

# Chance-Constrained AC Optimal Power Flow for Unbalanced Distribution Grids

Kshitij Girigoudar\*, Ashley M. Hou\*, and Line A. Roald

**Abstract**—The growing penetration of distributed energy resources (DERs) is leading to continually changing operating conditions, which need to be managed efficiently by distribution grid operators. The intermittent nature of DERs such as solar photovoltaic (PV) systems as well as load forecasting errors not only increase uncertainty in the grid, but also pose significant power quality challenges such as voltage unbalance and voltage magnitude violations. This paper leverages a chance-constrained optimization approach to reduce the impact of uncertainty on distribution grid operation. We first present the chance-constrained optimal power flow (CC-OPF) problem for distribution grids and discuss a reformulation based on constraint tightening that does not require any approximations or relaxations of the three-phase AC power flow equations. We then propose two iterative solution algorithms capable of efficiently solving the reformulation. In the case studies, the performance of both algorithms is analyzed by running simulations on the IEEE 13-bus test feeder using real PV and load measurement data. The simulation results indicate that both methods are able to enforce the chance constraints in in- and out-of-sample evaluations.

**Index Terms**—AC optimal power flow, chance constraints, distribution grids, uncertainty, voltage unbalance

## NOMENCLATURE

### Sets and Indices

$\mathcal{N}$	Set of nodes excluding slack bus, $ \mathcal{N}  = n$
0	Substation node (slack bus) index
$i, k, l \in \mathcal{N}$	Node index
$\Phi$	Set of phases, $\{a, b, c\}$
$\phi \in \Phi$	Phase index
$\Omega$	Uncertainty set, $ \Omega  = M$
$\omega \in \Omega$	Uncertainty realization
$\kappa$	Tuning iteration count

### Random Variables

$p_{L,i,\omega}^\phi, P_{L,i,\omega}$	Uncertain active power load demand
$q_{L,i,\omega}^\phi, Q_{L,i,\omega}$	Uncertain reactive power load demand
$\delta p_{L,i,\omega}^\phi, \delta P_{L,i,\omega}$	Deviation of uncertain active power load demand from average value
$p_{G,i,\omega}^\phi, P_{G,i,\omega}$	Uncertain active power generation of solar PV
$\delta p_{G,i,\omega}^\phi, \delta P_{G,i,\omega}$	Deviation of uncertain active power generation of solar PV from average value
$p_{i,\omega}^\phi, P_{i,\omega}$	Uncertain active power injection
$q_{i,\omega}^\phi, Q_{i,\omega}$	Uncertain reactive power injection

\*The first two authors contributed equally to this work. K. Girigoudar, A. M. Hou, and L. A. Roald are with the Department of Electrical and Computer Engineering, University of Wisconsin-Madison, USA (e-mail: girigoudar@wisc.edu, amhou@wisc.edu, roald@wisc.edu).

$Y_{\bar{v},i}^\phi$

Constraint violation indicator random variable

$\hat{E}_{\bar{v},i}^\phi$

Constraint empirical violation probability

$\hat{E}_{\bar{v}}^{\max}$

Worst case empirical violation probability for upper voltage constraint

### Parameters

$\gamma_{L,i}^\phi, \Gamma_{L,i}$

Power ratio of load demand

$\bar{p}_{L,i}^\phi, \bar{P}_{L,i}$

Average active power load demand

$\bar{q}_{L,i}^\phi, \bar{Q}_{L,i}$

Average reactive power load demand

$\bar{p}_{G,i}^\phi, \bar{P}_{G,i}$

Average active power generation of solar PV

$|s_{G,i}^\phi|, |S_{G,i}|$

Apparent power rating of solar PV

$\bar{p}_{i,\omega}^\phi, \bar{P}_{i,\omega}$

Average active power injection

$\bar{q}_{i,\omega}^\phi, \bar{Q}_{i,\omega}$

Average reactive power injection

$\epsilon_q$

Violation probability of solar PV inverter reactive power limits,  $\epsilon_q \in [0, 1]$

$G, B$

Nodal conductance and susceptance matrices of nodal admittance matrix  $Y$

$\underline{v}, \bar{v}$

Voltage magnitude limits

$\epsilon_v$

Violation probability of voltage magnitude limits,  $\epsilon_v \in [0, 1]$

$\bar{q}_{G,i}^\phi, \underline{q}_{G,i}^\phi$

Upper and lower PV inverter limits w.r.t. average generation

$\bar{\lambda}_{q,i}^\phi, \bar{\Lambda}_q$

Upper reactive power constraint tightening

$\underline{\lambda}_{q,i}^\phi, \underline{\Lambda}_q$

Lower reactive power constraint tightening

$\bar{\lambda}_{v,i}^\phi, \bar{\Lambda}_v$

Upper voltage constraint tightening

$\underline{\lambda}_{v,i}^\phi, \underline{\Lambda}_v$

Lower voltage constraint tightening

$\eta$

Tuning convergence tolerance

$s \in \mathbb{R}^+$

Tuning parameters

$s_{\min}, s_{\max}$

Upper and lower tuning parameter limits

$\sigma_{v,i}^\phi$

Estimated standard deviation of voltage magnitude

### Optimization Variables

$q_{G,i}^\phi, Q_{G,i}$

Reactive power generation of solar PV

$p_{G,0,\omega}^\phi, P_{G,0,\omega}$

Active power injection at substation corresponding to uncertain power injections

$q_{G,0,\omega}^\phi, Q_{G,0,\omega}$

Reactive power injection at substation corresponding to uncertain power injections

$p_{G,0}^\phi, P_{G,0}$

Active power injection at substation corresponding to average power injections

$q_{G,0}^\phi, Q_{G,0}$

Reactive power injection at substation corresponding to average power injections

$ v_{i,\omega}^\phi ,  V_{i,\omega} $	Voltage magnitude corresponding to uncertain power injections
$\theta_{i,\omega}^\phi, \Theta_{i,\omega}$	Voltage angle corresponding to uncertain power injections
$ v_i^\phi ,  V_i $	Voltage magnitude corresponding to average power injections
$\theta_i^\phi, \Theta_i$	Voltage angle corresponding to average power injections
$v_{dl,\omega}^-, v_{ql,\omega}^-$	Real and imaginary components of negative sequence voltage phasor $v_{l,\omega}^-$ corresponding to uncertain power injections
$v_{dl,\omega}^+, v_{ql,\omega}^+$	Real and imaginary components of positive sequence voltage phasor $v_{l,\omega}^+$ corresponding to uncertain power injections
$v_{dl}^-, v_{ql}^-$	Real and imaginary components of negative sequence voltage phasor $v_{l,\omega}^-$ corresponding to average power injections
$v_{dl}^+, v_{ql}^+$	Real and imaginary components of positive sequence voltage phasor $v_{l,\omega}^+$ corresponding to average power injections
<i>Functions</i>	
$\mathbb{P}_\omega(\cdot)$	Chance constraint probability
$C(\cdot), S(\cdot)$	Cosine and sine components of branch angle matrix
$f(\cdot)$	Power flow equations (short-hand)
$f_{q,i}^\phi(\cdot)$	Empirical distribution of reactive power limit
$f_{v,i}^\phi(\cdot)$	Empirical distribution of voltage limit

## I. INTRODUCTION

The increasing integration of distributed energy resources (DERs), such as rooftop solar photovoltaic (PV) systems and electric vehicles, poses both operational challenges and opportunities for distribution grid operations. Large scale penetration of DERs can lead to increases in power injection uncertainty and load variability. This is particularly challenging at the distribution level, where DERs are typically not equally distributed throughout the feeder. At the same time, the presence of DERs provide new opportunities for control, including using offline design of individual voltage droop control curves [1], [2], limited communication of real-time measurements [3], [4], or direct dispatching of inverters based on frequent resolving of centralized optimal power flow (OPF) problems [5]–[8]. While centralized OPF provides optimal set-points, limited real-time system measurements and communication delays can impact the frequency at which the set-points are updated. For example, smart meters can typically take measurements every 15 minutes, but this data is often only available to system operators with several hours delay [9]. As a result, it may be important to identify control set-points ensure system security for extended periods of time, with intervals ranging from 15 minutes (if good communication systems exist) to an entire season (if low complexity is desired). Significant variability and uncertainty in the load and DER power

injections is present across these time horizons, which, if not appropriately accounted for, can lead to constraint violations and high voltage unbalance between control set-point updates. To mitigate this problem, we propose treating the variable load and DER injections as uncertain and formulating the centralized OPF problem as a stochastic optimization problem.

Approaches for stochastic OPF with uncertain renewable energy generation and load have been well studied in the context of transmission systems [10]–[14]. However, methods developed for transmission systems often leverage DC power flow representations and single-phase equivalents to represent the system, which are not applicable to distributions grids, who are inherently unbalanced and exhibit high R/X ratios. Furthermore, while transmission operations focus primarily on congestion management and system balancing, distribution utilities focus on managing voltage magnitudes, voltage unbalance, and other power quality issues for end customers. As a result, there is a range of stochastic OPF models that have been developed specifically for distribution grids. This includes robust and distributionally robust methods [15]–[17], stochastic approximation techniques [18], and chance-constrained formulations [19]. These models focus primarily on managing or reducing voltage magnitude violations [15], [17] while minimizing objectives such as cost of energy [16], [17], [19], deviation from a desired power withdrawal at the substation [17], or losses [15], [18], [19]. They often leverage power flow formulations that rely on a radial network topology [20]–[24], include approximate representations of unbalance [25]–[28], and sometimes take advantage of iterative solution algorithms such as forward-backward sweep [6], [29], [30]. Many of these formulations are linear approximations [2], [20], [25] or convex relaxations [22], [23], [26]–[28], which may converge to solutions that are not actually AC feasible. Methods that do consider the full, non-linear, non-convex AC power flow formulation [16] or guarantee convergence to a solution that satisfies those equations [19] typically only consider balanced, single-phase systems.

In this paper, we take a more comprehensive view on voltage management than previous work and focus specifically on minimizing voltage unbalance, which causes damage to three-phase motors [31]. Existing approximate models [25]–[28] are inadequate for a detailed analysis of voltage unbalance, as they assume that the system is nearly balanced to start with. Instead, we build our optimization problem around a full, detailed representation of the three-phase power flow, including single-phase, two-phase, and untransposed three-phase lines that serve unbalanced loads [24].

To minimize voltage unbalance while ensuring that voltage magnitude constraints and DER capacity limits are satisfied, we adopt a chance-constrained formulation. In this model, we ensure that the voltage magnitude and inverter limits are enforced with high probability using single chance constraints. Chance constraints have the benefit of offering an intuitive trade-off between optimality (i.e., lowering voltage unbalance) and system robustness or reliability (i.e., limiting constraint violations) by adjusting the desired violation probability pa-

rameter. We note an important distinction between voltage and inverter limits: voltage limits are soft constraints, i.e., small and infrequent voltage violations may be acceptable and there is no inherent mechanism in place to enforce that they are always within bounds, while inverter constraints are hard limits, i.e., inverters have control systems in place to ensure that the apparent power from the inverter always stays at or below the maximum threshold as specified by the IEEE Standard 1547-2018 [32]. Therefore, voltage constraint violations reflect actual over- or under-voltages in the system, while inverter constraint violations reflect a situation in which the inverters will provide less apparent power than our system model assumes. A better system model therefore incorporates the capping of inverter power, i.e., curtailing the inverter reactive power to stay with the limit. However, modeling this type of capping behavior within the optimization problem can be challenging.

In general, chance-constrained optimization problems are hard to solve [33], as the probabilistic constraints must be reformulated to obtain a tractable representation. Methods to reformulate constraints in distribution grid applications include moment-based reformulations assuming Gaussian uncertainty [34] as well as distributionally agnostic methods via the use of, e.g., conservative convex approximations [35], [36], scenario approaches [37], or data-driven robust methods [17]. In this work, we leverage two data-driven, iterative approaches to solving AC chance-constrained optimal power flow (CC-OPF), inspired by existing methods for transmission networks [38]–[41] and balanced distribution networks [19]. The key idea is to represent each chance constraint using the nominal constraint with the addition of a tightening term (also known as an uncertainty margin) that secures the constraint against uncertainties. An optimally chosen tightening term will yield solutions that satisfy the chance constraints with the exact desired probability level. The methods iterate between solving an approximate, deterministic OPF with fixed tightening terms and using available uncertainty data to evaluate the resulting solution and update the tightening terms. A more principled approach to updating or tuning the tightening terms was developed in [42] and applied to transmission networks formulated with a DC linearized power flow approximation. This tuning-based approach solves a sequence of simpler optimization problems, resulting in better computationally tractability, while leveraging detailed system models in the evaluation process, where we simulate system behavior, to allow for the integration of more complex behavior such as inverter capping.

A particular challenge for stochastic and data-driven methods in distribution grids is the characterization of the uncertainty. Assumptions that may be true when considering large numbers of customers (e.g., the law of large numbers) no longer apply to the small number of consumers in distribution feeders, where both load and DER power injections are hard to accurately forecast and exhibit large variability over time. The most flexible way of modeling these distributions is using scenarios. However, the quality of the solution may be

sensitive to the choice of scenarios, and therefore choosing an appropriate set of samples (e.g., based on historical data) is a challenging yet important aspect that often is overlooked. To remedy this, we investigate how the number of data points and choice of scenario sets impact the solution quality. Using realistic load and solar PV data from Pecan Street [43], we assess how the solution changes as we use either entire days of data or randomly sampled data points across multiple days.

