{R}$ is the expected change of network losses (in kVA) of all sections downstream of e on ϕ from adding GET k devices, and $S^\phi_e \in \mathbb{R}$ is the rated power (in kVA) of e on ϕ which is assumed given. For certain equipment whose S^ϕ_e is expressed in A (e.g., overhead lines or switches), their S^ϕ_e can be converted into kVA by multiplying the nominal voltage. The first term in (6)'s numerator can be interpreted as the expected apparent power contributed by GET k devices. We simplify (6) by assuming $\mathbb{E}[\Delta S^\phi_{e,\mathrm{loss}}(p)] \ll n_k(p)\Pr^\phi_{e,k}(p)\mathbb{E}[S^\phi_{e,k}]$. To see this, let $\mathbb{E}[I^\phi_{e,k}] \in \mathbb{R}$ and $\mathbb{E}[I^\phi_{l,k}] \in \mathbb{R}$ be the expected currents *contributed by* GET k flowing through e and through each section l downstream of e on ϕ, respectively, $\mathbb{E}[V^\phi_{e,k}] \in \mathbb{R}$ be the expected voltage of e on ϕ, and $\mathbb{E}[\delta V^\phi_{l,k}] \in \mathbb{R}$ be the expected voltage drop on section l. Hence, we have

$$n_k(p)\Pr^\phi_{e,k}(p)\mathbb{E}[S^\phi_{e,k}] = \mathbb{E}[V^\phi_{e,k}]\mathbb{E}[I^\phi_{e,k}] \tag{7}$$

$$\begin{aligned}\mathbb{E}[\Delta S^\phi_{e,\mathrm{loss}}(p)] &= \sum\nolimits_l \mathbb{E}[I^\phi_{l,k}]\mathbb{E}[\delta V^\phi_{l,k}] \\ &\le \sum\nolimits_l \mathbb{E}[I^\phi_{l,k}]\overline{\delta V} \\ &= \overline{\delta V}\sum\nolimits_l \mathbb{E}[I^\phi_{l,k}]\end{aligned} \tag{8}$$

where $\overline{\delta V}$ is an upper bound of the voltage drops for all sections. Given that in general $\overline{\delta V} \ll \mathbb{E}[V^\phi_{e,k}]$ and $\sum_l \mathbb{E}[I^\phi_{l,k}] \approx \mathbb{E}[I^\phi_{e,k}]$ if phase angles of currents are similar, we can establish $\mathbb{E}[\Delta S^\phi_{e,\mathrm{loss}}(p)] \ll n_k(p)\Pr^\phi_{e,k}(p)\mathbb{E}[S^\phi_{e,k}]$ by (7) and (8). Hence $g^\phi_{e,k}$ can be approximated by,

$$g^\phi_{e,k}(p) \approx \mathrm{sgn}(k)\frac{n_k(p)\Pr^\phi_{e,k}(p)\mathbb{E}[S^\phi_{e,k}]}{S^\phi_e}. \tag{9}$$

By (5), we have $\mathrm{d}/\mathrm{d}p\left(z^\phi_{e,k}(p)\right) = \mathrm{d}/\mathrm{d}p\left(g^\phi_{e,k}(p)\right)$. To obtain an analytical form of $\mathrm{d}/\mathrm{d}p\left(g^\phi_{e,k}(p)\right)$, $\Pr^\phi_{e,k}(p)$ and $\mathbb{E}[S^\phi_{e,k}]$ need to be computed.

1) Computation of $\Pr^{\phi}_{e,k}(p)$*:* To derive an analytical form for $\Pr^{\phi}_{e,k}(p)$, we first denote $\mathcal{K}$ as the set of all customers, and let $\mathcal{K}^{\phi}_{e} = \{i \in \mathcal{K} \mid i \text{ is downstream of } e \text{ and on phase } \phi\}$ be the set of customers who are downstream of equipment e on phase ϕ. Suppose for now that n^i_k, the number of GET k devices allocated to customer i at any p, is proportional to their expected number of GET devices, which is $\mathbb{E}[n^i_k] = \sum_{n=0,1,\dots} n \Pr^i_{\text{num},k}(n)$. We have,

$$n^i_k(p) = \begin{cases} n_k(p)\mathbb{E}[n^i_k] / \sum_{i\in\mathcal{K}} \mathbb{E}[n^i_k], & \text{if } p < \overline{p}^i_k, \\ \overline{n}^i_k, & \text{if } p \geq \overline{p}^i_k. \end{cases} \tag{10}$$

Here, $\overline{n}^i_k$ is the maximum number of devices that i is allowed to add, i.e., $\overline{n}^i_k = \text{argmax}_n\{\Pr^i_{\text{num},k}(n) > 0\}$, and $\overline{p}^i_k$ is the penetration rate at which $\overline{n}^i_k$ devices would have been added to i. Given that such $\overline{p}^i_k$ may differ from one customer to another, for a given p, some customer, say i, may already have hit the upper bound ($p > \overline{p}^i_k$) while another customer j still has room to add more devices ($p \leq \overline{p}^j_k$). In such a case, the total number of devices actually added to the network (denoted as $\tilde{n}_k$) will be less than the desired n_k.

Let $\Delta n_k(p) = n_k(p) - \tilde{n}_k(p)$ be the gap between the desired n_k and the actual $\tilde{n}_k$ at p, and $\mathcal{K}_{>p,k} = \{i \in \mathcal{K} | p < \overline{p}^i_k\}$ be the set of customers who still have room to add more devices. Then for each $i \in \mathcal{K}_{>p,k}$, we can compute the available room to add more GET k devices, which is $\Delta^i_k = \overline{n}^i_k - n^i_k > 0$. To bridge the gap of Δn_k, each customer in $\mathcal{K}_{>p,k}$ should receive some additional number of devices to n^i_k in (10). Such an additional number $n^{i,+}_k$ for $i \in \mathcal{K}_{>p,k}$ is computed by,

$$n^{i,+}_k(p) = \Delta n_k(p) \frac{\Delta^i_k}{\sum_{i\in\mathcal{K}_{>p,k}} \Delta^i_k}.$$

The computed $n^{i,+}_k$ for a customer i never exceeds their available room Δ^i_k because $\Delta n_k \leq \sum_{i\in\mathcal{K}_{>p,k}} \Delta^i_k$. Note that for $i \notin \mathcal{K}_{>p,k}$, $n^{i,+}_k = 0$.

Hence, the number of devices added to customer i can be modified from (10) by the following equation,

$$n^i_k(p) = \begin{cases} n^{i,\text{old}}_k(p) + n^{i,+}_k(p), & \text{if } p < \overline{p}^i_k, \\ \overline{n}^i_k, & \text{otherwise}, \end{cases} \tag{11}$$

where $n^{i,\text{old}}_k$(p) is computed by (10).

The probability $\Pr^{\phi}_{e,k}(p)$ can then be expressed as the ratio of the total numbers of devices that can be installed to all customers in $\mathcal{K}^{\phi}_e$ over those in $\mathcal{K}$:

$$\Pr^{\phi}_{e,k}(p) = \frac{\sum_{i\in\mathcal{K}^{\phi}_e} n^i_k(p)}{\sum_{i\in\mathcal{K}} n^i_k(p)}. \tag{12}$$

2) Calculation of $\mathbb{E}[S^{\phi}_{e,k}]$*:* Recall that $\mathbb{E}[S^{\phi}_{e,k}]$ is the expected load offset of one GET k device for all customers in $\mathcal{K}^{\phi}_e$. Let $s^i_k(t) \in \mathbb{C}$ be the complex load offset value at time t for 1 GET k device added to i. Based on $\mathcal{L}_k$ and $\Pr^i_{\mathcal{L},k}(j)$, one can compute the expected value of $s^i_k(t)$.

$$\mathbb{E}[s^i_k(t)] = \sum_j l^j_k(t) \Pr^i_{\mathcal{L},k}(j). \tag{13}$$

To calculate $\mathbb{E}[S^{\phi}_{e,k}]$, one can sum up $\mathbb{E}[s^i_k(t)]$ for all customers in $\mathcal{K}^{\phi}_e$ and then divided by the number of customers in $\mathcal{K}^{\phi}_e$. Specifically,

$$\mathbb{E}[S^{\phi}_{e,k}] = \frac{1}{|\mathcal{K}^{\phi}_e|} \left| \sum_{i\in\mathcal{K}^{\phi}_e} \mathbb{E}[s^i_k(t)] \right|, \tag{14}$$

where $|\mathcal{K}^{\phi}_e|$ is the cardinality of $\mathcal{K}^{\phi}_e$. Note that in (14), the magnitude is taken to the summation since $\mathbb{E}[s^i_k(t)]$ may be a complex value.

