Risk-Based Hosting Capacity Analysis in Distribution Systems

Avinash N. Madavan, Nathan Dahlin, Subhonmesh Bose, and Lang Tong

***Abstract*—Solar hosting capacity analysis (HCA) assesses the ability of a distribution network to host distributed solar generation without seriously violating distribution network constraints. In this paper, we consider risk-sensitive HCA that limits the risk of network constraint violations with a collection of scenarios of solar irradiance and nodal power demands, where risk is modeled via the conditional value at risk (CVaR) measure. First, we consider the question of maximizing aggregate installed solar capacities, subject to risk constraints and solve it as a second-order cone program (SOCP) with a standard conic relaxation of the feasible set of the power flow equations. Second, we design an incremental algorithm to decide whether a configuration of solar installations has acceptable risk of constraint violations, modeled via CVaR. The algorithm circumvents explicit risk computation by incrementally constructing inner and outer polyhedral approximations of the set of acceptable solar installation configurations from prior such tests conducted. Our numerical examples study the impact of risk parameters, the number of scenarios and the scalability of our framework.**

***Index Terms*—Hosting capacity analysis, distributed solar, conditional value at risk.**

I. Introduction

ACCORDING to a recent study in [1], usable rooftop space in the United States can accommodate solar energy production sufficient to meet the nation's current electricity demands. Bolstered by federal and state level renewable portfolio standards and legislation, both utility and small-scale distributed installation projects are slated to expand and bring this potential capacity to realization [2]. Given this anticipated rapid growth, we ask: *is the grid ready to host distributed solar at scale*?

In the absence of careful planning, high levels of distributed generation can lead to unacceptable voltage rise and imbalances, protection system malfunction and excessive harmonics [3], necessitating real-time curtailment. For example, solar production curtailments, exceeding 15% of total potential output, have often been necessary in California [4] to avoid violation of system constraints. Today, these interventions typically target utility-scale solar installations. Soon, similar measures may be required in the context of distributed solar installations in low/medium-voltage distribution networks as well, if the impacts of such projects are not properly assessed as they proliferate. Thus, reliable analysis and planning tools are required. The focus of this paper lies in answering *how much installed solar capacity can a distribution network host, without jeopardizing safe operational limits for the network?* Such quantification studies are categorized as *hosting capacity analyses* (HCA). In developing HCA as a planning tool for evolving distribution systems, one must account for the range of operational conditions such systems may face over a planning horizon. By operational conditions, we mean the spatio-temporal characteristics of load and solar irradiance scenarios. Planning for a future system cannot rely solely on historical operating conditions, as future characteristics can substantially differ from the past. Aside from the aforementioned penetration of renewables on the generation side, network load characteristics are expected to change significantly with the increasing adoption of new technologies such as at-home batteries and electric vehicles. To account for such a breadth of operating conditions, stochastic HCA approaches, as surveyed in [5], are more appropriate for planning purposes than deterministic methods that perform analysis based on one or a few historical scenarios, e.g., those studied in [6], [7], [8].

Stochastic HCA assesses acceptibility of solar capacity installations on the basis of system constraint violation *statistics*. Evaluation of such statistics often incurs both modeling and computational challenges. For example, the probability of constraint violation considered in existing approaches, e.g. [9], [10], requires identification of a suitable probability distribution over operating conditions. As illustrated in [11], solar generation is not well characterized by parametric distributions. Even when an accurate parametric distribution is available, outside of a limited collection of distribution families, solution of stochastic optimization problems with parametric families is generally difficult. This difficulty can arise from the nonconvexity introduced by stochastic constraints, e.g., those that seek to restrict the probability of violation. The challenge can also arise due to the need to compute suitable expectations that require computationally intensive numerical integration, as argued in [12].

The computational difficulties associated with exact calculation of statistics can be alleviated by instead relying on a combination of scenarios sampled from generative models and/or

Manuscript received 19 April 2022; revised 9 September 2022 and 5 December 2022; accepted 13 December 2022. Date of publication 7 February 2023; date of current version 26 December 2023. This work was supported by NSF CAREER under Grant 2048065. The work of Lang Tong was supported by NSF under Grant 1932501. Paper no. TPWRS-00564-2022. *(Corresponding author: Avinash N. Madavan.)*

Avinash N. Madavan, Nathan Dahlin, and Subhonmesh Bose are with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: madavan2@illinois.edu; dahlin@illinois.edu; boses@illinois.edu).