In summary, the contributions of the paper are as follows:

- 1) We present a chance-constrained formulation of the three-phase AC OPF problem which minimizes voltage unbalance by controlling inverter set-points. Our formulation extends existing work in the literature by considering the full three-phase, unbalanced AC power flow physics and incorporating the effect of inverter capping.
- 2) We propose a computationally tractable reformulation of our chance-constrained problem, extending the idea of uncertainty margins from [38], [39] to a distribution grid setting. We then compare two data-driven algorithms for identifying appropriate constraint tightening values: directly using the results of a Monte Carlo simulation (inspired by transmission grid AC CC-OPF in [39]) or tuning a safety factor (inspired by a transmission grid DC CC-OPF in [41]). Relative to prior work, these updated algorithms include a number of adaptations to make them applicable to the unbalanced distribution grid setting.
- 3) We perform a detailed case study using real load and PV data, where we investigate three important aspects of data-driven chance-constrained optimization in distribution grids:
  - a) We investigate how properties of the uncertainty data (i.e., sampling method and correlation between samples) can influence the resulting solutions.
  - b) We compare our two methods with respect to their ability to enforce chance constraint satisfaction.
  - c) We assess whether the proposed methods are able to effectively integrate the effect of hard PV inverter limits via the use of inverter capping.

The remainder of the paper is organized as follows: Section II presents the uncertainty modeling and formulation of the AC CC-OPF problem. Section III details the analytical reformulation and also provides a description of the iterative quantile-based and tuning-based solution algorithms. We introduce the test case and data sets used in the case study in Section IV and present the numerical results and analysis in Sections V, VI and VII. Section VIII concludes.

## II. PROBLEM FORMULATION

In this section, we describe our model of the PV inverters and load in the distribution grid. We then present a formulation of the three-phase chance-constrained AC OPF problem.

### A. Notation

We use the following phasor notation:  $X = |X|\angle\theta = X_d + jX_q$ , where  $X$  is a complex phasor,  $|X|$  and  $\theta$  represent the magnitude and angle components,  $X_d$  and  $X_q$  denote the real

and imaginary components, and  $j = \sqrt{-1}$ . All optimization variables are denoted using **bold** symbols. We use a subscript  $\omega$  to represent dependency on a given uncertainty realization  $\omega$  (discussed in detail in Section II-B). All scalar values are denoted using small letters, while all vector counterparts and matrices are denoted using capital letters. We represent the element-wise product of two vectors using  $\odot$ .

### B. Uncertainty Modeling

Due to the small number of households served by different parts of a distribution feeder, we do not observe the same smoothing effect as in transmission grids [24]. The electricity consumption and PV production of individual households is highly variable, as illustrated by the data in Fig. 1 with measurements obtained from Pecan Street for a set of households in New York [43]. Not only is the load very challenging to forecast, but the data is not necessarily well described by standard probability distributions. To capture the realistic probability distributions and dependence structure of the solar PV and loads, we propose to directly leverage historical data (which may, in many cases, be obtained by utilities from smart meters after a delay [9]). We denote the full data set as  $\Omega$ , with individual realizations represented as  $\omega \in \Omega$ . This sample set may include generation and load across certain days (i.e., representative days) or randomly drawn samples from a larger set of historical data (where the resulting sample set has no temporal structure).

We next describe our load and PV modeling assuming a network where  $\mathcal{N}$  denotes the set of three-phase nodes and  $\Phi = \{a, b, c\}$  represents the set of phases.

1) *Modeling of loads:* For each node  $i \in \mathcal{N}$ , we denote the three-phase active and reactive power consumption for a given realization  $\omega$  as  $P_{L,i,\omega} = [p_{L,i,\omega}^a \ p_{L,i,\omega}^b \ p_{L,i,\omega}^c]^\top$  and

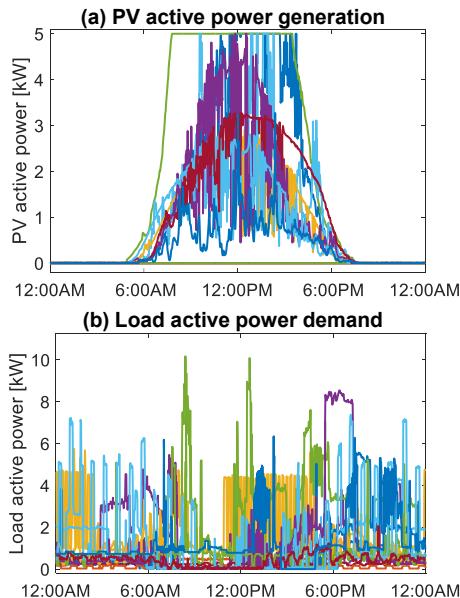


Fig. 1. One-minute resolution Pecan Street data for 15 houses from a single day. (a) PV active power generation. (b) Load active power demand.

$Q_{L,i,\omega} = [q_{L,i,\omega}^a \ q_{L,i,\omega}^b \ q_{L,i,\omega}^c]^\top$ , respectively. If reactive power measurements are not available (as will be the case in our case study), we assume that the loads operate with a constant power factor  $pf_i^\phi$ , giving rise to the following reactive power consumption:

$$Q_{L,i,\omega} = \Gamma_{L,i} \odot P_{L,i,\omega}, \quad \forall i \in \mathcal{N}, \omega \in \Omega, \quad (1)$$

where the constant factor  $\Gamma_{L,i} = [\gamma_{L,i}^a \ \gamma_{L,i}^b \ \gamma_{L,i}^c]^\top$  is computed based on the power factor  $pf_i^\phi$  with

$$\gamma_{L,i}^\phi = \sqrt{(1 - pf_i^\phi)^2 / (pf_i^\phi)^2}, \quad \forall \phi \in \Phi, i \in \mathcal{N}.$$

We define the average active load  $\bar{P}_{L,i} = [\bar{p}_{L,i}^a \ \bar{p}_{L,i}^b \ \bar{p}_{L,i}^c]^\top$  and average reactive load  $\bar{Q}_{L,i} = [\bar{q}_{L,i}^a \ \bar{q}_{L,i}^b \ \bar{q}_{L,i}^c]^\top$  as the sample average across all realizations in our sample set  $\Omega$  with

$$\bar{p}_{L,i}^\phi = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} p_{L,i,\omega}^\phi, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \quad (2a)$$

$$\bar{q}_{L,i}^\phi = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} q_{L,i,\omega}^\phi, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \quad (2b)$$

and express the uncertainty of the load consumption as deviations from this average,

$$\delta P_{L,i,\omega} = P_{L,i,\omega} - \bar{P}_{L,i}, \quad \forall i \in \mathcal{N}, \omega \in \Omega, \quad (3a)$$

$$\delta Q_{L,i,\omega} = Q_{L,i,\omega} - \bar{Q}_{L,i}, \quad \forall i \in \mathcal{N}, \omega \in \Omega. \quad (3b)$$

#### 2) Modeling of active power generation from solar PV:

The active power generation of the solar PV inverters under realization  $\omega$  is defined as  $P_{G,i,\omega} = [p_{G,i,\omega}^a \ p_{G,i,\omega}^b \ p_{G,i,\omega}^c]^\top$  and we define the average generation  $\bar{P}_{G,i} = [\bar{p}_{G,i}^a \ \bar{p}_{G,i}^b \ \bar{p}_{G,i}^c]^\top$  and uncertain deviation  $\delta P_{G,i,\omega}$  in a similar way as for the loads, i.e.,

$$\bar{p}_{G,i}^\phi = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} p_{G,i,\omega}^\phi, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \quad (4a)$$

$$\delta P_{G,i,\omega} = P_{G,i,\omega} - \bar{P}_{G,i}, \quad \forall i \in \mathcal{N}, \omega \in \Omega. \quad (4b)$$

We assume that the utility does not wish to curtail active power and therefore consider the active power of the PV inverters as uncontrollable random variables.

### C. Reactive Power Control from Solar PV Inverters

We assume that all solar PV inverters are equipped with smart inverters and provide opportunities for reactive power control as outlined in the IEEE Standard 1547-2018 [32]. For the purposes of this paper, we assume that the inverter is operating in the constant reactive power mode, where it provides reactive power according to a given set-point,

$$\mathbf{Q}_{G,i} = [q_{G,i}^a \ q_{G,i}^b \ q_{G,i}^c]^\top, \quad \forall i \in \mathcal{N}.$$

This setpoint is provided by the utility (without accounting for, e.g., local voltage measurements) and serves as a decision variable in our problem. Our goal is to identify a suitable set-point  $\mathbf{Q}_{G,i}$  that will remain the same for all uncertainty realizations  $\omega$  and contributes to minimizing voltage unbalance while keeping voltage magnitudes within bounds. Note that other types of reactive power control, such as the voltage

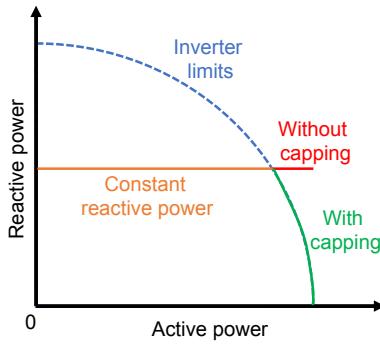


Fig. 2. Inverter reactive power control with and without capping.

droop-control outlined in [32], could also be considered but is deferred to future work.

We further assume that the active and reactive power of the PV inverter must adhere to limits on apparent power. For a single-phase PV inverter connected to phase  $\phi$  at node  $i$  with apparent power rating  $|s_{G,i}^\phi|$ , the inverter reactive power  $q_{G,i}^\phi$  is constrained by

$$\mathbb{P}_\omega((q_{G,i}^\phi)^2 + (p_{G,i,\omega}^\phi)^2 \leq |s_{G,i}^\phi|) \geq 1 - \epsilon_q, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \omega \in \Omega. \quad (5)$$

Because the active power generation  $p_{G,i,\omega}^\phi$  varies according to uncertainty realizations  $\omega$ , this constraint is enforced as a chance constraint with acceptable violation probability  $\epsilon_q \in [0, 1]$ . The chosen reactive power set-point  $q_{G,i}^\phi$  is guaranteed to be feasible with probability  $1 - \epsilon_q$ .

Fig. 2 illustrates the PV inverter limit (blue dashed line), which is assumed to be a hard constraint, and the reactive power is constant (orange line) with respect to the active power. If the reactive power remains fixed beyond a certain level of active power generation, the PV inverter is overloaded and operates in the region where the inverter limits are violated, indicated by the red line. Operating under overloaded conditions for long duration of time can lead to premature failure of the inverter, potentially requiring manual intervention. To avoid this, we consider capping the inverter to be within the specified limits by reducing the reactive power, as shown by the green curve. This also complies with the IEEE Std 1547-2018 [32], which states that DERs are required to respect their apparent power limits. For such cases, the violation probability  $\epsilon_q$  should be interpreted as the probability that an inverter is not able to provide the desired reactive power to the grid.

#### D. Grid Model

We consider a three-phase distribution grid with one slack node and a set of remaining nodes  $\mathcal{N}$ , with  $n = |\mathcal{N}|$ . While we limit our case study to radial distribution grids, our model makes no assumptions of radiality and could thus also be applied to distribution grids that are operated in non-radial configurations. Distribution grids typically have several single and two-phase nodes. However, to simplify notation, we consider all nodes to have three phases. For the single and two-phase nodes, the corresponding entries for the missing

phases are set to zero. The resulting total number of single-phase connections in the distribution grid is  $3(n + 1)$ .

The distribution substation is chosen as the slack node with index  $i = 0$ . We assume that there is one solar PV inverter and one load at each single-phase connection of every node  $i \in \mathcal{N}$ . If any node connected to a phase has no source or load, we set the corresponding entries to zero.

1) *Grid parameters*: Following [44], the critical distribution grid components (including distribution lines, cables, and transformers) are modeled using an overall nodal admittance matrix

$$Y = G + jB \in \mathbb{C}^{3(n+1) \times 3(n+1)}, \quad (6)$$

where  $G, B \in \mathbb{R}^{3(n+1) \times 3(n+1)}$  denote the nodal conductance and susceptance matrices, respectively.

2) *Nodal power injections*: The active power injection  $P_{i,\omega} \in \mathbb{R}^3$  at node  $i \in \mathcal{N}$  varies with the uncertainty realization  $\omega$ . It is modeled as the sum of the average active generation  $\bar{P}_{G,i} \in \mathbb{R}^3$  and average load  $\bar{P}_{L,i} \in \mathbb{R}^3$  with corresponding deviations  $\delta P_{G,i,\omega} \in \mathbb{R}^3$  and  $\delta P_{L,i,\omega} \in \mathbb{R}^3$ ,

$$P_{i,\omega} = \underbrace{(\bar{P}_{G,i} + \delta P_{G,i,\omega})}_{P_{G,i,\omega}} - \underbrace{(\bar{P}_{L,i} + \delta P_{L,i,\omega})}_{P_{L,i,\omega}}, \quad \forall i \in \mathcal{N}, \omega \in \Omega. \quad (7)$$

The reactive power injection  $Q_{i,\omega} \in \mathbb{R}^3$  at node  $i \in \mathcal{N}$  is modeled as the difference between the controllable reactive power generation of solar PV inverters  $Q_{G,i}$  and the reactive power load  $Q_{L,i,\omega}$ , i.e.,  $Q_{i,\omega} = Q_{G,i} - Q_{L,i,\omega}$ . The average active and reactive power injections at each node are defined as  $\bar{P}_i = \bar{P}_{G,i} - \bar{P}_{L,i}$  and  $\bar{Q}_i = Q_{G,i} - Q_{L,i}$ , respectively.