D. Drift velocity of voltage level

As GET devices are added to the distribution network, voltage levels will be affected. For example, undervoltage conditions are likely to happen to areas with a high penetration rate of EVs, whereas overvoltage conditions may occur where high penetration levels of distributed energy resources (e.g. PVs) are installed on the network. In this section, the computation of the drift velocity $u(x,p)$ is outlined, where $m(x,p)$ is the PDF of voltage level x at various penetration rate p. Unless specified otherwise, all complex variables in this section refer to per-unit values.

We first consider a simple case where *one* GET k device is installed to a customer. We can then compute the current injected by such GET k device, Δi, by the following equation, from which the change of voltage of this customer can be approximated,

$$\Delta i = \frac{\Delta v}{z_{\text{th}}} \approx \left(\frac{\Delta S}{v}\right)^*, \tag{15}$$

where $v \in \mathbb{C}$ is the customer's voltage before installing the GET device, $z_{\text{th}} \in \mathbb{C}$ is the equivalent network impedance at the customer's location, $\Delta S \in \mathbb{C}$ is the complex power of the GET device, and $\Delta v \in \mathbb{C}$ is the change of voltage after installing the GET device. The $*$ denotes the complex conjugate.

It should be noted that Δv in (15) is an approximated value for the following two reasons:

1) The equivalent impedance z_{th} is assumed to be computed from a linearized network. In reality, the distribution network is not fully linear (due to load/generation models, tap changers and voltage-controlled devices, etc.).
2) Even if the network is fully linear, the current injected by the GET device is non-linear since it depends on the voltage $(v + \Delta v)$ after the GET device is connected. However, on (15)'s right-hand side, the voltage v before adding GET devices is used, thus any voltage difference is neglected.

Model (15) can be extended to simultaneously installing multiple GET devices to multiple customers in a matrix form,

$$\begin{aligned} \Delta \boldsymbol{v}(p) &= \boldsymbol{Z} \Delta \boldsymbol{i}(p) \\ &= \left[\Delta v^{\phi}_{o,k}(p), o = 1, 2, \cdots, N, \phi = A, B, C\right]^{\top} \end{aligned} \tag{16}$$

where $\Delta \boldsymbol{v} \in \mathbb{C}^{3N}$ is the voltage change of all N nodes of the network, and $\Delta \boldsymbol{i} \in \mathbb{C}^{3N}$ is the vector of currents injected to each node due to the GET devices installed. It is assumed that each node is three-phase, and for a non-connected phase,

the voltage and current injection are zero. The matrix $\boldsymbol{Z} \in \mathbb{C}^{3N\times 3N}$ is the sensitivity impedance of voltage with respect to current injection, which can be computed from the modified augmented nodal analysis (MANA) formulation [33], [34].

When multiple GET devices are connected to a node o, the current injected to o on phase ϕ, which is a single element of $\Delta\boldsymbol{i}$, can be computed using:

$$\Delta i^{\phi}_{o,k}(p) \approx n_k(p)\Pr{}^{\phi}_{o,k}(p)\left(\frac{\mathbb{E}[S^{\phi}_{o,k}]}{v^{\phi}_o(p^0)}\right)^*, \tag{17}$$

where $\Pr^{\phi}_{o,k}(p)$ is the probability of a GET k device directly installed to o on phase ϕ, $\mathbb{E}[S^{\phi}_{o,k}]$ is the expected load offset of a GET k device directly connected to o on ϕ, and $v^{\phi}_o(p^0)$ is the initial phase voltage of node o on phase ϕ which can be obtained from a power flow analysis.

To derive the analytical forms of $\Pr^{\phi}_{o,k}(p)$ and $\mathbb{E}[S^{\phi}_{o,k}]$, let $\mathcal{K}^{\phi}_o = \{i \in \mathcal{K} | i \text{ is connected to } o \text{ and on phase } \phi\}$ be the set of customers who are directly connected to node o and on phase ϕ. Then $\Pr^{\phi}_{o,k}(p)$ can be computed in a similar way as in (12), such that,

$$\Pr{}^{\phi}_{o,k}(p) = \frac{\sum_{i\in\mathcal{K}^{\phi}_o} n^i_k(p)}{\sum_{i\in\mathcal{K}} n^i_k(p)},$$

where $n^i_k(p)$ is given by (11). Recall that $\mathcal{K}$ is the set of all customers on the network. The expected load offset $\mathbb{E}[S^{\phi}_{o,k}]$ can be obtained from (14) where $\mathcal{K}^{\phi}_e$ is replaced by $\mathcal{K}^{\phi}_o$.

Having computed the current injection to each node on each phase from (17), we can construct the column vector $\Delta\boldsymbol{i}(p)$ by,

$$\Delta\boldsymbol{i}(p) = \left[\Delta i^{\phi}_{o,k}(p), o = 1, 2, \cdots, N, \phi = A, B, C\right]^{\top}.$$

Using the resulting $\Delta\boldsymbol{i}(p)$ in (16), we can obtain the column vector $\Delta\boldsymbol{v}(p)$ which has the voltage changes to all nodes on all phases. The voltage of node o on phase ϕ can be written as,

$$v^{\phi}_{o,k}(p) \triangleq v^{\phi}_o(p^0) + \Delta v^{\phi}_{o,k}(p).$$

As x refers to the voltage level and $v^{\phi}_{o,k}(p)$ is the per-unit voltage in phasor, we let $x = |v^{\phi}_{o,k}(p)|$, and the drift velocity of x can be analytically computed by taking the derivative of $|v^{\phi}_{o,k}(p)|$ with respect to p. Due to space limitation, the resulting long expression for the derivative is omitted, but it is straightforward to obtain it using the chain rule and algebraic operations with all the components of $v^{\phi}_{o,k}(p)$ expressed in this section.

III. Numerical solution

A. Numerical solution to the FPE

Let the solution to the FPE be $M = \{m(x,p)\}_{p\in\mathcal{P}}$ which is a sequence of PDFs for each p in a discretized set of penetration rates $\mathcal{P} = \{p^0, p^0 + \Delta p, p_0 + 2\Delta p, \cdots, p^{\max}\}$ with a step size Δp. The conservation law is to be satisfied such that the cumulative probability of each PDF must sum up to 1. For advection-diffusion partial differential equations such as the Fokker-Planck equation under the conservation law, finite-volume method (FVM) [35], [36] is a suitable class of numerical methods to solve them. Using an implicit scheme, we can write the following generalized equation:

$$m(x,p) = f\left(m(x,p+\Delta p), u(x,p+\Delta p)\right), \tag{18}$$

where the mapping $f : \mathbb{R}^{n_x} \mapsto \mathbb{R}^{n_x}$ depends on the selected discretization scheme and n_x is the number of discretized points of x. For some schemes such as backwards Euler or Crank–Nicolson, the mapping f is linear such that it can be represented by a matrix $\boldsymbol{S}_u \in \mathbb{R}^{n_x\times n_x}$ which depends on $u(x, p+\Delta p)$. Hence, (18) can be written as:

$$m(x,p) = \boldsymbol{S}_u m(x, p+\Delta p).$$

By taking the inverse of $\boldsymbol{S}_u$, we can compute $m(x, p+\Delta p)$ given $m(x,p)$ by the following equation:

$$m(x, p+\Delta p) = \boldsymbol{S}^{-1}_u m(x,p).$$

Due to space limitation, the form of the $\boldsymbol{S}_u$ matrix is not expressed in this paper, but the readers can refer to [35], [36] for more details. It is remarked that $\boldsymbol{S}_u$ is tridiagonal, where its diagonal elements are always non-zero, and lower/upper-diagonal elements on each row are also non-zero and differ from any diagonal element. Hence, for any given row of $\boldsymbol{S}_u$, it must be linearly independent of any other row. Therefore, $\boldsymbol{S}_u$ has always full rank and is non-singular.