Lang Tong is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14680 USA (e-mail: lt35@cornell.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPWRS.2023.3238846.

Digital Object Identifier 10.1109/TPWRS.2023.3238846

historical datasets. However, in order to obtain reliable HCA results, a large number of samples need to be used. Therefore, under scenario-based stochastic HCA, scalability of the proposed tools becomes a critical consideration.

The choice of statistic impacts the nature and utility of results obtained via stochastic HCA. While existing approaches control for the probability of violations under candidate installations, they do not necessarily capture the *extent* or severity of said violations. Clearly, both the frequency and extent of violations are relevant to the health of distribution network components, and should be considered in the development of stochastic HCA methods.

In view of these central aspects of stochastic HCA, the key innovation of our work lies in the use of the *conditional value at risk* (CVaR) measure to process the statistics of constraint violation, as we explain in Section III. Already popular in finance (e.g., see [13], [14]), the favorable mathematical properties of CVaR are popularizing its use in engineering domains, including power systems. For example, CVaR is used to encode risk aversion in objectives of multi-stage renewable investment planning and economic dispatch problems in [15] and [16], respectively. As will become evident, these mathematical properties allow efficient algorithm design to tackle the scalability challenge of HCA.

A variety of methods have been proposed for estimation of distribution network hosting capacity. Under an EPRI approach widely recognized and adopted by many utilities, Monte Carlo simulation is first used to generate installation configurations consistent with PV adoption patterns and increasing levels of renewable penetration. Hosting capacity is then assessed by considering constraint violations under min and max midday loading levels [17]. This method was examined in detail across a number of distribution feeder case studies in [18]. Similarly, [19] randomly generates installation configurations for each of a number of penetration levels and assesses hosting capacity via evaluation of the probability of voltage constraint violations at each level under historical load time series. In [20], hosting capacity is determined via the outcome of a mixed integer linear program which jointly optimizes network investment and emission costs together with hosting capacity. Acceptable generation configurations satisfy network constraints, e.g. voltage level across each of a set of renewable generation and loading level scenarios. Due to the combinatorial nature of the problem, heuristics are introduced to reduce the search space size.

The specific contributions of this paper are as follows. First in Section IV, we propose a CVaR-sensitive hosting capacity maximization problem that seeks to identify the maximum aggregate capacity of solar installations in a distribution network, and describe a scenario-based approach to solve it via *convex optimization*. This convexity follows from our use of CVaR, as well as a standard second-order cone programming-based relaxation of the DistFlow power flow model. The vast majority of risk measures introduce nonconvexity, and require explicit evaluation of the risk of constraint violations across scenarios in order to assess the risk of a particular installation configuration. As a result, one must consider solar installations one at a time, assess the risk level for each, and update the configuration suitably to optimize hosting capacity. In contrast, CVaR preserves the convexity of our power flow model and allows for formulation of an optimization problem that directly optimizes over realizable installations according to risk level, defined through scenarios, and quality as captured by the chosen objective function. Convexity allows us to leverage mature off-the-shelf convex optimization tools and avoids the pitfalls of constraint linearization, mixed integer programming, or the design of complex, parameter-sensitive algorithms to explore the space of capacity configurations that require careful parameter-tuning, e.g., in [18], [19], [20]. Our numerical experiments demonstrate the impact of risk parameters and the number of scenarios on the solution quality and runtime.

Second, we propose a method in Section V to quickly determine whether a candidate solar installation configuration has acceptable risk of network constraint violations, as evaluated via CVaR. Our proposed method is incremental, i.e., we utilize the results of prior tests to expedite determination as to whether a new candidate capacity configuration is acceptable. The properties of the CVaR measure dictate that the resulting set of acceptable configurations is convex. Thus, we construct and refine convex inner and outer approximations to this set as more configurations are tested. These approximations often permit us to quickly certify whether a new candidate configuration is acceptable, without having to explicitly test for it using all scenarios. Our numerical examples illustrate that acceptability certification time decreases rapidly, as the knowledge accumulates over time from prior tests. We end the paper in Section VII with concluding remarks.