The active power balance is maintained by the substation, which supplies the difference between the active power generation and load for each scenario as well as any additional power needed to cover the power losses. We treat the active power injection at the substation as a decision variable  $P_{G,0,\omega}$ . Similarly, the substation node also guarantees reactive power balance. We define the reactive power injection at the substation as a decision variable  $Q_{G,0,\omega}$ , which covers the difference between production, consumption, and losses.

3) *Voltage representation*: The three-phase OPF is implemented in the polar coordinate frame using the phase-to-neutral voltage magnitude and angle variables at every node  $i \in \mathcal{N}$ . We denote the voltage magnitudes and angles corresponding to the average power injections  $\bar{P}_i, \bar{Q}_i$  using  $|\mathbf{V}_i| = [|\mathbf{v}_i^a| \ |\mathbf{v}_i^b| \ |\mathbf{v}_i^c|]^\top$  and  $\Theta_i = [\theta_i^a \ \theta_i^b \ \theta_i^c]^\top$ , respectively. As the active and reactive power injections change, the voltages change as well. The voltage magnitude and angle for a given realization  $\omega$  is denoted by  $|\mathbf{V}_{i,\omega}| = [|\mathbf{v}_{i,\omega}^a| \ |\mathbf{v}_{i,\omega}^b| \ |\mathbf{v}_{i,\omega}^c|]^\top$  and  $\Theta_{i,\omega} = [\theta_{i,\omega}^a \ \theta_{i,\omega}^b \ \theta_{i,\omega}^c]^\top$ . The distribution substation is considered as the reference for voltage angle measurements. We assume the voltage is independent of  $\omega$  and fixed at the substation, i.e.,

$$|\mathbf{V}_0| \angle \Theta_0 = [1 \angle 0^\circ \quad 1 \angle -120^\circ \quad 1 \angle 120^\circ]^\top. \quad (8)$$

All other voltage magnitudes are constrained by

$$\mathbb{P}_\omega(|\mathbf{v}_{i,\omega}^\phi| \leq \bar{v}) \geq 1 - \epsilon_v, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \omega \in \Omega, \quad (9)$$

$$\mathbb{P}_\omega(|\mathbf{v}_{i,\omega}^\phi| \geq \underline{v}) \geq 1 - \epsilon_v, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \omega \in \Omega, \quad (10)$$

where  $\underline{v}$  and  $\bar{v} \in \mathbf{R}$  are the respective lower and upper voltage magnitude bounds. Because the voltage magnitude  $|\mathbf{v}_{i,\omega}^\phi|$  are dependent on the uncertainty realization  $\omega$ , these constraints are enforced as with acceptable violation probability  $\epsilon_v \in [0, 1]$ . The voltage magnitude constraints can be considered as *soft* constraints, where a constraint violation indicates an under- or over-voltage. The violations of soft constraints is acceptable if the duration and magnitude are small.

4) *Power flow*: Following [44], we express the power balance equation at node  $i \in \{0, \mathcal{N}\}$  using

$$\begin{aligned} P_{i,\omega} = & |\mathbf{V}_{i,\omega}| \odot \sum_{k \in \{0, \mathcal{N}\}} \left[ G_{ik} \odot \mathbf{C}(\Theta_{ik,\omega}) \right] |\mathbf{V}_{k,\omega}| \\ & + |\mathbf{V}_{i,\omega}| \odot \sum_{k \in \{0, \mathcal{N}\}} \left[ B_{ik} \odot \mathbf{S}(\Theta_{ik,\omega}) \right] |\mathbf{V}_{k,\omega}|, \end{aligned} \quad (11a)$$

$$\begin{aligned} Q_{i,\omega} = & |\mathbf{V}_{i,\omega}| \odot \sum_{k \in \{0, \mathcal{N}\}} \left[ G_{ik} \odot \mathbf{S}(\Theta_{ik,\omega}) \right] |\mathbf{V}_{k,\omega}| \\ & - |\mathbf{V}_{i,\omega}| \odot \sum_{k \in \{0, \mathcal{N}\}} \left[ B_{ik} \odot \mathbf{C}(\Theta_{ik,\omega}) \right] |\mathbf{V}_{k,\omega}|, \end{aligned} \quad (11b)$$

where  $G_{ik}, B_{ik} \in \mathbb{R}^{3 \times 3}$  represent the real and imaginary sub-matrices of the nodal admittance matrix  $Y$  for a three-phase branch  $ik$ . We represent the cosine and sine components of the branch angle matrix  $\Theta_{ik,\omega} \in \mathbb{R}^{3 \times 3}$  using  $\mathbf{C}(\Theta_{ik,\omega})$  and  $\mathbf{S}(\Theta_{ik,\omega})$ , respectively. Note that the equations (11) represents the AC power flow equations for all  $\omega \in \Omega$ . For the remainder of the paper, we will use

$$f(|\mathbf{V}_{i,\omega}|, \Theta_{i,\omega}, P_{i,\omega}, Q_{i,\omega}) = 0, \quad \forall i \in \{0, \mathcal{N}\}, \omega \in \Omega,$$

as a shorthand representation of the power balance constraints defined in (11).

### E. Objective Function

Our objective for the OPF problem is to minimize voltage unbalance. We utilize the IEC standard 61000-2-2 [45] commonly referred to as Voltage Unbalance Factor (VUF) to define voltage unbalance. For a three-phase node  $l$ , the square of VUF for the uncertainty realization  $\omega$  is expressed as

$$\text{VUF}_{l,\omega}^2 = \frac{|\mathbf{v}_{l,\omega}^-|^2}{|\mathbf{v}_{l,\omega}^+|^2} = \frac{(\mathbf{v}_{dl,\omega}^-)^2 + (\mathbf{v}_{ql,\omega}^-)^2}{(\mathbf{v}_{dl,\omega}^+)^2 + (\mathbf{v}_{ql,\omega}^+)^2}, \quad (12)$$

where  $\mathbf{v}_{dl,\omega}^-$ ,  $\mathbf{v}_{ql,\omega}^-$  and  $\mathbf{v}_{dl,\omega}^+$ ,  $\mathbf{v}_{ql,\omega}^+$  are the rectangular form representation of negative sequence voltage  $\mathbf{v}_{l,\omega}^-$  and positive sequence voltage  $\mathbf{v}_{l,\omega}^+$ , respectively, and expressed as non-linear functions of our voltage variables  $\mathbf{V}_{l,\omega}, \Theta_{l,\omega}$  using

$$\mathbf{v}_{dl,\omega}^- = \text{Re}\{\mathbf{v}_{l,\omega}^-\}, \mathbf{v}_{ql,\omega}^- = \text{Im}\{\mathbf{v}_{l,\omega}^-\}, \text{ where} \quad (13a)$$

$$\mathbf{v}_{l,\omega}^- = |\mathbf{v}_{l,\omega}^a| \angle \theta_{l,\omega}^a + |\mathbf{v}_{l,\omega}^b| \angle (\theta_{l,\omega}^b - 120^\circ) + |\mathbf{v}_{l,\omega}^c| \angle (\theta_{l,\omega}^c + 120^\circ),$$

$$\mathbf{v}_{dl,\omega}^+ = \text{Re}\{\mathbf{v}_{l,\omega}^+\}, \mathbf{v}_{ql,\omega}^+ = \text{Im}\{\mathbf{v}_{l,\omega}^+\}, \text{ where} \quad (13b)$$

$$\mathbf{v}_{l,\omega}^+ = |\mathbf{v}_{l,\omega}^a| \angle \theta_{l,\omega}^a + |\mathbf{v}_{l,\omega}^b| \angle (\theta_{l,\omega}^b + 120^\circ) + |\mathbf{v}_{l,\omega}^c| \angle (\theta_{l,\omega}^c - 120^\circ).$$

### F. Chance-Constrained Optimal Power Flow

The three-phase AC CC-OPF problem is formulated as follows:

$$\begin{aligned} \min_{\substack{P_{G,0}, Q_G, \\ |\mathbf{V}|, \Theta}} \quad & \sum_{\omega \in \Omega} \sum_{l \in \mathcal{N}} \text{VUF}_{l,\omega}^2 & & \text{(CC-OPF)} \\ \text{s.t.} \quad & f(|\mathbf{V}_{i,\omega}|, \Theta_{i,\omega}, P_{i,\omega}, Q_{i,\omega}) = 0, \quad \forall i \in \{0, \mathcal{N}\}, \omega \in \Omega, \\ & \mathbb{P}_\omega(|\mathbf{v}_{i,\omega}^\phi| \leq \bar{v}) \geq 1 - \epsilon_v, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \\ & \mathbb{P}_\omega(|\mathbf{v}_{i,\omega}^\phi| \geq \underline{v}) \geq 1 - \epsilon_v, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \\ & \mathbb{P}_\omega((\mathbf{q}_{G,i}^\phi)^2 + (p_{G,i,\omega}^\phi)^2 \leq |s_{G,i}^\phi|) \geq 1 - \epsilon_q, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \\ & |\mathbf{V}_0| \angle \Theta_0 = [1 \angle 0^\circ \quad 1 \angle -120^\circ \quad 1 \angle 120^\circ]^\top, \end{aligned}$$

where the optimization decision variables for each uncertainty realization  $\omega \in \Omega$  are the voltage magnitudes at all nodes  $|\mathbf{V}| \in \mathbb{R}^{3(n+1)}$ , voltage angles at all nodes  $\Theta \in \mathbb{R}^{3(n+1)}$ , active power generation at the substation  $P_{G,0} \in \mathbb{R}^3$ , and reactive power generation at all nodes  $Q_G \in \mathbb{R}^{3(n+1)}$ .

In general, we interpret the chance constraints (5) and (10) as the likelihood that the distribution grid operator will need to take extra control actions in real time to protect the system. By choosing high acceptable violation probabilities  $\epsilon_q$  and  $\epsilon_v$ , the system is at a higher risk of insecure operation as it may necessitate frequent deployment of real-time controls, which are not always available. Alternatively, choosing a low violation probability is expensive, but makes system operation safer and less stressful for the operator [39].

## III. ANALYTICAL REFORMULATION AND SOLUTION METHODS

As formulated, CC-OPF is intractable due to the chance constraints (5) and (10). Not only is the feasible region defined by the chance constraints non-convex, evaluating the feasibility of a solution is generally difficult as it requires multi-dimensional integration. Furthermore, it can also be computationally expensive to enforce the power balance constraints (11) for all  $\omega \in \Omega$ , particularly if the uncertainty set  $\Omega$  is large. As a result, it is necessary to reformulate these constraints into deterministic counterparts in order to obtain a tractable formulation that can be efficiently solved. Our problem setting yields two additional complications. First, we do not assume any distributional assumptions on the uncertainty in the problem. Second, we must account for the uncertainty in the full, non-linear, non-convex AC power flow constraints. Consequently, CC-OPF falls under a challenging class of problems, where typical analytical reformulation methods cannot be applied.

To solve this problem, we take an iterative, data-driven approach that combines an approximate problem formulation with sample-based evaluations to successively adapt the formulation, as illustrated in Fig. 3. Inspired by [39], [41], we replace (11) by a single set of deterministic power flow equations and represent the chance constraints (5) and (10) using their nominal counterpart plus a fixed tightening term that secures the system against uncertainty by functioning as an uncertainty

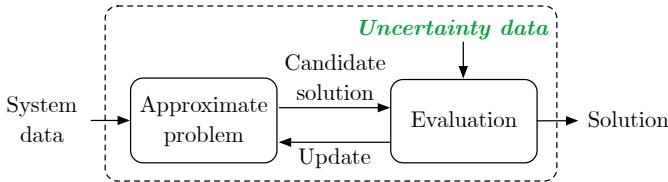


Fig. 3. The general iterative process, consisting of solving an approximate problem formulation and using the results of a sample-based evaluation to update the formulation.

margin. The tightening terms are successively updated using evaluated feedback from the available uncertainty data.

The primary advantage of the iterative approach is that the solving of the three-phase AC OPF problem is *decoupled* from the consideration of the uncertainty, which is wholly captured via the tightenings. This allows us to leverage the available uncertainty data without increasing the computational burden of solving the optimization problem. Thus, in contrast to other sample-based methods such as the scenario approach, the size of the optimization does not grow with the number of samples [42]. Furthermore, we are able to use the full non-linear, non-convex AC power flow equations, without needing to perform any explicit linearizations or approximations. Because we deliberately choose to use an approximate reformulation that is simple to solve, the resulting algorithm can also take advantage of the computational efficiency of commercial solvers.

A challenge of using this style of deterministic reformulation is accurately determining and updating the values of the uncertainty margins. We need to ensure that the uncertainty margins are large enough that the solution to the reformulated problem is ensured to be feasible to the original chance constrained problem, yet small enough so that we do not obtain an overly conservative solution. We thus explore two methods for iteratively updating constraint tightenings: a direct quantile-based method (inspired by [39]) and a tuning-based method (based on [41]). The remainder of this section details the main components of the iterative approach, including the two tuning variations.

#### A. Approximate Problem Formulation

As the approximate problem formulation, we choose to use a generalized reformulation based on [39]. In the following, we detail the reformulated constraints and objective function.