B. Calibration of $m(x,p)$

Due to the approximations made in (9) and (17), the drift velocity term $u(x,p)$ computed in Section II-C for equipment loading levels and in Section II-D for voltage levels may become less accurate when (i) the penetration rate is high enough such that magnitudes of voltage drop/rise caused by GET devices are non-negligible, and (ii) transformers with load-tap changers, voltage regulators and Volt/VAR devices are active on the network causing non-linearity to voltages and network loadings. Under these situations, the obtained sequence of $m(x,p)$ by numerically solving the FPE may not accurately reflect the evolution of network states as a function of the penetration rate. To improve the accuracy, a calibration process is proposed.

Based on (11) and (13), we can compute the expected number of GET k devices $n^i_k(p)$ that should be added to customer i at penetration p and the expected offset of one GET k device $\mathbb{E}[s^i_k(t)]$ at the time of analysis t. We can then compute an aggregated offset value of $n^i_k(p)\mathbb{E}[s^i_k(t)]$ which is to be added to each customer i in the network model. A power flow analysis will be performed to the modified network to extract the "true" mean network state $\tilde{x}(p)$. In the ideal case, such $\tilde{x}(p)$ should be identical to the mean value $\mathbb{E}[x(p)]$ which is computed from $\mathbb{E}[x(p)] = \int_{-\infty}^{+\infty} xm(x,p)\mathrm{d}x$. If $\mathbb{E}[x(p)] \neq \tilde{x}(p)$, then $m(x,p)$ should by calibrated to $m(x+\Delta x, p)$ where $\Delta x(p) \triangleq \tilde{x}(p) - \mathbb{E}[x(p)]$ is the shift value at p.

To maintain the computation efficiency, the power flow analysis of the network with aggregated offset values added to customers is not done for all $p \in \mathcal{P}$. Rather, the power flow is only performed at a few selected penetration rates,

and the calibration is interpolated for all penetration rates. The process is as follow. Suppose that $\mathcal{P}_{\text{pf}}$ is the set of selected penetration rates for power flow analyses, then we construct $\mathcal{D} = \{\Delta x(p) = \tilde{x}(p) - \mathbb{E}[x(p)]\}_{p\in\mathcal{P}_{\text{pf}}}$ which are differences between the mean $\mathbb{E}[x(p)]$ from the REM and the power flow mean values $\tilde{x}(p)$ at these penetration rates. Then for any $p \in \mathcal{P}$, we compute $\Delta x(p)$ by,

$$\Delta x(p) = \frac{p - p_-}{p_+ - p_-}\left[\Delta x(p_+) - \Delta x(p_-)\right] + \Delta x(p_-),$$

where $p_- \triangleq \max\{\mathcal{P}_{\text{pf}} \cap [p^0, p]\}$ is the greatest element in $\mathcal{P}_{\text{pf}}$ that is less than or equal to p, $p_+ \triangleq \min\{\mathcal{P}_{\text{pf}} \cap [p, p^{\max}]\}$ is the least element in $\mathcal{P}_{\text{pf}}$ that is greater than or equal to p, and $\Delta x(p_-), \Delta x(p_+) \in \mathcal{D}$. The calibrated PDF $m(\tilde{x}, p) = m(x + \Delta x(p), p)$ is reported as the REM results. The number of power flow analyses required by the calibration process is independent of the number of network sections for analysis but only corresponds to the number of penetration rates included in $\mathcal{P}_{\text{pf}}$, which is a much smaller set than $\mathcal{P}$. Thus, the calibration process adds negligible computation efforts to the REM.

It is remarked that this calibration step is optional. It does not produce any PDF, but rather it "shifts" the mean values of the PDFs computed numerically from the FPE. For large-scale distribution networks such as the IEEE-8500 test feeder used in the next section, voltage control (e.g., regulators) and VAR control (e.g., switchable shunt capacitors) devices are usually installed. Without the calibration step, the PDFs obtained by solving the FPE while neglecting impacts of these Volt/VAR devices may result in observable errors even at low penetration rates, and become much less accurate as the penetration rate increases. On the other hand, the calibration step can be skipped without impacts on the accuracy for small networks or networks without Volt/VAR control devices.

C. Combined impact of multiple GETs

In the sections above, we present the methodology to study the impact of one GET type (EV or PV) installed to the network at various penetration rates. Throughout this section, the subscript k in the notation is conserved for clarity. Recall that $M_k = \{m(x, p_k)\}_{p_k\in\mathcal{P}_k}$ is the sequence of PDFs computed and calibrated from the REM for each $p_k \in \mathcal{P}_k$ when only GET type k exists on the network. In the following, we explain how the model can be extended to the analysis of networks where multiple GET types exist.

We start with the case in which two GET types k_1 and k_2 are installed to the network. We first compute $M_{k_1|k_2} = \{m(x, p_{k_1})|p_{k_2}\}_{p_{k_1}\in\mathcal{P}_{k_1}}$ when the penetration rate of GET type k_2 is fixed at p_{k_2} as a parameter. To do so, the numerical FVM in Section III-A is first used to solve for $m(x, p_{k_1})|p_{k_2}$ from the FPE under the drift velocity of $u(x, p_{k_1})|p_{k_2}$. By assuming the impacts of GET types k_1 and k_2 are independent, we have $u(x, p_{k_1})|p_{k_2} = u(x, p_{k_1}) + u(x, p_{k_2})$ by superposition. If necessary, the method in Section III-B is then applied to calibrate the resulting $m(x, p_{k_1})|p_{k_2}$. We repeat this process to construct $M_{k_1,k_2} = \{M_{k_1|k_2}\}_{p_{k_2}\in\mathcal{P}_{k_2}}$ which is a sequence of $M_{k_1|k_2}$ for each $p_{k_2} \in \mathcal{P}_{k_2}$. Each element of M_{k_1,k_2} represents the network state PDF under the combined impacts of GET types k_1 and k_2 with penetrations p_{k_1} and p_{k_2}, respectively.

For combined impacts of more than two GET types, the same process is adopted to compute $M_{k_1,k_2,\cdots}$, where sequences of PDFs are solved and calibrated by varying the penetration rate of one GET type at a time.

IV. Test results

In this section, results from the stochastic impact analysis are illustrated on a test feeder.

A. Test setup

The IEEE-8500 test feeder [37] is selected to demonstrate REM's ability to conduct a rapid stochastic impact analysis. The following modifications to the network are made:

- The network contains 1177 spot loads and a total of 4205 customers, where some spot loads contain more than 1 customer (the original IEEE-8500 network assumes each spot load models 1 customer).
- The substation transformer is changed to Wye-wye configuration such that its loading on each phase can be more easily seen.

To assess impacts of EVs and PVs on the network, we assume that the following information are given.