II. The Hosting Capacity Analysis Problem

To formulate HCA formally, we begin by presenting the equivalent single-phase power flow model of an n-bus three-phase balanced radial distribution network.[1] Let $\mathcal{V}$ denote the set of all n buses, and $\overline{\mathcal{V}}$ denote the same set, save the feeder bus. Derived from Kirchhoff's laws, the power flow equations identify voltages and real/reactive power flows over distribution lines that can sustain real and reactive power injections given by $/vpinj(\omega) \in \mathbb{R}^{n-1}$ and $q_{\text{inj}}(\omega) \in \mathbb{R}^{n-1}$ across buses in $\overline{\mathcal{V}}$, where ω encodes an injection scenario. To describe these equations, we introduce additional notation. For the $n-1$ (directed) distribution lines, let $\boldsymbol{R}(\omega) \in \mathbb{R}^{n-1}$ and $\boldsymbol{X}(\omega) \in \mathbb{R}^{n-1}$ collect their resistances and reactances. Define $\boldsymbol{P}(\omega) \in \mathbb{R}^{n-1}$ and $\boldsymbol{Q}(\omega) \in \mathbb{R}^{n-1}$ as the sending-end real and reactive power flows in scenario ω over said lines. Also, let $\boldsymbol{L}(\omega) \in \mathbb{R}^{n-1}$ describe the squared current magnitude over these lines. Let $\boldsymbol{W}(\omega) \in \mathbb{R}^{n}$ collect the squared voltage magnitudes in scenario ω across buses in $\mathcal{V}$.

Let $\boldsymbol{B} \in \mathbb{R}^{(n-1)\times n}$ denote the edge-to-node incidence matrix, i.e., $B_{\ell i} = 1$ and $B_{\ell j} = -1$ if nodes i and j denote the sending and receiving ends of edge ℓ, respectively. Let the edges be ordered by receiving end node index. We let $\bar{\boldsymbol{B}} \in \mathbb{R}^{(n-1)\times(n-1)}$ be

[1]While we present risk-sensitive HCA with a balanced three-phase distribution network, we remark that the generalization to unbalanced systems offers no additional conceptual difficulty.

the reduced incidence matrix B with the column corresponding to the feeder bus removed.

With this notation, we write the power flow equations via the "DistFlow" model as

$$p_{\text{inj}}(\omega) = \bar{B}^{\mathsf{T}} P(\omega) + R \odot L(\omega), \tag{1a}$$

$$q_{\text{inj}}(\omega) = \bar{B}^{\mathsf{T}} Q(\omega) + X \odot L(\omega), \tag{1b}$$

$$BW(\omega) = 2[R \odot P(\omega) + X \odot Q(\omega)] - (R^2 + X^2) \odot L(\omega), \tag{1c}$$

$$[B_+ W(\omega)] \odot L(\omega) = [P(\omega)]^2 + [Q(\omega)]^2, \tag{1d}$$

where $\odot$ denotes the Hadamard, or element-wise, product. Here, B_+ is the matrix resulting from taking $\max\{B_{\ell i}, 0\}$ for all ℓ and i, and $B_+ W(\omega)$ is used to capture the sending-end voltage. This form of the power flow equations for radial distribution networks has been derived in [21]. The first two relations encode power balance at each bus, excluding the feeder. The last two equations describe Kirchhoff's voltage law, expressed in terms of squared line currents and real/reactive power flows. For simplicity, we use

$$(P(\omega), Q(\omega), W(\omega), L(\omega)) \in \mathscr{D}(p_{\text{inj}}(\omega), q_{\text{inj}}(\omega)), \tag{2}$$

to describe the set of all $(P(\omega), Q(\omega), W(\omega), L(\omega))$ that satisfy (1) for a specified power injection profile defined by $p_{\text{inj}}(\omega)$ and $q_{\text{inj}}(\omega)$.