1) *Power flow*: Rather than enforcing the power flow equations in (11) for all uncertainty realizations, we enforce a single set of power flow equations at node  $i \in \{0, \mathcal{N}\}$  which are functions of the average power injections,  $\bar{P}_i$ ,  $\bar{Q}_i$  and corresponding voltage variables  $|\mathbf{V}_i|$ ,  $\Theta_i$ . The shorthand representation of the power balance constraint becomes

$$f(|\mathbf{V}_i|, \Theta_i, \bar{P}_i, \bar{Q}_i) = 0, \quad \forall i \in \{0, \mathcal{N}\}.$$

2) *Inverter limits*: We replace the inverter limits defined by quadratic chance constraints in (5) with deterministic box constraints, consisting of a nominal limit with a tightening term. The upper and lower limits for a PV inverter connected

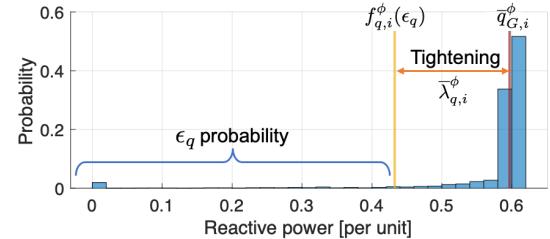


Fig. 4. The histogram represents the probability distribution function of an example inverter reactive power upper limit,  $f_{q,i}^\phi(\cdot)$ . The constraint tightening  $\bar{\lambda}_{q,i}^\phi$ , illustrated with the orange line, is given by the distance between the nominal limit  $\bar{q}_{G,i}^\phi$  (brown line) and the  $\epsilon_q$  quantile of the empirical distribution of the upper reactive power limit  $f_{q,i}^\phi(\epsilon_q)$  (yellow line).

to phase  $\phi \in \Phi$  at node  $i \in \mathcal{N}$  are calculated using the average generation  $\bar{p}_{G,i}^\phi$  as follows:

$$\bar{q}_{G,i}^\phi = \sqrt{|s_{G,i}^\phi|^2 - (\bar{p}_{G,i}^\phi)^2}, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \quad (15a)$$

$$\underline{q}_{G,i}^\phi = -\sqrt{|s_{G,i}^\phi|^2 - (\bar{p}_{G,i}^\phi)^2}, \quad \forall \phi \in \Phi, i \in \mathcal{N}. \quad (15b)$$

The tightenings for the upper and lower limits are denoted by  $\bar{\lambda}_{q,i}^\phi, \underline{\lambda}_{q,i}^\phi \in \mathbb{R}^+$ , respectively. For notational convenience, we also denote the tightenings using vectors

$$\underline{\Lambda}_q = \left[ [\underline{\lambda}_{q,i}^\phi]_{\phi \in \Phi}^\top \right]_{i \in \mathcal{N}}^\top, \quad \bar{\Lambda}_q = \left[ [\bar{\lambda}_{q,i}^\phi]_{\phi \in \Phi}^\top \right]_{i \in \mathcal{N}}^\top.$$

Note that these tightenings can be pre-calculated since the inverter reactive power limits depend only on the uncertainty samples which are known a priori. E.g., for the upper inverter limit, we evaluate the limit  $\bar{q}_{G,i,\omega}^\phi$  for each uncertainty sample  $\omega \in \Omega$  using Eq. (15a) with the sampled value  $p_{G,i,\omega}$  rather than the averaged value  $\bar{p}_{G,i}^\phi$ . By calculating this limit under all uncertainty realizations, we obtain an empirical distribution for the upper reactive power generation limit, which we denote as  $f_{q,i}^\phi(\cdot)$ . The empirical distribution for the lower reactive power generation limit would simply be the negative, i.e.,  $-f_{q,i}^\phi(\cdot)$ . We then find the desired quantiles of the empirical distribution and use them to directly calculate the constraint tightenings. For the chance constraint on the upper reactive power limit to hold, we require the upper limit to be set to  $f_{q,i}^\phi(\epsilon_q)$ , which is the  $\epsilon_q$  quantile of the empirical distribution of the upper reactive power limit. We can observe that the appropriate uncertainty margin to ensure the chance constraint holds corresponds to the difference between the nominal inverter lower limit  $\underline{q}_{G,i}^\phi$  and  $f_{q,i}^\phi(\epsilon_q)$ . Thus, the constraint tightenings for reactive power limits are calculated as follows:

$$\bar{\lambda}_{q,i}^\phi = \bar{q}_{G,i}^\phi - f_{q,i}^\phi(\epsilon_q), \quad \forall \phi \in \Phi, i \in \mathcal{N},$$

$$\underline{\lambda}_{q,i}^\phi = -f_{q,i}^\phi(1 - \epsilon_q) - \underline{q}_{G,i}^\phi, \quad \forall \phi \in \Phi, i \in \mathcal{N}.$$

This process is also illustrated in Fig. 4. We note that once these tightening values are calculated, they remain fixed throughout the entirety of the iterative algorithm.

3) *Voltage limits*: We similarly replace the chance constraints on the upper and lower voltage magnitude constraints

(10) with the corresponding nominal limits and tightenings, defined by  $\bar{\lambda}_{v,i}^\phi, \underline{\lambda}_{v,i}^\phi \in \mathbb{R}^+$ , respectively. The vector representations are

$$\underline{\lambda}_V = \left[ \left[ \lambda_{V,i}^{\phi \in \Phi} \right]_{i \in \mathcal{N}} \right]^\top, \quad \bar{\lambda}_V = \left[ \left[ \bar{\lambda}_{V,i}^{\phi \in \Phi} \right]_{i \in \mathcal{N}} \right]^\top.$$

In contrast to the inverter limits, the voltage constraint tightenings need to be calculated and updated within the iterative algorithm since they require obtaining a solution from the approximate OPF problem in order to be determined.

4) *Objective function*: We calculate the deterministic VUF by replacing the voltage variables  $V_{l,\omega}, \Theta_{l,\omega}$  with their nominal counterparts  $|V_l|, \Theta_l$  (corresponding to average power injections) in (12) and (13).

5) *Reformulated problem*: The resulting approximate optimization problem obtained from using the reformulation described above is deterministic and takes the following form:

$$\begin{aligned} & \min_{P_{G,0}, Q_G, |\mathbf{V}|, \Theta} \sum_{l \in \mathcal{N}} \text{VUF}_l^2 & & \text{(CCR-OPF)} \\ \text{s.t.} \quad & f(|V_i|, \Theta_i, \bar{P}_i, \bar{Q}_i) = 0, \quad \forall i \in \{0, \mathcal{N}\}, \\ & |\mathbf{v}_i^\phi| \leq \bar{v} - \bar{\lambda}_{v,i}^\phi, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \\ & |\mathbf{v}_i^\phi| \geq v + \underline{\lambda}_{v,i}^\phi, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \\ & q_{G,i}^\phi \leq \bar{q}_{G,i}^\phi - \bar{\lambda}_{q,i}^\phi, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \\ & q_{G,i}^\phi \geq \underline{q}_{G,i}^\phi + \underline{\lambda}_{q,i}^\phi, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \\ & \mathbf{V}_0 \angle \Theta_0 = [1 \angle 0^\circ \quad 1 \angle -120^\circ \quad 1 \angle 120^\circ]^\top. \end{aligned}$$

We denote a solution of CCR-OPF using  $\mathbf{X} = (P_{G,0}, Q_G, |\mathbf{V}|, \Theta)$ .

### B. Solution Evaluation

After we solve the approximate problem (CCR-OPF) and obtain a candidate solution  $\mathbf{X}$ , we use the available uncertainty samples to evaluate the conservativeness of the voltage constraint tightenings. We can either evaluate the solution  $\mathbf{X}$  by creating an empirical distribution and performing a quantile evaluation (as done to obtain the inverter limit tightenings in Section III-A2) or by estimating the empirical violation probability of the solution. Both methods are detailed below.

1) *Quantile evaluation*: We calculate power flows for each sample in our uncertainty set using the candidate solution  $\mathbf{X}$ . By doing so, we obtain empirical distributions for the voltage magnitude around the candidate voltage solution  $|\mathbf{v}_i^\phi|$  at node  $i \in \mathcal{N}$  connected to phase  $\phi \in \Phi$ , which we denote with  $f_{v,i}^\phi(\cdot)$ . Similarly to the evaluation of the inverter limit tightenings, we use the empirical distribution to directly evaluate the voltage magnitude value at our desired quantiles. As a result, we obtain  $f_{v,i}^\phi(1 - \epsilon_v)$  and  $f_{v,i}^\phi(\epsilon_v)$  as the respective upper  $1 - \epsilon_v$  and lower  $\epsilon_v$  quantiles of the empirical distribution  $f_{v,i}^\phi(\cdot)$ .

2) *Empirical violation probability evaluation*: Alternatively, we can use the uncertainty samples to estimate the

empirical violation probability of the candidate solution  $\mathbf{X}$  to ascertain whether it satisfies the chance constraint. As an example, let us take the upper voltage magnitude constraint for node  $i \in \mathcal{N}$  connected to phase  $\phi \in \Phi$  and samples  $p_{G,i,\omega}^\phi$  and  $p_{L,i,\omega}^\phi$ . We define an indicator random variable  $Y_{\bar{v},i}^\phi(\mathbf{X}, p_{G,i,\omega}^\phi, p_{L,i,\omega}^\phi) \in \{0, 1\}$ , which evaluates to 0 if the upper voltage constraint holds and 1 otherwise, i.e.,

$$Y_{\bar{v},i}^\phi(\mathbf{X}, p_{G,i,\omega}^\phi, p_{L,i,\omega}^\phi) = \begin{cases} 0 & \text{if } |\mathbf{v}_{i,\omega}^\phi| \leq \bar{v}, \\ 1 & \text{otherwise.} \end{cases} \quad (17)$$

We obtain the empirical violation probability for this constraint, denoted  $\hat{E}_{\bar{v},i}^\phi(\mathbf{X}) \in [0, 1]$ , by taking the average evaluated on all samples, i.e.,

$$\hat{E}_{\bar{v},i}^\phi(\mathbf{X}) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} Y_{\bar{v},i}^\phi(\mathbf{X}, p_{G,i,\omega}^\phi, p_{L,i,\omega}^\phi). \quad (18)$$

We perform similar evaluations to obtain the empirical violation probabilities of the voltage lower limits  $\hat{E}_{\underline{v},i}^\phi(\mathbf{X})$ , inverter upper limits  $\hat{E}_{\bar{q},i}^\phi(\mathbf{X})$ , and inverter lower limits  $\hat{E}_{\underline{q},i}^\phi(\mathbf{X})$ . For a solution  $\mathbf{X}$ , we define the worst case empirical violation probability for a single constraint type as the maximum empirical violation probability observed across all constraints of that type. For, e.g., the worst case empirical violation probability for the voltage upper limit is

$$\hat{E}_{\bar{v}}^{\max}(\mathbf{X}) = \max_{i \in \mathcal{N}, \phi \in \Phi} \{\hat{E}_{\bar{v},i}^\phi\}. \quad (19)$$

We further define the worst case empirical across all voltage constraints as

$$\hat{E}_v^{\max}(\mathbf{X}) = \max\{\hat{E}_{\bar{v}}^{\max}(\mathbf{X}), \hat{E}_{\underline{v}}^{\max}(\mathbf{X})\}. \quad (20)$$

3) *Handling of Inverter Limits*: As noted in Section II-C, we assume PV inverter limits are hard constraints. Consequently, as we evaluate the solution  $\mathbf{X}$  in this step, we propose two methods to deal with the situation where the inverter limit specified in (5) is violated:

- (i) We allow inverters to be overloaded and observe violation of the inverter limits.
- (ii) We cap the inverter reactive power to be within the apparent power limit.

We note that the latter approach (i.e., with capping) is a more realistic scenario. While the capping of the inverter output is difficult to represent as a constraint within the optimization problem itself, we can easily perform this a posteriori during the solution evaluation step.

### C. Iterative Schemes

We next describe the two iterative algorithms used to solve the deterministic CCR-OPF.

1) *Quantile-based method*: Based on [19], [38], [39], this method uses available samples to generate empirical distribution functions for the voltage variables. We then follow Section III-B1 to find the desired quantiles of the empirical distribution, which are used to calculate the tightenings for the voltage magnitude limits. The quantile-based method comprises of the following steps:

(1) *Initialize*: Set the iteration count to  $\kappa = 0$  and voltage magnitude tightenings  $\bar{\Lambda}_v^{(0)}, \underline{\Lambda}_v^{(0)}$  to zero. The reactive power generation tightenings  $\bar{\Lambda}_q, \underline{\Lambda}_q$  are calculated according to the procedure described in Section [III-A2](#).

(2) *Solve approximate problem*: Solve CCR-OPF using fixed tightenings  $\bar{\Lambda}_v^{(\kappa)}, \underline{\Lambda}_v^{(\kappa)}, \bar{\Lambda}_q, \underline{\Lambda}_q$  to obtain a candidate operating point  $\mathbf{X}^{(\kappa+1)}$ .

(3) *Solution evaluation*: Perform the quantile evaluation described in Section [III-B1](#) to obtain the upper  $1 - \epsilon_v$  and lower  $\epsilon_v$  quantiles of the empirical distribution of the voltage magnitude,  $f_{v,i}^{\phi,(\kappa+1)}(1 - \epsilon_v)$  and  $f_{v,i}^{\phi,(\kappa+1)}(\epsilon_v)$ .