- 4 levels of EV charging power are considered (1.8kW, 3.6kW, 6.6kW, 7.2kW). Charging may start at any hour during the day and can last 2, 4, or 8 hours. Hence the set $\mathcal{L}_{\text{EV}}$ contains 288 charging profiles, where each profile $l^j_{\text{EV}}(t)$ is time-series data over a 24-hour period.
- 4 levels of PV nominal generation capacity are considered (1.6kW, 2.4kW, 3.6kW, 4kW). We consider 8 solar radiance profiles under different weather types [38], hence we have a total of 32 PV generation profiles in $\mathcal{L}_{\text{PV}}$.
- Each customer is assumed to have at most 1 EV.[1] Quantity probability $\Pr^i_{\text{num,EV}}(1)$ for customer $i = 1, 2, \ldots, 4205$ is randomly generated from the uniform distribution $\mathcal{U}[0, 1]$, and $\Pr^i_{\text{num,EV}}(0) = 1 - \Pr^i_{\text{num,EV}}(1)$. The probability $\Pr^i_{\mathcal{L},\text{EV}}(j)$ for adopting charging profile $j \in [1, 288]$ is also randomly generated from $\mathcal{U}[0, 1]$, and $\Pr^i_{\mathcal{L},\text{EV}}$ is normalized such that $\sum_j \Pr^i_{\mathcal{L},\text{EV}}(j) = 1$.
- Each customer is assumed to have at most 1 PV.[1] Quantity probability $\Pr^i_{\text{num,PV}}(1)$ for customer $i = 1, 2, \ldots, 4205$ is randomly generated from $\mathcal{U}[0, 1]$, and $\Pr^i_{\text{num,PV}}(0) = 1 - \Pr^i_{\text{num,PV}}(1)$. To consider the correlation between the weather and the generation profiles, we first randomly generate the probabilities of the 8 solar radiance profiles from $\mathcal{U}[0, 1]$, denoted as $\Pr_{\text{solar}}$. Then for each customer i we generate the probabilities of PV generation capacity from $\mathcal{U}[0, 1]$, denoted as $\Pr^i_{\text{PVgen}}$. Then, we let $\Pr^i_{\mathcal{L},\text{PV}} = \Pr_{\text{solar}} \times \Pr^i_{\text{PVgen}}$. Finally, $\Pr^i_{\mathcal{L},\text{PV}}$ is normalized such that $\sum_j \Pr^i_{\mathcal{L},\text{PV}}(j) = 1$.

The information assumed above can be obtained from socio-economic or statistical studies. Taking EV as an example, in [26], [28] the adoption rate of electric vehicles in terms of different demographic and socio-economic characteristics

[1] The REM approach is not restricted to this assumption. Rather, this assumption is made to reduce the amount of the probability data required and, hence, simplifies the simulation.

(such as education levels, age group, household income, etc.) are studied. The results can be used to quantify the customer's quantity probabilities. Probability distributions of EV charging start time, state-of-charge (SOC), and EV travel information (mileage and duration per trip) are collected in [11], [12], and can be used to generate EV charging profiles along with the adoption probability. Although there may exist much more possible profiles from the statistical studies, in this paper we use a smaller set of representative EV charging and PV generation profiles (288 and 32, respectively) for illustration purpose. As increasing the number of profiles only affects the calculations of the expected value in (13) and (14), it has negligible impacts on the efficiency of the REM. Thus, the number of profiles is not limited in the REM approach, nor does it affect its efficiency.

The modified network is modeled in the CYME software which is also used to perform power flow analysis. The numerical solution and the calibration process of the REM analysis are implemented in `Python`.

B. Result validation

For validation of accuracy, a Monte Carlo simulation-based approach is also developed, where samples are constructed by randomly sampling from the same data set as given above. Although such a Monte Carlo simulation-based approach is slow in convergence rate and requires extensive computation for power flow analysis, the obtained results can be considered accurate once convergence is reached. Stochastic analysis results based on power flow solutions, i.e., the empirical distributions of loading levels of equipment and voltage levels of network sections, serve as benchmarks for the REM and is used to evaluate its performance.

C. Test results – EV

The analysis time is set to 8:00PM when peak load usually occurs, and 2000 samples are constructed for Monte Carlo simulations.

1) Loading level at selected penetration levels: Figure 1 illustrates the loading levels of the substation transformer when EVs are installed to the modified IEEE-8500 network at 3 penetration rates (10%, 30%, 50%). Because the power flow solution to a Monte Carlo sample is likely to diverge if the penetration rate is too high, completing the Monte Carlo simulations of 2000 samples with feasible power flow solutions takes too long. For this reason, the maximum penetration rate of 50% is chosen for Monte Carlo simulations. However, no limitation to the maximum penetration rate exists for our REM approach (i.e., see Section IV-C3 and Figure 3 below). Results from the REM with and without calibrations at these penetration rates are compared with empirical distributions from Monte Carlo simulations. It should be noted that the y-axis represents the probability density and the total area under each curve sums up to 1.

2) Voltage level at selected penetration levels: Figure 2 illustrates the probability distributions of the voltage level (in percentage) of a section called "M1125994". This section is 4.8 km from the substation and 1.2 km from a downstream

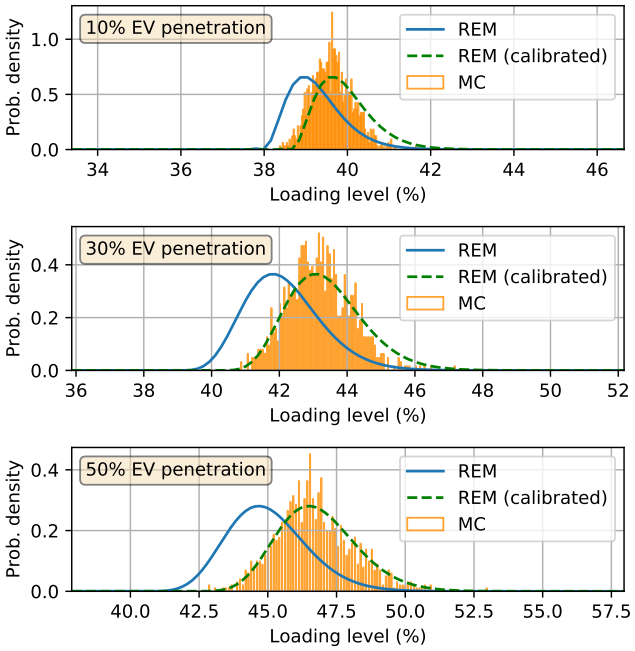

Fig. 1. Comparison of probabilities of the substation transformer loading levels at various EV penetration rates, REM results (blue) vs. calibrated REM results (green dashed) vs. Monte Carlo (orange)

voltage regulator. Due to its long distance from the substation, this section is expected to suffer from undervoltage when EV penetration is high. While, the substation voltage is regulated at 1.05 p.u., Figure 2 confirms that this section has a small probability of having undervoltage issues at 50% EV penetration if the lower voltage limit is set at 0.95 p.u.

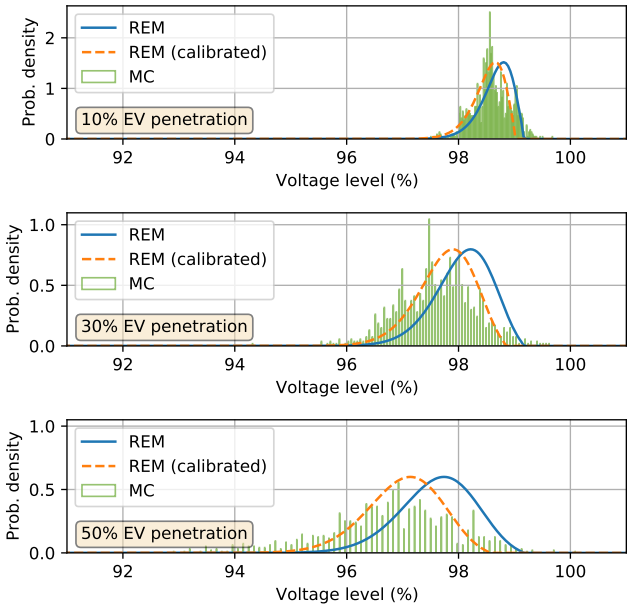

Fig. 2. Comparison of probabilities of voltage levels the M1125994 section at various EV penetration rates, REM results (blue) vs. calibrated REM results (orange dashed) vs. Monte Carlo (green)

3) Network state estimation across a wide range of penetration rates: The REM provides a rapid estimation of

network states (minimum, average, maximum levels) across a wide range of penetration rates, based on the computed PDFs. It also enables fast assessment for probabilities of abnormal conditions on the network (overloading, under/over-voltages) as the penetration rate increases. For example, Figure 3 shows the average, minimum, and maximum loading levels as well as overloading probabilities of a main line section near the substation of the IEEE-8500 network up to 100% EV penetration rate. Here, we take the mean value of the calibrated PDF at each p as the average loading level, and minimum and maximum loading levels are $\pm$ 2$\times$ standard deviation from the mean value, respectively. This type of results can provide insightful information on what penetration level can be supported by the network before severe abnormalities occur.