(4) *Update tightenings*: Use the evaluated quantiles to directly update the voltage magnitude constraint tightenings,

$$\begin{aligned}\bar{\lambda}_{v,i}^{\phi} &= f_{v,i}^{\phi,(\kappa+1)}(1 - \epsilon_v) - |v_i^{\phi,(\kappa+1)}|, \quad \forall \phi \in \Phi, \forall i \in \mathcal{N}, \\ \underline{\lambda}_{v,i}^{\phi} &= |v_i^{\phi,(\kappa+1)}| - f_{v,i}^{\phi,(\kappa+1)}(\epsilon_v), \quad \forall \phi \in \Phi, \forall i \in \mathcal{N}.\end{aligned}$$

(5) *Check convergence*: Terminate when the upper and lower voltage constraint tightenings converge below their respective pre-specified tolerance levels  $\bar{\eta}_v, \underline{\eta}_v \in \mathbb{R}^+$ , i.e., both

$$\max\{|\bar{\Lambda}_v^{(\kappa+1)} - \bar{\Lambda}_v^{(\kappa)}|, |\underline{\Lambda}_v^{(\kappa+1)} - \underline{\Lambda}_v^{(\kappa)}|\} \leq \bar{\eta}_v, \quad \max\{|\bar{\Lambda}_v^{(\kappa+1)} - \bar{\Lambda}_v^{(\kappa)}|, |\underline{\Lambda}_v^{(\kappa+1)} - \underline{\Lambda}_v^{(\kappa)}|\} \leq \underline{\eta}_v$$

are satisfied. Return final solution  $\mathbf{X}^{(\kappa+1)}$ . Otherwise, increase the iteration count to  $\kappa = \kappa + 1$  and return to step (2).

2) *Tuning-based method*: This solution algorithm employs a tuning-based approach adapted from [\[41\]](#). Rather than directly using the quantile evaluation to calculate the uncertainty margins, we define the tightenings as parameterized by a single-dimensional tuning parameter, denoted  $s \in \mathbb{R}^+$ . This parameter is then iteratively adjusted based on the empirical violation probability evaluation of the current candidate solution. We follow [\[41\]](#) and use a simple bisection search to adjust  $s$  (although we note that various other tuning procedures can also be used).

Before detailing the algorithm steps, we first define the tightening as a product between the tuning parameter  $s$  and an approximation or estimation of the standard deviation of the nominal constraint. The tightenings are symmetric for the upper and lower voltage magnitude constraints and take the following form:

$$\bar{\lambda}_{v,i}^{\phi} = \underline{\lambda}_{v,i}^{\phi} = s \cdot \sigma_{v,i}^{\phi}, \quad \forall \phi \in \Phi, i \in \mathcal{N}, \quad (21)$$

where  $\sigma_{v,i}^{\phi} \in \mathbb{R}^+$  represents the estimated standard deviation of the voltage magnitude at node  $i \in \mathcal{N}$  connected to phase  $\phi \in \Phi$ . To obtain this estimate, we solve CCR-OPF using voltage limit tightenings  $\bar{\Lambda}_v = \underline{\Lambda}_v = 0$  and inverter limit tightenings  $\bar{\Lambda}_q, \underline{\Lambda}_q$  initialized according to Section [III-A2](#). We obtain an operating point  $\mathbf{X}^{(0)}$  and evaluate the empirical distributions  $f_{v,i}^{\phi,(\kappa)}(\cdot)$  according to Section [III-B2](#). We set  $\sigma_{v,i}^{\phi}$  as the standard deviation of  $f_{v,i}^{\phi,(\kappa)}(\cdot)$ .

The operating point  $\mathbf{X}^{(0)}$  is also used to initialize upper bound  $s_{\max}^{(0)}$  and lower bound  $s_{\min}^{(0)}$  of the tuning parameter. We aim to initialize these bounds such that the mid-point of the bounds (which will be the initial tuning parameter value  $s^{(0)}$ )

since we are performing a bisection search) will approximate the initial tightenings of the quantile-based method. The initial bounds are as follows:

$$s_{\min}^{(0)} = 0, \quad (22a)$$

$$s_{\max}^{(0)} = \left( \max_{i \in \mathcal{N}, \phi \in \Phi} \{ |v_i^{\phi,(\kappa)}| - f_{v,i}^{\phi,(\kappa)}(\epsilon_v) \} \right) \frac{2}{\sigma_{v,i}^{\phi}}. \quad (22b)$$

The lower bound corresponds to the case where all tightenings are zero. The upper bound calculates the maximum difference between the voltage magnitude operating point and  $\epsilon_v$  quantile across all single phase connections. The resulting value is divided by the standard deviation estimate and doubled.

(1) *Initialize*: Set the iteration count to  $\kappa = 0$  and calculate the inverter limit tightenings  $\bar{\Lambda}_q, \underline{\Lambda}_q$  as described in Section [III-A2](#). Initialize the tuning parameter bounds  $s_{\min}^{(0)}, s_{\max}^{(0)}$  according to [\(22\)](#) and calculate the initial value of the tuning parameter by taking the mid-point of the bounds  $s^{(0)} = (s_{\max}^{(0)} - s_{\min}^{(0)})/2 + s_{\min}^{(0)}$ . Calculate  $\bar{\Lambda}_v^{(0)}, \underline{\Lambda}_v^{(0)}$  according to [\(21\)](#).

(2) *Solve approximate problem*: Solve CCR-OPF using fixed tightenings  $\bar{\Lambda}_v^{(\kappa)}, \underline{\Lambda}_v^{(\kappa)}, \bar{\Lambda}_q, \underline{\Lambda}_q$  to obtain candidate operating point  $\mathbf{X}^{(\kappa+1)}$ .

(3) *Solution evaluation*: Perform the empirical violation probability evaluation described in Section [III-B2](#) to calculate the worst case empirical violation probability  $\hat{E}_v^{\max}(\mathbf{X})$  as defined in [\(20\)](#).

(4) *Update tightenings*:  
 If  $\hat{E}_v^{\max}(\mathbf{X}^{(\kappa+1)}) \leq \epsilon_v$ , decrease  $s$  by setting  $s^{(\kappa+1)} = s^{(\kappa)}$  to obtain a less conservative solution.  
 If  $\hat{E}_v^{\max}(\mathbf{X}^{(\kappa+1)}) > \epsilon_v$ , increase  $s$  by setting  $s^{(\kappa+1)} = s^{(\kappa)}$  to obtain a more conservative solution.  
 Obtain a new tuning parameter  $s^{(\kappa+1)} = (s_{\max}^{(\kappa+1)} - s_{\min}^{(\kappa+1)})/2 + s_{\min}^{(\kappa+1)}$  and update the tightenings according to [\(21\)](#).

(5) *Check convergence*: Terminate if the worst case violation probability is within a tolerance level  $\eta \in \mathbb{R}^+$  of the desired  $\epsilon_v$ , i.e.,  $|\hat{E}_v^{\max}(\mathbf{X}^{(\kappa+1)}) - \epsilon_v| \leq \eta$ , or if the upper and lower bounds have converged within tolerance  $\eta_s \in \mathbb{R}^+$ , i.e.,  $|s_{\max}^{(\kappa)} - s_{\min}^{(\kappa)}| \leq \eta_s$ . Return the solution with the lowest objective that satisfies  $|\hat{E}_v^{\max}(\mathbf{X}^{(\kappa+1)}) - \epsilon_v| \leq 0$ . Otherwise, increase iteration count  $\kappa = \kappa + 1$ , and go back to step (2).

3) *Comparison of iterative schemes*: It is evident that both quantile-based and tuning-based methods share large similarities, mainly differing in how the uncertainty margins are updated in each iterations. The most significant difference between the two methods is that the quantile-based method has more flexibility in its tightening updates due to the ability to calculate different tightenings for each of the voltage constraints separately. In contrast, the tuning method relies on adjusting a single tuning parameter  $s$ , which is used across all voltage constraints. However, the drawback of the flexibility of the quantile-based method is that there is a higher potential for overfitting to the sampled uncertainty data.

The updating of the tightenings of the two methods are

loosely related. For the tuning-based method, the tightenings are defined as the product of  $s$  and an estimate of the standard deviation. We can interpret tuning  $s$  as choosing the number of standard deviations away from the average value to set the voltage limits. If the distributions  $f_{v,i}^\phi$  were known, then we could choose the number of standard deviations to directly correspond to the  $\epsilon_v$  or  $1 - \epsilon_v$  quantiles of  $f_{v,i}^\phi$ . In other words, tuning essentially functions as an indirect method for choosing a quantile.

#### IV. CASE STUDY SET-UP

In the case studies, we perform extensive numerical simulations on the IEEE 13-node radial distribution feeder [46] using realistic uncertainty data for solar PV generation and load demand to assess the performance of the iterative algorithms detailed in Section III. Our goal is to find reactive power set-points for the inverters that are valid for the whole day, i.e., we assume that the distribution of possible load and solar PV values is representative for power injection profiles across the whole day.

We investigate how properties of the uncertainty data set (i.e., sampling procedure, number of samples, correlation between data points) can impact the resulting solutions. We further compare the performance of the two algorithms in terms of chance constraint feasibility and objective optimality using both in- and out-of-sample evaluations. In this section, we describe the test case and data sets used to obtain the results in the subsequent sections.

The optimization problem and algorithms are implemented in Julia [47] using JuMP [48] with the solver Ipopt [49].

##### A. Feeder Description

As our test system, we use the modified IEEE 13-node feeder [46], which consists of 15 houses represented as single-phase connections at seven nodes as shown in Fig. 5. To ensure that the generation and demand data is realistic, we use Pecan Street [43] residential data with 1-minute resolution. This data comprises of measurements from residential homes in New York, which are available for 25 full days between May 2019 and August 2019. The measurements include rooftop solar PV generation and load profiles modelling electric vehicle charging behavior, HVAC, refrigerator, and other appliance use. Recall that an example of the load and solar PV profiles for a single day was shown in Fig. 1. The Pecan Street data was scaled by a factor of 20 to match the existing loads in IEEE 13-node feeder. We assume each house has a single-phase rooftop solar PV system with a maximum inverter rating of 100kVA.

##### B. Sampling Procedure

For our numerical simulations, we split the available data for in-sample and out-of-sample evaluations.

1) *Out-of-sample data*: We first randomly choose five days of data from the 25 day data set to be used for an out-of-sample evaluation. These samples are the same across all experiments.

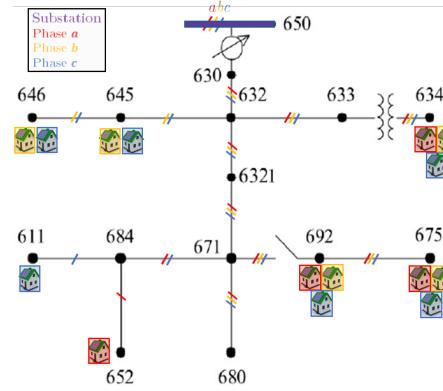


Fig. 5. Modified IEEE-13 node feeder.

2) *In-sample data*: We draw  $M$  samples from the remaining 20 days to be used in the iterative algorithms and in-sample evaluation. These  $M$  samples are drawn using two types of sampling methods with three size variations:

- (i) *Full day samples*: We randomly choose either 1, 2 or 4 days from the set of 20 days, corresponding to  $M = 1440, 2880$ , or  $5760$  data points, respectively. This data set corresponds to the common practice of using data from a set of “representative days”. In this case, there may be strong correlations among subsequent time steps, but all time steps will be represented in the data used in the algorithms.
- (ii) *Random samples*: All data points from the set of 20 days are pooled together and  $M = 1440, 2880$ , or  $5760$  samples are randomly drawn from the data pool. This data set assumes that the data from the 20 days represents a probability distribution of the data and draws i.i.d. samples from this set. This does not guarantee that all time steps in the day are represented, but is likely to represent a wider variety of operating conditions as it includes data from more days.

##### C. Investigations

In the following two sections, we perform a variety of analyses to examine the performance of our methods and assess the different ways of modeling hard inverter limits:

- Section V investigates how **different data sampling procedures** impact the performance of our methods. Specifically, we analyze the performance of the two iterative algorithms with the full day and random samples across several different algorithm replications, and aim to assess which sampling procedure is most suitable.
- After determining the most suitable sampling approach, Section VI provides a **detailed comparison between the two iterative algorithms**. We look at the resulting constraint tightenings, inverter and voltage set points, and constraint violations evaluated on in- and out-of-sample data.
- Section VII investigates algorithm performance when the effect of **inverter capping** is incorporated.

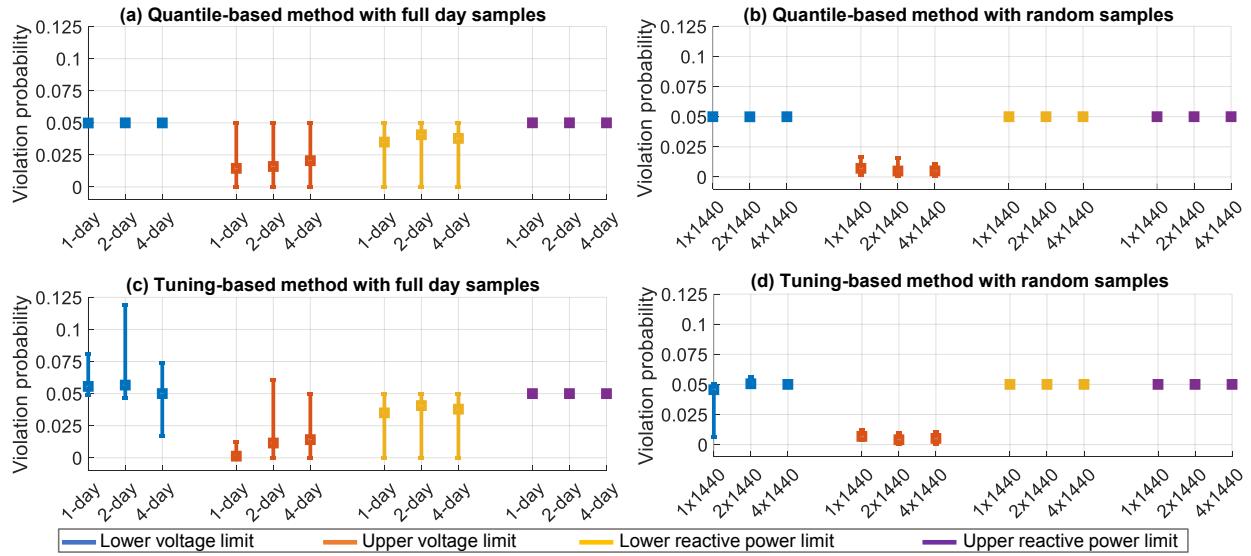


Fig. 6. In-sample evaluation of violation probabilities for quantile-based and tuning-based methods with 1, 2, 4 full day samples and 1440, 2880, 5760 random samples for 10 replications. The worst-case empirical violation probability for all constraint types are shown. The square represent the average values over 10 replications and the bars show the variability of violation probabilities (minimum and maximum values) across 10 replications.

## V. CASE STUDY I: COMPARISON OF SAMPLING METHODS

We first run numerical experiments to assess the performance of both the quantile and tuning-based methods under each of the data set sampling variations detailed in Section IV-B2. For each of the six sampling methods (full day or random samples with  $M = 1440, 2880$ , or  $5760$ ), we perform 10 algorithm replications, each using a different, independent sample draw. We set the desired chance constraint violation probability to be  $\epsilon_v = \epsilon_q = 0.05$ .

In Figs. 6 and 7, we plot the worst-case empirical violation probabilities for each constraint type (19) evaluated on in- and out-of-sample data, respectively. The worst-case violation probabilities for the lower voltage magnitude, upper voltage magnitude, lower inverter reactive power, and upper inverter reactive power constraints are illustrated in the figures using the blue, red, yellow, and purple bars, respectively. The average across the 10 replications is shown using the square, while the bars represent the range of worst-case violation probabilities obtained across replications. The figures illustrate the results for solutions obtained by the quantile-based (top) and tuning-based (bottom) methods, when using full day samples consisting of 1, 2, and 4 days (left) and  $M = 1 \times 1440, 2 \times 1440$ , and  $4 \times 1440$  randomly sampled data points (right).

We assess solution optimality by calculating the total VUF obtained the CCR-OPF solutions. Fig. 8 shows the average, minimum, and maximum values across 10 replications of the in-sample VUF for solutions obtained by each algorithm and data set variation. We also evaluate the solution quality on the out-of sample data by calculating the difference between the out-of-sample VUF and in-sample VUF normalized by the in-sample VUF. The out-of-sample VUF for each replication is obtained by calculating the total unbalance averaged across the five evaluation days. Fig. 9 illustrates the average, minimum, and maximum values across 10 replications of the normalized

VUF for each algorithm and data set variation, evaluated on out-of-sample data.

### A. Comparison of In-Sample Results

We compare the in-sample results of both methods using the various data sets.

1) *Empirical violation probability*: To assess the feasibility of solutions to the original chance-constrained problem, we compare the in-sample violation probabilities of solutions obtained from the quantile-based method (top) and the tuning-based method (bottom) in Fig 6. We observe that the quantile-based method limits the worst-case violation probabilities of all constraint types to the desired  $\epsilon_v = \epsilon_q = 0.05$  for all cases. The tuning-based method is unable to consistently obtain a worst case violation probability of 0.05 using the full day data, particularly for the upper voltage constraints, but mostly achieves violation probabilities close to 0.05 when using randomly sampled data. This may be a consequence of the tuning-based method having less flexibility than the quantile-based method in adapting the tightenings. While the quantile-based methods adjusts all tightenings individually based on the quantile evaluation, the tuning-based method uses the same tuning parameter  $s$  across all constraints, including both upper and lower voltage magnitude constraints. As a result, the quantile-based method provides more “fine-tuned” and less conservative tightenings.

2) *Sampling procedure*: In Fig. 6, we observe that, for both algorithms, the range of worst-case probabilities is much smaller when using random samples (right) than when using full day samples (left). This likely occurs since drawing random samples results in a data set that captures a wider range of possible uncertainty realizations.

3) *VUF*: Due to the less conservative tightening approach of the quantile-based method, we observe in Fig. 8 that the

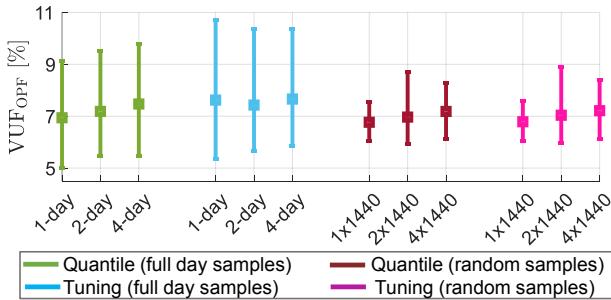


Fig. 8. In-sample evaluation of total VUF for quantile-based and tuning-based methods with 1, 2, 4 full day samples and 1440, 2880, 5760 random samples for 10 replications. The square represent the average values over 10 replications and the bars show the variability of VUF (minimum and maximum values) across 10 replications.

average VUF values for the quantile-based method (green and maroon bars) are slightly lower than the VUF obtained by the tuning-based method (light blue and magenta bars) which solves a more constrained problem resulting in higher unbalance levels. Furthermore, by comparing the results obtained with the full day samples (green and light blue bars) and the results obtained using random samples (maroon and magenta bars), we see that the VUF tends to be slightly smaller on average and fall in a narrower range. This is true across all sample sizes (including the smallest sample size set). This indicates that the total set of operating conditions is much better represented by using randomly chosen samples rather than using full days of data.

### B. Comparison of Out-of-Sample Results

We next consider the out-of-sample performance of each method.

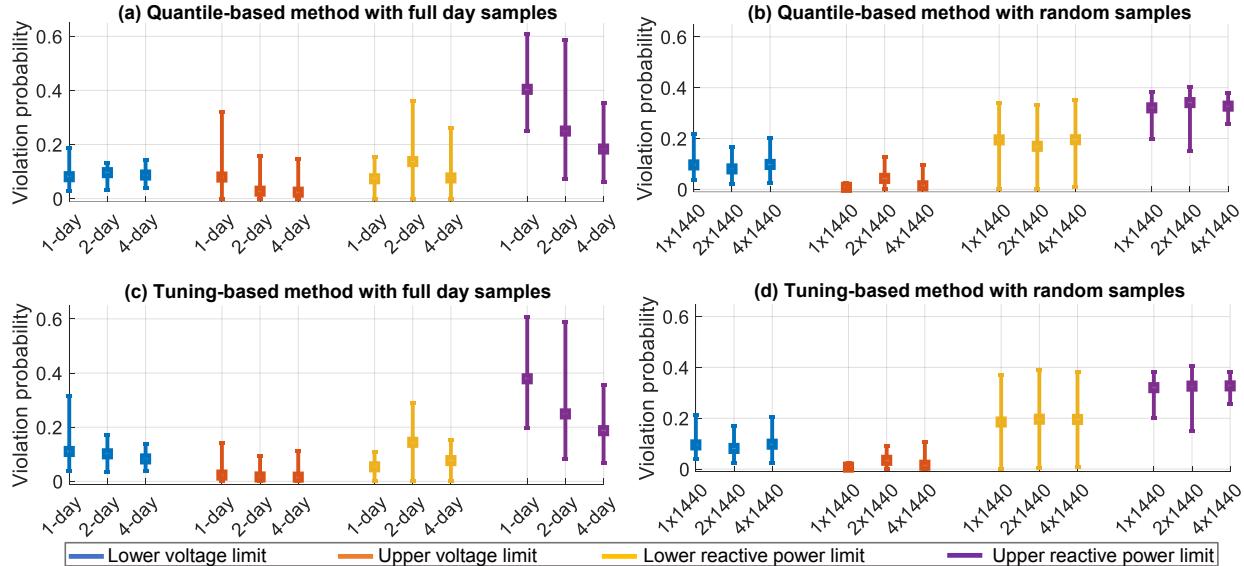


Fig. 7. Out-of-sample evaluation of violation probabilities for quantile-based and tuning-based methods with 1, 2, 4 full day samples and 1440, 2880, 5760 random samples for 10 replications. The worst-case empirical violation probability for all constraint types are shown. The square represent the average values over 10 replications and the bars show the variability of violation probabilities (minimum and maximum values) across 10 replications.

1) *Empirical violation probability*: In Fig. 7 we first observe that the maximum violation probability across the 10 replications is significantly higher than the desired violation probabilities  $\epsilon_v = \epsilon_q = 0.05$ , with empirical violation probabilities up to 0.6. This indicates that solutions obtained from either method are likely not feasible to the original chance constrained problem. When comparing the quantile-based results (top) and the tuning-based results (bottom), we observe that both method generally provide comparable out-of-sample violation probabilities.

2) *Sampling procedure*: The significant difference in the in- and out-of-sample violation probabilities additionally suggests that the data used within the optimization algorithm (i.e., to produce the in-sample results) is not representative of the out-of-sample data, contributing to the highly disappointing results. We do, however, observe that the solutions obtained using randomly selected samples (right) tend to have lower worst-case violation probabilities, indicating that randomly chosen samples may be an advantageous sampling procedure for both iterative algorithms. We further observe that the use of random samples results in solutions with smaller ranges (i.e., less variation across replications) of violation probabilities compared to full-day samples.

3) *VUF*: From Fig. 9 we observe that there is negligible difference between the results obtained by the quantile-based method (green and maroon bars) and results for the tuning-based method (light blue and magenta bars). Furthermore, the deviation in out-of-sample VUF from the in-sample VUF is higher when we use random samples (maroon and magenta bars) compared to full day samples (green and light blue bars). This is likely because the full day in-sample data is more representative of the out-of-sample data, where we also use full day samples.

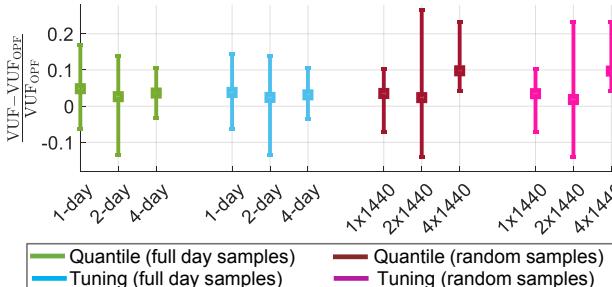


Fig. 9. Out-of-sample evaluation of normalized VUF for quantile-based and tuning-based methods with 1, 2, 4 full day samples and 1440, 2880, 5760 random samples for 10 replications. The square represent the average values over 10 replications and the bars show the variability of VUF (minimum and maximum values) across 10 replications.

Overall, we conclude that both methods perform in a similar manner, with the quantile-based method obtaining slightly better in-sample results.

## VI. CASE STUDY II: COMPARISON OF ITERATIVE ALGORITHMS

The subsequent sections discuss results of a single replication of each of the iterative algorithms. Based on the conclusions of Section V-B, we choose to use a data set comprising of  $M = 2880$  randomly sampled data points.

1) *Constraint tightenings*: We first compare the constraint tightenings obtained with the two different methods. Shown in Fig. 10 are the voltage magnitude set-points and constraint tightenings across all single-phase connections (left) and PV inverter reactive power set-points and constraint tightenings across all houses (right). The nominal constraints are plotted in a dashed red line, while the tightened constraints are shown as blue lines (quantile-based method) and orange lines (tuning-based method). The tightening is the difference between the red dashed and the respective solid lines. We first observe that the inverter reactive power constraints for both methods shown in Fig. 10(b) are identical. This is as expected because they are calculated in a similar manner as described in Section III-A2.

We consider the voltage magnitude constraints shown in Fig. 10(a). We first observe that the tightenings for the upper and lower voltage constraints obtained with the tuning-based method (orange lines) are symmetric. This is as expected, because the voltage tightening of the tuning-based method is given by the product between the standard deviation of the voltage magnitude at each node (which is the same for the upper and lower bound) and the tuning parameter  $s$  (which is shared among all voltage constraints) and is thus the same for both the upper and lower bound. We further notice that the tuning-based method is able to find the tuning parameter  $s$  that results in a tightening of the lower voltage bound (orange line) that closely matches the tightening determined by the quantile-based method (blue line). This empirically demonstrates that tuning  $s$  to be the correct number of standard deviations closely approximates the  $\epsilon_v$  quantile for this constraints. On the other hand, the tightening of the upper voltage limits do not match as closely, i.e., there is a larger difference between

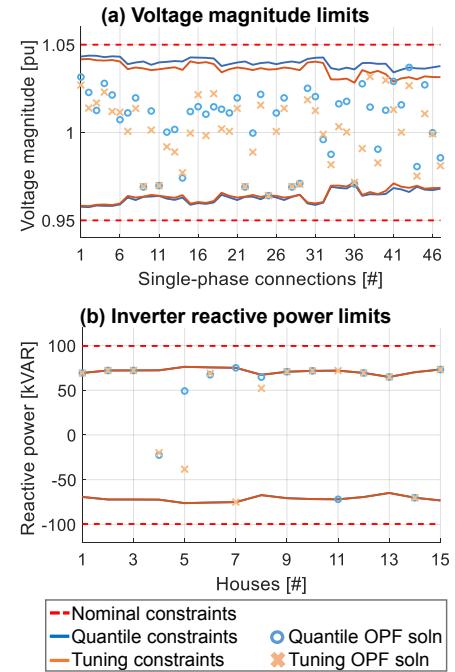


Fig. 10. Comparison of CCR-OPF solution and constraint tightenings resulting from the quantile-based and tuning-based methods for single replication using  $M = 2880$  random samples.

the tightenings obtained with the tuning-based method (orange line) and the quantile-based method (blue line). This indicates that the probability distribution of the voltage magnitudes is not symmetric. This non-symmetry cannot be captured by the tuning-based method as explained above.