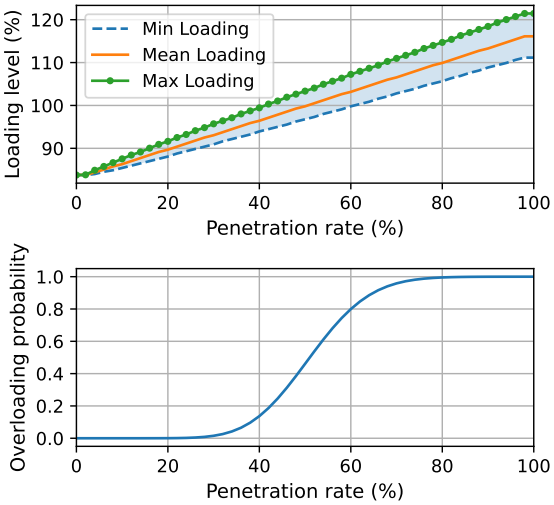

Fig. 3. Loading levels (top) and overloading probabilities (bottom) of a line section vs. penetration rates of EVs added to the IEEE-8500 network

4) Network locations susceptible to abnormal conditions: In the last experiment related to EVs, we select a total of 183 network sections on which the initial equipment loading level is greater than 50% or the initial voltage level is less than 0.965 p.u. before any EV is installed to the network. These sections are likely to have equipment overloading or undervoltage issues as EVs are installed to the network. The impact analysis is performed to these sections to estimate network states and probabilities of abnormal conditions at different EV penetration rates. Figure 4 visualizes the locations of the network sections with the calculated probabilities of abnormal conditions at 50% EV penetration rate. Here, we create *Overloading probability* and *Undervoltage probability* keywords in CYME to highlight the locations according to the computed overloading and undervoltage probabilities, respectively. Almost all selected network equipment have some probability of overloading. The line sections near the substation have an overloading probability close to 50%, which is consistent with the results of Figure 3. The undervoltage issue is less severe on the network mainly due to the presence of voltage regulators (indicated by the green droplets). Only some sections that are far from the substation and upstream of a voltage regulator show some probability of undervoltage, and this is where the section M1125994 is located.

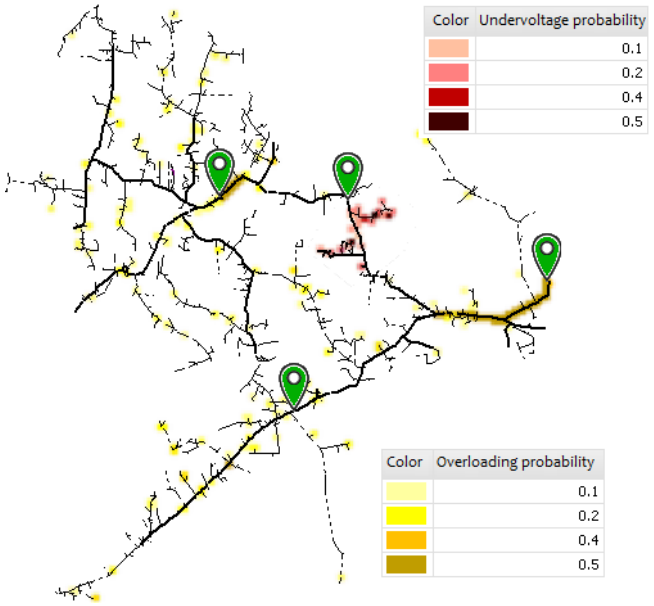

Fig. 4. IEEE-8500 network sections susceptible to equipment overloading and undervoltage issues at 50% EV penetration rate

D. Test results – PV

The analysis time is set to 11:00AM when the solar irradiance starts to reach the peak. While customers can have diversified EV charging profiles, the PV generation profiles are correlated with the weather type, thus the analysis should be performed conditional on each weather type. Then weighted average results based on probabilities of all weather types are computed. Here, we present only the results when considering PV generation profiles during sunny days to avoid running Monte Carlo simulations for other weather types. In Monte Carlo simulations, 2000 samples are constructed.

1) Loading level at selected penetration levels: Figure 5 illustrates the loading levels of the substation transformer when PV are installed to the network at 3 penetration rates (10%, 30%, 50%). Similar to the EV case, 50% penetration is selected for comparison purpose. Results from the REM with and without calibrations are compared with empirical distributions from Monte Carlo simulations.

2) Voltage level at selected penetration levels: Figure 6 illustrates the probability distributions of the voltage level (in percentage) of the M1125994 section when PVs are installed. Although no overvoltage (by taking 1.05 p.u. as the upper voltage limit) occurs to this section due to its distance from the substation, it is observed that REM well captures the trend of increasing voltage levels with the PV penetration. Therefore, overvoltage would possibly occur as the penetration continues to increase, and calibrated results from the REM approach would accurately indicate the probabilities.

3) Network locations susceptible to overvoltage: In this experiment, we select a total of 114 locations whose voltage level is greater than 1.04 p.u. before any PV is installed to the network. When setting the upper voltage limit to 1.05 p.u., more than half of these locations have overvoltage probabilities at 50% PV penetration. The calibrated probability values from

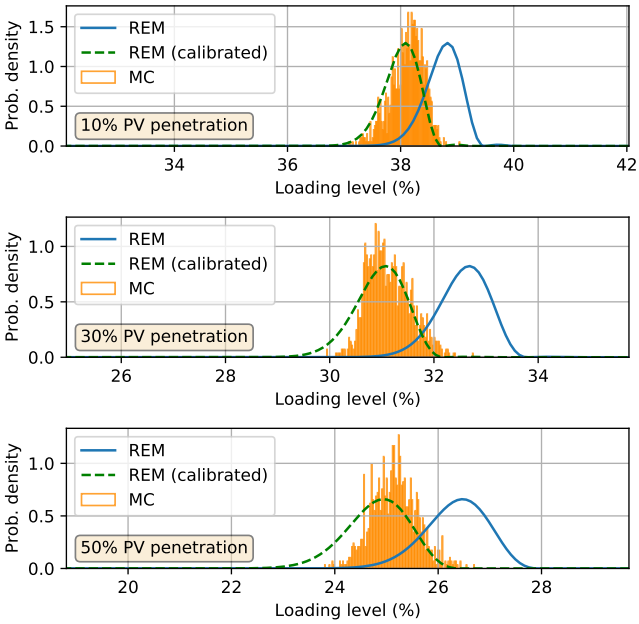

Fig. 5. Comparison of probabilities of the substation transformer loading levels at various PV penetration rates, REM results (blue) vs. calibrated REM results (green dashed) vs. Monte Carlo (orange)

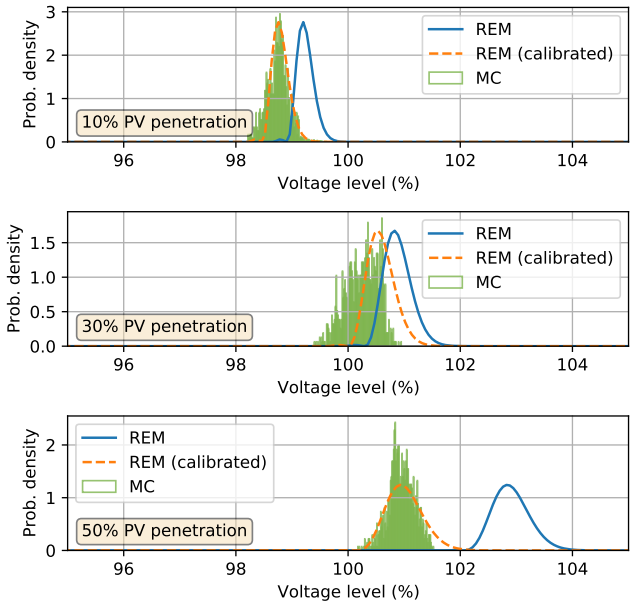

Fig. 6. Comparison of probabilities of voltage levels of the M1125994 section at various PV penetration rates, REM results (blue) vs. calibrated REM results (orange dashed) vs. Monte Carlo (green)

REM are used to create the *Overvoltage probability* keyword in CYME to highlight network sections in Figure 7. It is observed that sections near the substation have close to 100% probability of overvoltage, while some sections immediately downstream of voltage regulators also have overvoltage issues with different probabilities.