2) *OPF solutions*: As a result of the variations in the constraint tightenings, we obtain slightly different solutions with the two methods. The VUF of the nominal solution is 6.3% for the quantile-based method and 6.5% for the tuning-based method. This is as expected, since the quantile-based methods has lower constraint tightenings and thus a slightly larger feasible space, allowing for better solutions.

The nominal voltage magnitude and reactive power obtained with each method are shown in Fig. 10, with blue circles representing the quantile-based solutions and the yellow crosses representing the tuning-based solutions. When comparing the reactive power set-points obtained with the two methods in Fig. 10(b), we see that many of the set-points are either the maximum or minimum allowable reactive power injections and that the two methods mostly produce similar reactive power set-points. However, there are also notable differences. For example, the inverters at houses 7 and 11 have high reactive power injection, but with opposite signs. These differences in the reactive power set-points leads to differences in the voltage magnitude values, shown in Fig. 10(a) with the quantile-based method producing generally higher voltage magnitudes than the tuning-based method. In particular, we observe that the quantile-based method produces a nominal voltage magnitude at node 43 that violates the tightened constraint of the tuning-based method. This shows how smaller constraint tightenings

can enable solutions with a lower nominal VUF value.

3) *Constraint violations by node and inverter*: We next compare the performance of the methods in terms of constraint violations for the same replication discussed above. To assess the differences in performance, we compute the voltage magnitude and inverter apparent power for each in-sample and out-of-sample realization. Fig. 11 shows the box-whisker plots of the results for the in-sample (top) and out-of-sample (bottom) data using the quantile-based solution. We omit the results obtained for the tuning-based method, as they are qualitatively similar to the ones seen in Fig. 11.

We first compare the in-sample and out-of-sample results for the voltage magnitudes (left). The voltage magnitude violations in the in-sample evaluation (shown in Fig. 11(a)) are both larger in magnitude and occur at a greater number of nodes compared to the out-of-sample evaluation (shown in Fig. 11(c)). A similar trend is observed when we compare the in-sample and out-of-sample constraint violations of the inverters. While the largest constraint violations are of similar magnitudes in both the in-sample and out-of-sample data, the number of nodes experiencing violations is much larger in the in-sample evaluation. Moreover, we see that the variability in the data for both voltage magnitudes and apparent power is larger in the in-sample compared to the out-of-sample distributions. This further suggests that the in-sample data is not representative of the out-of-sample distributions and may be due to the difference in the sampling method for the in-sample evaluation, where random samples are used, and the out-of-sample evaluation, where 5 full days of data is used. Furthermore, this indicates the necessity of ensuring the days used for the out-of-sample evaluation are representative of

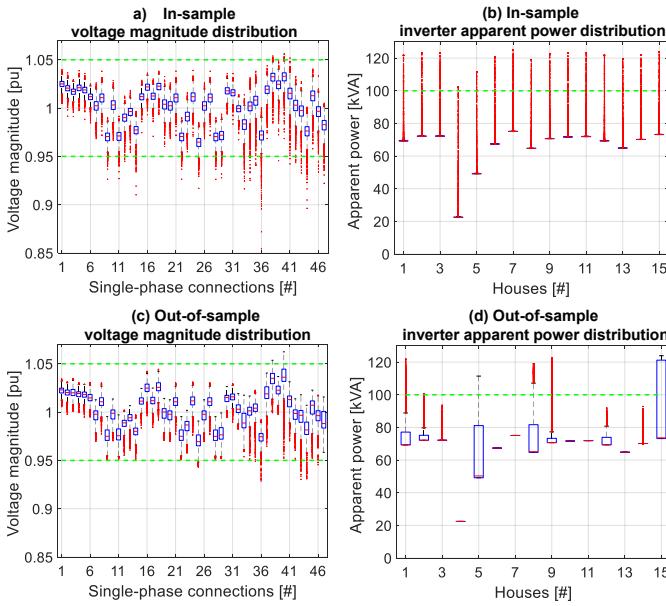


Fig. 11. Box-whisker plots for voltage magnitude and inverter apparent power distribution calculated using Monte Carlo simulations for the quantile-based method for single replication using  $M = 2880$  random samples. The green dashed lines represent the lower and upper voltage magnitude limits and the maximum inverter apparent power limits.

actual operating conditions.

4) *Constraint violations by time of day*: We observed above that there are considerable voltage magnitude and inverter limit violations in the out-of-sample results. Since the out-of-sample evaluation is performed using full day data, we can analyze how the violations vary across the day. Fig. 12 shows violations over time for the quantile-based solution. The top plots show the percentage of single-phase nodes experiencing voltage violations (left) or inverters experiencing apparent power violations (right) at each time step of the day, averaged over all five of the out-of-sample evaluation days. The bottom plots show the magnitude of the worst case violation across all nodes or inverters and all evaluation days at each time step. In Fig. 13 we plot the analogous results for the solution obtained with the tuning-based method.

In Fig. 12(a) and 13(a), the upper voltage violations (orange lines) typically occur around mid-day when the PV generation variability is high, whereas the lower voltage limit violations (blue lines) occur in the evening (between 6pm and 9pm) and at night (between midnight and 2am) when the load variability is high. However, we can see in Figures 12(c) and 13(c) that the magnitude of either voltage magnitude violations are relatively small, with a maximum of around 0.02 p.u.

With respect to the inverter limits shown in Fig. 12(b) and 13(b), both methods have inverter upper limit violations (purple lines) occurring across the entirety of the day-time when the PV generation is non-zero. In contrast to the voltage violation results, the percentage of nodes experiencing violations is relatively high, peaking at around 30% around 1pm for the quantile-based method and 24% for the tuning-based method. A noticeable difference between the two methods is that tuning-based method yields a set point with inverter lower limit violations throughout the day-time, whereas no such violations occur with quantile-based method results. The maximum magnitude of inverter upper limit violation is high for both methods in Fig. 12(d) and 13(d), at 73.2 kVAR throughout the almost entirety of the day-time period. The maximum magnitude of inverter lower limit violation for the tuning-based method is also high at 38.3 kVAR.

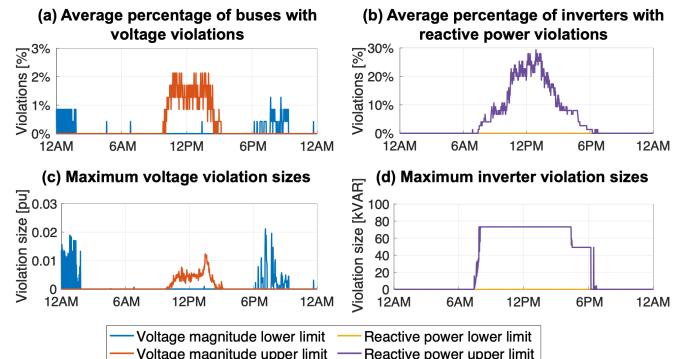


Fig. 12. Time-series plots for out-of-sample voltage magnitude and inverter reactive power limits violations for quantile-based method for single replication using  $M = 2880$  random samples. The top plots show average number of violations throughout the entire day. The bottom plots illustrate the size of maximum violation throughout the entire day.

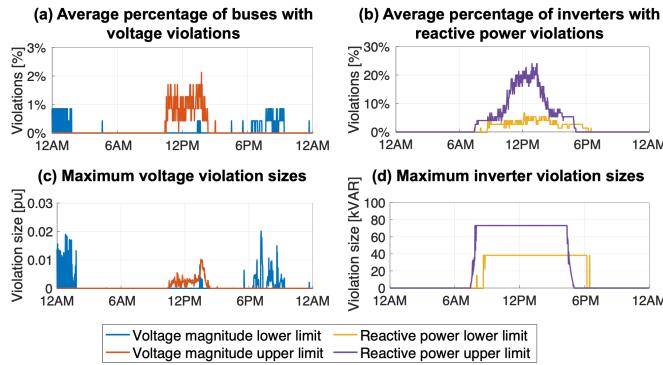


Fig. 13. Time-series plots for out-of-sample voltage magnitude and inverter reactive power limits violations for tuning-based method for single replication using  $M = 2880$  random samples. The top plots show average number of violations throughout the entire day. The bottom plots illustrate the size of maximum violation throughout the entire day.

## VII. CASE STUDY III: CONSIDERING THE IMPACT OF INVERTER CAPPING

A strategy to mitigate inverter violations is capping the inverter reactive power injections to comply with the IEEE 1547-2018 standard [32]. In this section, we cap the inverter reactive power injections to their corresponding limits during the solution evaluation step described in Section III-B. We use the same set of  $M = 2880$  randomly sampled data points as in Section VI and investigate the impact of inverter capping on the voltage violations and total VUF. For all results in this section, we show results only for the quantile-based method since the results of the tuning-based method are very similar.

1) *Comparison of empirical distributions:* We first examine the impact of capping the inverter reactive power injections on the voltage and inverter violations. Shown in Fig. 14 are the voltage magnitude (left) and inverter apparent power (right) distributions for both in-sample (top) and out-of-sample (bottom) data. As expected, we do not observe any inverter limit violations in Figures 14(b) and 14(d) due to the capping. By comparing the voltage distribution plots in Fig. 14(c) to the results obtained in Figs. 11(c) without capping, we see that there are less violations of the upper voltage magnitude limit (at nodes 38 and 40) and smaller violation sizes for the lower voltage magnitude limit (nodes 36 and 44) for the out-of-sample data. The decrease in the number of overvoltages and undervoltage violation size is likely due to the reduced availability of reactive power injections because of capping.

Recall that while the reactive power injections are capped in the solution evaluation step, the reactive power limits in CCR-OPF are tightened using the  $\epsilon_q$  and  $1 - \epsilon_q$  quantiles of the empirical distribution. So, we next investigate how changing  $\epsilon_q$  impacts the constraint tightenings, resulting VUF, and out-of-sample voltage magnitude violations.

2) *Impact of higher violation probability:* We assess the impact of using a higher desired violation for the reactive power constraints. Here, we keep  $\epsilon_v = 0.05$  while using  $\epsilon_q = 0.15$  in the quantile-based algorithm with reactive power capping. Fig. 15 compares the constraint tightenings and operating points for  $\epsilon_q = 0.05$  (blue lines) and  $\epsilon_q = 0.15$  (orange lines).

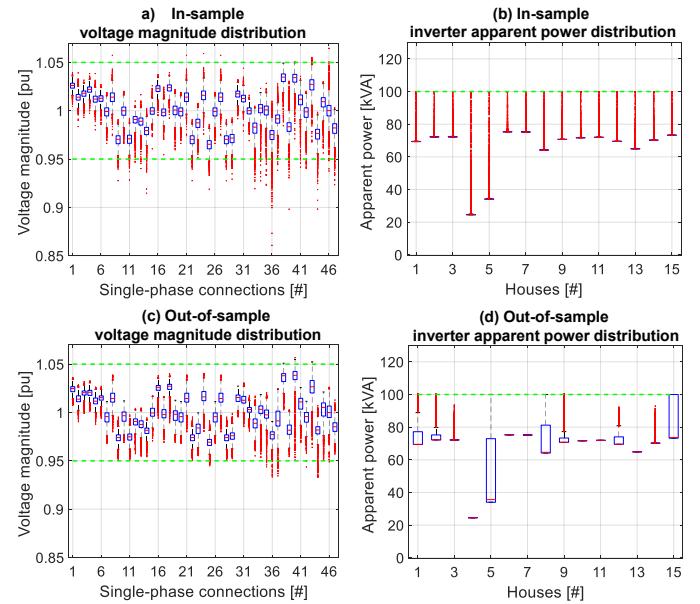


Fig. 14. Box-whisker plots for voltage magnitude and inverter apparent power distribution calculated using Monte Carlo simulations with capping for the quantile-based method for single replication using  $M = 2880$  random samples with inverter capping. The green dashed lines represent the lower and upper voltage magnitude limits and the maximum inverter apparent power limits.

(orange lines). The voltage magnitude tightenings in Fig. 15(a) remain almost identical since we use  $\epsilon_v = 0.05$  for both cases. However, by choosing a higher violation probability  $\epsilon_q = 0.15$ , we obtain a wider range for the reactive power limits in Fig. 15(b), which are very close to the nominal constraints (red dashed lines).

Tables I and II summarize the in- and out-of-sample results for a single replication of the quantile-based method using  $\epsilon_q = 0.05$  without capping (from Section VI),  $\epsilon_q = 0.05$  with

TABLE I  
IN-SAMPLE RESULTS FOR QUANTILE-BASED METHOD WITH CAPPING FOR SINGLE REPLICATION WITH 2880 RANDOM SAMPLES.

Case	$\hat{E}_v^{\max}$	$\hat{E}_{\bar{v}}^{\max}$	$\hat{E}_q^{\max}$	$\hat{E}_{\bar{q}}^{\max}$	VUF (%)
No capping	0.05	0.005	0.05	0.05	6.32
Capping, $\epsilon_q=0.05$	0.05	0.007	0.0	0.0	6.61
Capping, $\epsilon_q=0.15$	0.05	0.05	0.0	0.0	5.85

TABLE II  
OUT-OF-SAMPLE RESULTS FOR QUANTILE-BASED METHOD WITH CAPPING FOR SINGLE REPLICATION WITH 2880 RANDOM SAMPLES.

Case	$\hat{E}_v^{\max}$	$\hat{E}_{\bar{v}}^{\max}$	$\hat{E}_q^{\max}$	$\hat{E}_{\bar{q}}^{\max}$	VUF (%)
No capping	0.019	0.126	0.0	0.36	6.97
Capping, $\epsilon_q=0.05$	0.019	0.016	0.0	0.0	6.81
Capping, $\epsilon_q=0.15$	0.018	0.029	0.0	0.0	6.21

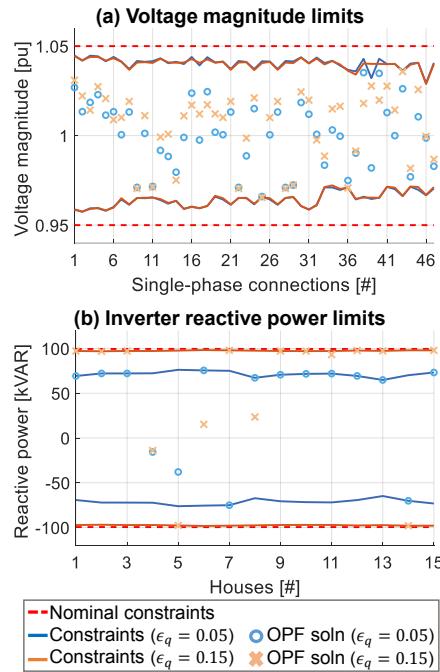


Fig. 15. Comparison of CCR-OPF solution and constraint tightenings resulting from the quantile-based with capping for single replication using  $M = 2880$  random samples with different  $\epsilon_q$ .

capping, and  $\epsilon_q = 0.15$  with capping. All results use the same sample set ( $M = 2880$  randomly drawn samples). We compare the worst-case violation probabilities for all constraint types and the VUF. In Table I we observe that none of the in-sample violation probabilities in any of the three cases exceed 0.05. Furthermore, when inverter capping is used, the violation probabilities for the inverter limits  $\hat{E}_q^{\max}$ ,  $\hat{E}_{\bar{q}}^{\max}$  are zero. By capping with  $\epsilon_q = 0.05$ , there is a slight increase in the VUF compared to the VUF for case without capping. This increase is likely due to the limited reactive power support resulting from capping. By setting  $\epsilon_q = 0.15$ , we increase the feasible space for reactive power injections, leading to a better solution with the lowest VUF value among all three cases.

A similar trend can be observed in the out-of-sample results shown in Table II. For the case without capping, we see that the out-of-sample violation probabilities for the inverter upper limits  $\hat{E}_{\bar{q}}^{\max}$  and voltage upper limits  $\hat{E}_v^{\max}$  are higher than our desired violation probability of 0.05. By capping with  $\epsilon_q = 0.05$ , we are able to achieve violation probabilities that are considerably below 0.05. By setting  $\epsilon_q = 0.15$ , we obtain the lowest VUF value at the cost of increasing the voltage upper limit violations.

### VIII. CONCLUSION

In this paper, we develop iterative, data-driven algorithms for solving the chance-constrained AC OPF for unbalanced distribution grids. We reformulate the chance constraints into deterministic constraints consisting of the nominal constraints tightened with uncertainty margins. The optimal constraint tightenings are calculated using an iterative approach that

alternates between solving a deterministic OPF with fixed tightenings and using sample-based evaluations to update the tightening terms. We propose two methods to perform the iterative updates: directly using the results of a Monte Carlo simulation (quantile-based method) or tuning using a safety parameter (tuning-based method).

Both methods were tested by running numerical simulations on the IEEE 13-bus test feeder using real residential load and PV data. Our case studies demonstrate that both methods perform in a similar manner and are able to enforce the chance constraints in the in-sample evaluation. The out-of-sample results were considerable improved by capping the DER power outputs. Furthermore, simulation results indicate that using randomly chosen samples across multiple days is the advantageous sampling procedure for both methods, as opposed to using representative full day sample sets.

For future work, we will focus on investigating the impact of using linear approximations of the power flow equations in order to solve the chance-constrained problem for large, realistic distribution feeders. We further plan to explore effective ways to identify a multi-dimensional tuning parameter for the tuning-based method so that we can separately tune the individual chance constraints, and will apply results from [42] to ensure that solutions obtained with the tuning-based methods provide rigorous feasibility guarantees.

### REFERENCES

- [1] A. Eggli, S. Karagiannopoulos, S. Bolognani, and G. Hug, "Stability analysis and design of local control schemes in active distribution grids," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 1900–1909, 2020.
- [2] K. E. Antoniadou-Plytaria, I. N. Kouveliotis-Lysikatos, P. S. Georgilakis, and N. D. Hatzigaryiou, "Distributed and decentralized voltage control of smart distribution networks: Models, methods, and future research," *IEEE Transactions on smart grid*, vol. 8, no. 6, pp. 2999–3008, 2017.
- [3] K. Turitsyn, P. Šulc, S. Backhaus, and M. Chertkov, "Distributed control of reactive power flow in a radial distribution circuit with high photovoltaic penetration," in *IEEE PES general meeting*. IEEE, 2010, pp. 1–6.
- [4] M. Yao, I. A. Hiskens, and J. L. Mathieu, "Mitigating voltage unbalance using distributed solar photovoltaic inverters," *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 2642–2651, 2020.
- [5] C. G. Bajo, S. Hashemi, S. B. Kjær, G. Yang, and J. Østergaard, "Voltage unbalance mitigation in LV networks using three-phase PV systems," in *2015 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2015, pp. 2875–2879.
- [6] S. Karagiannopoulos, P. Aristidou, and G. Hug, "A centralised control method for tackling unbalances in active distribution grids," in *Power Systems Computation Conference (PSCC)*, 2018.
- [7] X. Su, M. A. Masoum, and P. J. Wolfs, "Optimal PV inverter reactive power control and real power curtailment to improve performance of unbalanced four-wire LV distribution networks," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 3, pp. 967–977, 2014.
- [8] K. Girigoudar and L. A. Roald, "Linearized three-phase optimal power flow models for distribution grids with voltage unbalance," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 4214–4221.
- [9] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich, and E. Vaechedi, "Distribution system state estimation based on nonsynchronized smart meters," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2919–2928, 2015.
- [10] H. Zhang and P. Li, "Chance constrained programming for optimal power flow under uncertainty," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2417–2424, 2011.
- [11] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, "Analytical reformulation of security constrained optimal power flow with probabilistic constraints," in *2013 IEEE Grenoble Conference*. IEEE, 2013, pp. 1–6.

- [12] T. Summers, J. Warrington, M. Morari, and J. Lygeros, "Stochastic optimal power flow based on conditional value at risk and distributional robustness," *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 116–125, 2015.
- [13] —, "Stochastic optimal power flow based on convex approximations of chance constraints," in *2014 Power Systems Computation Conference*. IEEE, 2014, pp. 1–7.
- [14] J. Liang, D. D. Molina, G. K. Venayagamoorthy, and R. G. Harley, "Two-level dynamic stochastic optimal power flow control for power systems with intermittent renewable generation," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2670–2678, 2013.
- [15] X. Chen, W. Wu, and B. Zhang, "Robust capacity assessment of distributed generation in unbalanced distribution networks incorporating anm techniques," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 651–663, 2018.
- [16] T. Soares, R. J. Bessa, P. Pinson, and H. Moraes, "Active distribution grid management based on robust AC optimal power flow," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6229–6241, 2018.
- [17] R. Mieth and Y. Dvorkin, "Data-driven distributionally robust optimal power flow for distribution systems," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 363–368, 2018.
- [18] V. Kekatos, G. Wang, A. J. Conejo, and G. B. Giannakis, "Stochastic reactive power management in microgrids with renewables," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3386–3395, 2015.
- [19] S. Karagiannopoulos, L. Roald, P. Aristidou, and G. Hug, "Operational planning of active distribution grids under uncertainty."
- [20] M. D. Sankur, R. Dobbe, E. Stewart, D. S. Callaway, and D. B. Arnold, "A linearized power flow model for optimization in unbalanced distribution systems," *arXiv preprint arXiv:1606.04492*, 2016.
- [21] A. Bernstein, C. Wang, E. Dall'Anese, J.-Y. Le Boudec, and C. Zhao, "Load flow in multiphase distribution networks: Existence, uniqueness, non-singularity and linear models," *IEEE Trans. Power Systems*, vol. 33, no. 6, pp. 5832–5843, 2018.
- [22] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE Trans. power systems*, vol. 21, no. 3, pp. 1458–1459, 2006.
- [23] L. Gan, N. Li, U. Topcu, and S. H. Low, "Exact convex relaxation of optimal power flow in radial networks," *IEEE Trans. Automatic Control*, vol. 60, no. 1, pp. 72–87, 2014.
- [24] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2006.
- [25] D. B. Arnold, M. Sankur, R. Dobbe, K. Brady, D. S. Callaway, and A. Von Meier, "Optimal dispatch of reactive power for voltage regulation and balancing in unbalanced distribution systems," in *IEEE PES General Meeting*. IEEE, 2016, pp. 1–5.
- [26] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464–1475, 2013.
- [27] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Sys.*, vol. 27, no. 1, pp. 92–107, 2011.
- [28] C. Zhao, E. Dall'Anese, and S. H. Low, "Convex relaxation of OPF in multiphase radial networks with delta connection," in *Proceedings of the 10th IREP*, 2017, pp. 8885–8950.
- [29] P. Fortenbacher, M. Zellner, and G. Andersson, "Optimal sizing and placement of distributed storage in low voltage networks," in *2016 Power Systems Computation Conference (PSCC)*. IEEE, 2016, pp. 1–7.
- [30] S. Karagiannopoulos, P. Aristidou, and G. Hug, "Data-driven local control design for active distribution grids using off-line optimal power flow and machine learning techniques," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6461–6471, 2019.
- [31] E. Muljadi, R. Schiferl, and T. A. Lipo, "Induction machine phase balancing by unsymmetrical thyristor voltage control," *IEEE Trans. Industry Applications*, no. 3, pp. 669–678, 1985.
- [32] D. G. Photovoltaics and E. Storage, "IEEE standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces," *IEEE Std*, pp. 1547–2018, 2018.
- [33] X. Geng and L. Xie, "Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization," *Annual Reviews in Control*, vol. 47, pp. 341–363, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1367578819300306>
- [34] Y. Cao, Y. Tan, C. Li, and C. Rehtanz, "Chance-constrained optimization-based unbalanced optimal power flow for radial distribution networks," *IEEE Transactions on Power Delivery*, vol. 28, no. 3, pp. 1855–1864, 2013.
- [35] E. Dall'Anese, K. Baker, and T. Summers, "Optimal power flow for distribution systems under uncertain forecasts," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 7502–7507.
- [36] K. Baker, I. E. Dall'Anese, and T. Summers, "Distribution-agnostic stochastic optimal power flow for distribution grids," in *2016 North American Power Symposium (NAPS)*, 2016, pp. 1–6.
- [37] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis, "Convex relaxations of chance constrained AC optimal power flow," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2829–2841, 2018.
- [38] J. Schmidli, L. Roald, S. Chatzivasileiadis, and G. Andersson, "Stochastic AC optimal power flow with approximate chance-constraints," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, 2016, pp. 1–5.
- [39] L. Roald and G. Andersson, "Chance-constrained AC optimal power flow: Reformulations and efficient algorithms," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2906–2918, 2017.
- [40] L. A. Roald, D. K. Molzahn, and A. F. Tobler, "Power system optimization with uncertainty and AC power flow: Analysis of an iterative algorithm," in *10th IREP Symp. Bulk Power Syst. Dynamics Control*, 2017.
- [41] A. M. Hou and L. A. Roald, "Chance constraint tuning for optimal power flow," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. IEEE, 2020, pp. 1–6.
- [42] —, "Data-driven tuning for chance-constrained optimization: Two steps towards probabilistic performance guarantees," *IEEE Control Systems Letters*, vol. 6, pp. 1400–1405, 2022.
- [43] Pecan Street Inc, *Dataport: The world's largest energy data resource*. [Online]. Available: <https://dataport.pecanstreet.org/>
- [44] K. Girigoudar and L. A. Roald, "On the impact of different voltage unbalance metrics in distribution system optimization," *Electric Power Systems Research*, vol. 189, p. 106656, 2020.
- [45] IEC 61000-2-2, *EMC – Part 2-2: Environment – Compatibility Levels for Low Frequency Conducted Disturbances and Signalling in Public Low-Voltage Power Supply Systems*, 2002.
- [46] K. Schneider, B. Mather, B. Pal, C.-W. Ten, G. Shirek, H. Zhu, J. Fuller, J. Pereira, L. Ochoa, L. De Araujo *et al.*, "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3181–3188, 2017.
- [47] J. Bezanson, A. Edelman, S. Karpinski, and V. Shah, "Julia: A fresh approach to numerical computing," *SIAM Review*, vol. 59, no. 1, pp. 65–98, 2017. [Online]. Available: <https://doi.org/10.1137/141000671>
- [48] I. Dunning, J. Huchette, and M. Lubin, "JuMP: A modeling language for mathematical optimization," *SIAM Review*, vol. 59, no. 2, pp. 295–320, 2017.
- [49] A. Wächter and L. T. Biegler, "On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming," *Math. Programming*, vol. 106, no. 1, pp. 25–57, 2006.