E. Discussion of results

1) Impacts of EV: As the EV penetration level increases, more power demand is expected on the network due to EV

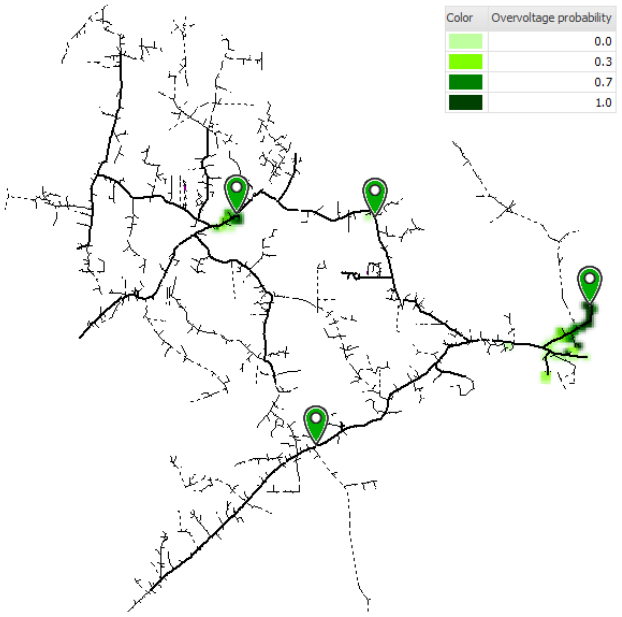

Fig. 7. IEEE-8500 network sections susceptible to overvoltage issues at 50% PV penetration rate

charging. This will increase the loading level of the substation transformer, and more voltage drops will be incurred on each network section. These can be confirmed from observations of Figure 1 and 2: empirical distributions of the substation transformer loading levels constructed from Monte Carlo simulations shift towards the heavier loading side, and distributions of the voltage level shift towards the lower voltage side. The same dynamics are also observable from the PDFs computed and calibrated from the REM approach. In addition, as the level of uncertainties increases with the penetration rate, so does the *dispersion* of the distributions which is well captured by both Monte Carlo and our method.

Tables I and II provide quantitative comparisons of the mean and standard deviation values for the substation transformer loading levels using the two methods at various EV penerations. Here, the standard deviation is the square root of the variance. For the Monte Carlo method, the variance is empirically computed from the result set of all Monte Carlo samples constructed. For the REM, let $m(x,p)$ and $m(\tilde{x},p)$ be the non-calibrated and calibrated PDFs, respectively. Hence, we have $\sigma^2 = \int_{-\infty}^{\infty}(x-\mathbb{E}[x])^2 m(x,p)\mathrm{d}x$ and $\tilde{\sigma}^2 = \int_{-\infty}^{\infty}(\tilde{x}-\mathbb{E}[\tilde{x}])^2 m(\tilde{x},p)\mathrm{d}\tilde{x}$ for the variance of $m(x,p)$ and $m(\tilde{x},p)$, respectively. From the calibration process, we have $\tilde{x} = x + \Delta x$, hence $\sigma^2 = \tilde{\sigma}^2$ and the standard deviation values with and without calibration are identical. For this reason, only one cell is used for the standard deviation (Std) of REM PDFs at each penetration rate in Tables II, IV, VI, and VIII. From Table I, the calibration step significantly reduces the errors of the mean values, i.e., from 4.32% to less than 0.5% at 50% penetration rate. From Table II, the Std errors are negligible, and the increasing values reflect higher levels of uncertainties with the penetrations.

Tables III and IV list the mean and standard deviation values of voltage levels of the section M1125994 at various EV

penetrations. The mean values from both methods are almost identical without the calibration step, and the calibration step further reduces the errors. The standard deviation values are also observed to increase with the penetration rates, which captures the higher levels of uncertainties.

TABLE I
MEAN VALUES OF SUBSTATION TRANSFORMER LOADING LEVELS AT VARIOUS EV PENETRATION RATES

Substation transformer loading levels (Mean)					
Penetration	Monte Carlo (p.u.)	Non-calibrated (p.u.) and *relative error*		Calibrated (p.u.) and *relative error*	
10%	0.396	0.389	*-1.97%*	0.397	*0.08%*
30%	0.432	0.419	*-3.06%*	0.432	*-0.08%*
50%	0.467	0.447	*-4.32%*	0.465	*-0.45%*

TABLE II
STANDARD DEVIATIONS OF SUBSTATION TRANSFORMER LOADING LEVELS AT VARIOUS EV PENETRATION RATES

Substation transformer loading levels (Std)			
Penetration	Monte Carlo (p.u.)	Calibrated/non-calibrated (p.u.) and *error* (p.u.)	
10%	1.31×10^{-2}	1.06×10^{-2}	*-2.5*$\times 10^{-3}$
30%	1.80×10^{-2}	1.42×10^{-2}	*-3.8*$\times 10^{-3}$
50%	2.13×10^{-2}	2.13×10^{-2}	*2.5*$\times 10^{-5}$

TABLE III
MEAN VALUES OF VOLTAGE LEVELS ON THE M1125994 SECTION AT VARIOUS EV PENETRATION RATES

Voltage levels (Mean)					
Penetration	Monte Carlo (p.u.)	Non-calibrated (p.u.) and *relative error*		Calibrated (p.u.) and *relative error*	
10%	0.9860	0.9881	*0.21%*	0.9866	*0.06%*
30%	0.9764	0.9821	*0.59%*	0.9790	*0.26%*
50%	0.9679	0.9774	*0.98%*	0.9714	*0.35%*

TABLE IV
STANDARD DEVIATIONS OF VOLTAGE LEVELS ON THE M1125994 SECTION AT VARIOUS EV PENETRATION RATES

Voltage levels (Std)			
Penetration	Monte Carlo (p.u.)	Calibrated/non-calibrated (p.u.) and *error* (p.u.)	
10%	3.26×10^{-3}	2.76×10^{-3}	*-4.96*$\times 10^{-4}$
30%	6.76×10^{-3}	5.92×10^{-3}	*-8.32*$\times 10^{-4}$
50%	1.18×10^{-2}	8.69×10^{-3}	*-3.11*$\times 10^{-3}$

2) Impacts of PV: It is generally expected that PVs have opposite impacts on the network to those by EVs, such that network equipment's loadings should decrease, and voltage rise usually happens on network sections as PV penetration increases. From Figures 5 and 6, it is observed that empirical distributions of the substation transformer loading levels shift towards the lighter loading side, while distributions of the voltage levels of the M1125994 section shift towards the higher voltage side. The PDFs computed and calibrated from our REM approach show similar dynamics.

Tables V and VI provide quantitative comparisons of the mean and standard deviation values using the two methods. As in the EV case, the calibration step reduces errors of the mean values, and errors of the standard deviation values are negligible.

TABLE V
MEAN VALUES OF SUBSTATION TRANSFORMER LOADING LEVELS AT VARIOUS PV PENETRATION RATES

Substation transformer loading levels (Mean)					
Penetration	Monte Carlo (p.u.)	Non-calibrated (p.u.) and *relative error*		Calibrated (p.u.) and *relative error*	
10%	0.381	0.388	*1.85%*	0.381	*-0.10%*
30%	0.311	0.327	*4.94%*	0.306	*-1.82%*
50%	0.252	0.265	*5.13%*	0.248	*-1.35%*

TABLE VI
STANDARD DEVIATIONS OF SUBSTATION TRANSFORMER LOADING LEVELS AT VARIOUS PV PENETRATION RATES

Substation transformer loading levels (Std)			
Penetration	Monte Carlo (p.u.)	Calibrated/non-calibrated (p.u.) and *error* (p.u.)	
10%	5.82×10^{-3}	5.19×10^{-3}	*-6.33*$\times 10^{-4}$
30%	8.31×10^{-3}	8.65×10^{-3}	*3.48*$\times 10^{-4}$
50%	8.70×10^{-3}	1.04×10^{-2}	*1.68*$\times 10^{-3}$

Tables VII and VIII list the mean and standard deviation values of voltage levels on the M1125994 section at various PV penetrations. The mean values from Monte Carlo and the REM have a relative difference of 1.85% in the worst case without the calibration step, and the calibration step further reduces the relative errors to below 0.35%.

TABLE VII
MEAN VALUES OF VOLTAGE LEVELS ON THE M1125994 SECTION AT VARIOUS PV PENETRATION RATES

Voltage levels (Mean)					
Penetration	Monte Carlo (p.u.)	Non-calibrated (p.u.) and *relative error*		Calibrated (p.u.) and *relative error*	
10%	0.9881	0.9920	*0.40%*	0.9876	*-0.05%*
30%	1.0026	1.0083	*0.57%*	1.0060	*0.34%*
50%	1.0096	1.0282	*1.85%*	1.0092	*-0.03%*

TABLE VIII
STANDARD DEVIATIONS OF VOLTAGE LEVELS ON THE M1125994 SECTION AT VARIOUS PV PENETRATION RATES

Voltage levels (Std)			
Penetration	Monte Carlo (p.u.)	Calibrated/non-calibrated (p.u.) and *error* (p.u.)	
10%	1.85×10^{-3}	1.21×10^{-3}	*6.39*$\times 10^{-4}$
30%	2.87×10^{-3}	1.81×10^{-3}	*-1.06*$\times 10^{-3}$
50%	2.32×10^{-2}	2.41×10^{-3}	*9.49*$\times 10^{-5}$

3) Quantile-Quantile plots: Besides comparing the mean and standard deviation values, we use Quantile-Quantile (Q-Q) plots to compare the calibrated PDFs from the REM against the empirical distributions from Monte Carlo simulations in Figure 8. Given that the latter are considered as the benchmark, we use them on the x-axis as the *theoretical* distribution and quantiles of the calibrated PDFs are plotted on the y-axis [39]. It is observed that in most plots the points in the center region mainly lie on a 45° straight line, indicating that the mean values of the calibrated PDFs from the REM well match those from Monte Carlo simulations. Some skewness and deviations from the straight line at both ends can also be observed. This could be due to the use of a constant diffusion velocity value in the FPE, which results in errors at the tail regions of the

PDFs. Diffusion velocity's dependency on the penetration rate and on the network model can be studied in future work.

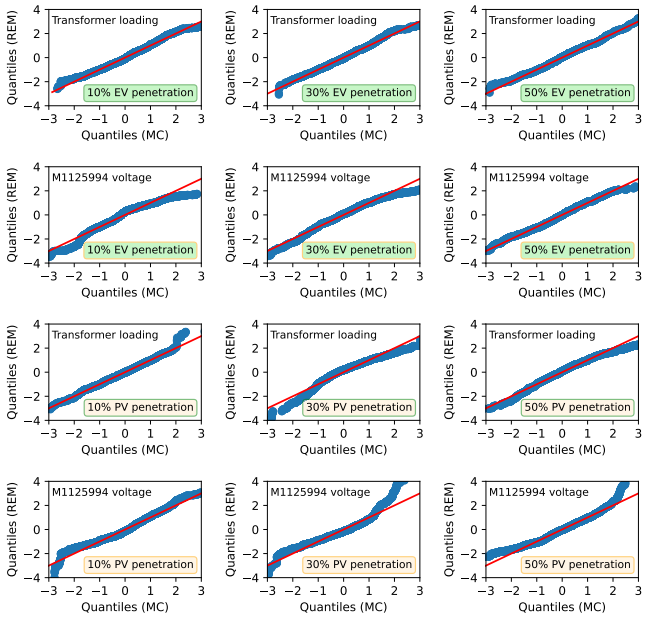

Fig. 8. Q-Q plots to compare quantiles of distributions from Monte Carlo simulations and calibrated PDFs obtained from the REM

4) Computation speed: All the experiments are performed on a workstation equipped with Intel Core i7-11800H @ 2.30GHz CPU and 16GB RAM. Table IX lists the computation time for experiments in Section IV-C when EVs are installed to the network. In Table IX, the time spent on numerically solving the FPE is reported in the "Non-calibrated" column. As the FPE is solved for *one location at a time*, the time is expected to increase linearly with the number of locations to be simulated. However, for the time in the "Calibration" column, since network states of *all locations* can be obtained from a single power flow, the time may only be slightly increased for data processing as the number of locations for simulation increases. The total time spent on each experiment using the REM is the sum of times in these two columns.

For the Monte Carlo simulation approach, since the results are based on power flow solutions of all the samples constructed, the time shown in Table IX is similar for all experiments, independent of the number of network sections included for stochastic impact analysis. For the REM approach, it is significantly faster to conduct an impact analysis to a single network section (within a few minutes including the calibration). However, in the last experiment, as 183 network sections are to be included for analysis, the time to obtain non-calibrated PDFs increases to 5.69 hours which is 190× that required for a single network section. On the other hand, as the time on calibrations does not depend linearly on the number of sections for analysis, it takes only 0.29 hour in the last experiment which is increased by only 10× on more data processing. Finally, although not shown in the table, the analysis time also depends linearly on the maximum penetration level for both approaches.

TABLE IX
COMPARISON OF COMPUTATION SPEED

Experiments	Time (hr)		
	Monte Carlo	Non-calibrated	Calibration
Substation transformer loading levels	20.84	0.03	0.03
Section M1125994 voltage levels	21.79	0.03	0.03
Line loading levels across a wide range of penetration rates	-	0.06	0.03
Identifying locations with potential abnormal conditions	21.90	5.69	0.29

V. CONCLUSION AND FUTURE WORK

In this paper, we present a rapid estimation approach to perform a stochastic analysis on the impact of EVs and PVs to power distribution networks. A calibration step is also proposed to improve the accuracy of the REM. Quantitative assessments on a large-scale realistic distribution network indicate that results from the REM well follow those from Monte Carlo simulations with minimal errors, hence Monte Carlo simulations can be avoided for such a stochastic analysis and the computation efficiency can be greatly improved.

We demonstrated the impacts of EVs and PVs to the distribution network at their respective "peak" time in this work. As a next step, time-series analysis (e.g., on a typical day) can be performed, and the combining impacts can be studied when EVs and PVs are simultaneously installed to the network. The approach can also be extended to analyze the impacts of other GETs to power networks. For example, DR programs or other control schemes for EVs and PVs can be designed to modify the set of offset profiles and to shift customers' adoption probabilities, and their effectiveness and impacts to the networks can be rapidly evaluated by the REM. As penetration levels of GETs are expected to increase over the next few years, an analysis framework dedicated to these technologies will allow utilities to properly plan and optimize their networks.

ACKNOWLEDGMENT

The authors would like to thank Francis Therrien, Lead Power System Engineer at CYME International T&D for his idea and support for applying MANA into the REM.

REFERENCES

[1] S. Habib, M. M. Khan, F. Abbas, L. Sang, M. U. Shahid, and H. Tang, "A comprehensive study of implemented international standards, technical challenges, impacts and prospects for electric vehicles," *IEEE Access*, vol. 6, pp. 13866–13890, 2018.

[2] M. Ashfaq, O. Butt, J. Selvaraj, and N. Rahim, "Assessment of electric vehicle charging infrastructure and its impact on the electric grid: A review," *International Journal of Green Energy*, vol. 18, no. 7, pp. 657–686, 2021.

[3] A. Dubey and S. Santoso, "Electric vehicle charging on residential distribution systems: Impacts and mitigations," *IEEE Access*, vol. 3, pp. 1871–1893, 2015.

[4] A. Dubey, S. Santoso, and M. P. Cloud, "Understanding the effects of electric vehicle charging on the distribution voltages," in *2013 IEEE Power Energy Society General Meeting*, pp. 1–5, 2013.

[5] A. Dubey, S. Santoso, and M. Cloud, "Comparative analysis of effects of electric vehicle loads on distribution system voltages," in *2014 IEEE PES T&D Conference and Exposition*, pp. 1–5, IEEE, 2014.

[6] J. Gomez and M. Morcos, "Impact of EV battery chargers on the power quality of distribution systems," *IEEE Transactions on Power Delivery*, vol. 18, no. 3, pp. 975–981, 2003.

[7] M. Cohen and D. Callaway, "Effects of distributed PV generation on California's distribution system, part 1: Engineering simulations," *Solar Energy*, vol. 128, pp. 126–138, 2016.

[8] M. A. Cohen, P. A. Kauzmann, and D. S. Callaway, "Effects of distributed PV generation on California's distribution system, part 2: Economic analysis," *Solar Energy*, vol. 128, pp. 139–152, 2016.

[9] A. Y. Elrayyah, M. Z. C. Wanik, and A. Bouselham, "Simplified approach to analyze voltage rise in LV systems with PV installations using equivalent power systems diagrams," *IEEE Transactions on Power Delivery*, vol. 32, no. 4, pp. 2140–2149, 2017.

[10] A. Rodriguez-Calvo, R. Cossent, and P. Frías, "Integration of PV and EVs in unbalanced residential LV networks and implications for the smart grid and advanced metering infrastructure deployment," *International Journal of Electrical Power & Energy Systems*, vol. 91, pp. 121–134, 2017.

[11] R.-C. Leou, C.-L. Su, and C.-N. Lu, "Stochastic analyses of electric vehicle charging impacts on distribution network," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1055–1063, 2014.

[12] S. Habib, M. M. Khan, F. Abbas, M. Numan, Y. Ali, H. Tang, and X. Yan, "A framework for stochastic estimation of electric vehicle charging behavior for risk assessment of distribution networks," *Frontiers in Energy*, vol. 14, no. 2, pp. 298–317, 2020.

[13] D. Salles, C. Jiang, W. Xu, W. Freitas, and H. E. Mazin, "Assessing the collective harmonic impact of modern residential loads—part i: Methodology," *IEEE Transactions on Power Delivery*, vol. 27, no. 4, pp. 1937–1946, 2012.

[14] C. Jiang, R. Torquato, D. Salles, and W. Xu, "Method to assess the power-quality impact of plug-in electric vehicles," *IEEE Transactions on Power Delivery*, vol. 29, no. 2, pp. 958–965, 2014.

[15] A. Navarro-Espinosa and L. F. Ochoa, "Probabilistic impact assessment of low carbon technologies in LV distribution systems," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2192–2203, 2016.

[16] R. Allan, A. L. Da Silva, and R. Burchett, "Evaluation methods and accuracy in probabilistic load flow solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 5, 1981.

[17] P. Zhang and S. Lee, "Probabilistic load flow computation using the method of combined cumulants and gram-charlier expansion," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 676–682, 2004.

[18] C.-L. Su, "Probabilistic load-flow computation using point estimate method," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1843–1851, 2005.

[19] Y. Xu, L. Mili, and J. Zhao, "Probabilistic power flow calculation and variance analysis based on hierarchical adaptive polynomial chaos-anova method," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3316–3325, 2019.

[20] D. Métivier, M. Vuffray, and S. Misra, "Efficient polynomial chaos expansion for uncertainty quantification in power systems," *Electric Power Systems Research*, vol. 189, p. 106791, 2020.

[21] K. Ye, J. Zhao, Y. Zhang, X. Liu, and H. Zhang, "A generalized computationally efficient copula-polynomial chaos framework for probabilistic power flow considering nonlinear correlations of pv injections," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107727, 2022.

[22] E. M. S. Duque, P. P. Vergara, P. H. Nguyen, A. Van Der Molen, and J. G. Slootweg, "Conditional multivariate elliptical copulas to model residential load profiles from smart meter data," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4280–4294, 2021.

[23] H. Sheng and X. Wang, "Probabilistic power flow calculation using non-intrusive low-rank approximation method," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3014–3025, 2019.

[24] F. Li, R. P. Malhamé, and J. Le Ny, "Mean field game based control of dispersed energy storage devices with constrained inputs," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016.

[25] K. Wang and M. L. Crow, "The Fokker-Planck Equation for power system stability probability density function evolution," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2994–3001, 2013.

[26] J. Saarenpää, M. Kolehmainen, and H. Niska, "Geodemographic analysis and estimation of early plug-in hybrid electric vehicle adoption," *Applied Energy*, vol. 107, pp. 456–464, 2013.

[27] M. A. Tamor, C. Gearhart, and C. Soto, "A statistical approach to estimating acceptance of electric vehicles and electrification of personal transportation," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 125–134, 2013.

[28] S. Carley, R. M. Krause, B. W. Lane, and J. D. Graham, "Intent to purchase a plug-in electric vehicle: A survey of early impressions in large us cites," *Transportation Research Part D: Transport and Environment*, vol. 18, pp. 39–45, 2013.

[29] S. Chen, T. Zhang, H. B. Gooi, R. D. Masiello, and W. Katzenstein, "Penetration rate and effectiveness studies of aggregated bess for frequency regulation," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 167–177, 2015.

[30] S. Akagi, R. Takahashi, A. Kaneko, M. Ito, J. Yoshinaga, Y. Hayashi, H. Asano, and H. Konda, "Upgrading voltage control method based on photovoltaic penetration rate," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 3994–4003, 2016.

[31] P. E. Kloeden and E. Platen, "Stochastic differential equations," in *Numerical Solution of Stochastic Differential Equations*, Springer, 1992.

[32] H. Risken, *The Fokker-Planck Equation Methods of Solution and Applications*. Springer Series in Synergetics, 18, Berlin, Heidelberg: Springer Berlin Heidelberg, 2nd ed. 1996. ed., 1996.

[33] J.-S. Lacroix, I. Kocar, and M. Belletête, "Accelerated computation of multiphase short circuit summary for unbalanced distribution systems using the concept of selected inversion," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1515–1522, 2013.

[34] I. Kocar, J. Mahseredjian, U. Karaagac, G. Soykan, and O. Saad, "Multiphase load-flow solution for large-scale distribution systems using MANA," *IEEE Transactions on Power Delivery*, vol. 29, no. 2, pp. 908–915, 2014.

[35] R. J. LeVeque and R. J. Leveque, *Numerical methods for conservation laws*, vol. 214. Springer, 1992.

[36] R. J. LeVeque *et al.*, *Finite volume methods for hyperbolic problems*, vol. 31. Cambridge university press, 2002.

[37] R. F. Arritt and R. C. Dugan, "The ieee 8500-node test feeder," in *IEEE PES T&D 2010*, pp. 1–6, 2010.

[38] "High-resolution solar radiation datasets." https://www.nrcan.gc.ca/energy/renewable-electricity/solar-photovoltaic/18409.

[39] J. M. Chambers, *Graphical methods for data analysis*. Chapman and Hall/CRC, 2018.

Feng Li (Student Member, IEEE) received the B.Eng degree in electrical engineering from McGill University, Montreal, Canada in 2007, and the M.A.Sc degree in electrical engineering from Polytechnique Montréal, Canada in 2017. In 2018, he joined Eaton's CYME International T&D, Brossard, Canada as a senior power system engineer. He is currently working toward the PhD degree in electrical engineering at Polytechnique Montréal. His research interests include modeling, analysis, and optimization of power distribution networks.

Ilhan Kocar (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in EE from Orta Doğu Teknik Üniversitesi, Ankara, Turkey, in 1998 and 2003, respectively, and the Ph.D. degree in EE from École Polytechnique de Montréal (affiliated with Université de Montréal), QC, Canada, in 2009. He is currently a professor at the Hong Kong Polytechnic University, with more than 20 years of experience in research, software development, and consulting. His research interests include power system analysis, modeling, and simulation.

Antoine Lesage-Landry (Member, IEEE) is an Assistant Professor in the Department of Electrical Engineering at Polytechnique Montréal, QC, Canada. He received the B.Eng. degree in Engineering Physics from Polytechnique Montréal, in 2015, and the Ph.D. degree in Electrical Engineering from the University of Toronto, ON, Canada, in 2019. From 2019 to 2020, he was a Postdoctoral Scholar in the Energy & Resources Group at the University of California, Berkeley, CA, USA. His research interests include optimization, online learning, and their application to power systems with renewable generation.