An injection scenario, ω specifies the profiles of solar irradiance, real power demand, and reactive power demand across buses in $\overline{\mathcal{V}}$, given by $\alpha(\omega) \in \mathbb{R}^{n-1}$, $p_{\text{D}}(\omega) \in \mathbb{R}^{n-1}$, and $q_{\text{D}}(\omega) \in \mathbb{R}^{n-1}$, respectively. If $\psi \in \mathbb{R}_+^{n-1}$ describes the vector of installed capacities of PV panels across buses in $\overline{\mathcal{V}}$ in terms of surface area, then the real and reactive power generation from solar panels across the same buses in scenario ω is given by

$$p_{\text{G}}(\omega) = \alpha(\omega) \odot \psi, \quad q_{\text{G}}(\omega) = \eta_{\text{G}} \odot \alpha(\omega) \odot \psi, \tag{3}$$

where $\eta_{\text{G}} \in \mathbb{R}_+^{n-1}$ is the ratio of reactive to real power produced by a solar panel. The value of η_{G} can be computed directly from the power factor (pf) according to $\eta_{\text{G}} = (1/\text{pf}^2 - 1)^{1/2}$. Today's inverters typically operate at unit power factors, for which $\eta_{\text{G}} = 0$. The generation profile assumes a direct scaling of the irradiance with the installed capacity–a premise that holds when inverters connecting solar panels track the maximum power point [22]. Given a configuration of installed capacities ψ across the distribution network, a single scenario ω of solar irradiance and power demands identifies the net real and reactive power injections into buses in $\overline{\mathcal{V}}$, respectively, as

$$p_{\text{inj}}(\omega) = p_{\text{G}}(\omega) - p_{\text{D}}(\omega), \; q_{\text{inj}}(\omega) = q_{\text{G}}(\omega) - q_{\text{D}}(\omega). \tag{4}$$

Suppose Ω defines a set of scenarios. Per Fig. 1, the solar installation configuration ψ and Ω then yield a set of nodal power injection vectors,

$$\mathcal{S}_{\text{inj}} := \{(p_{\text{inj}}(\omega), q_{\text{inj}}(\omega)) : \omega \in \Omega\}. \tag{5}$$

Elements of $\mathcal{S}_{\text{inj}}$ identify voltage magnitudes and power flows, per the DistFlow model in (2). As Fig. 1 portrays, the range of operational conditions over Ω produces a statistics of voltage

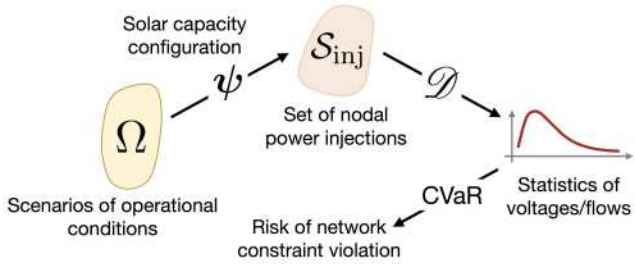

Fig. 1. The work-flow of risk-sensitive solar HCA.

magnitudes and power flows. HCA amounts to processing if the range of voltage magnitudes and power flows "adequately" satisfy engineering constraints of the distribution grid in aggregate.

To define the notion of sufficiency, we say that a scenario ω respects the network constraints, if

$$\underline{W} \le W(\omega) \le \overline{W}, \; P(\omega)^2 + Q(\omega)^2 \le \overline{S}^2, \tag{6}$$

for any $(P(\omega), Q(\omega), W(\omega), L(\omega)) \in \mathscr{D}(p_{\text{inj}}(\omega), q_{\text{inj}}(\omega))$. Here, $\underline{W} \in \mathbb{R}^n$ and $\overline{W} \in \mathbb{R}^n$ are the lower and upper squared voltage limits of each bus in $\mathcal{V}$, and $\overline{S} \in \mathbb{R}^{n-1}$ is the vector of line flow capacities. As the scenario ω ranges over Ω, there are a variety of ways in which one can define when the range of voltage magnitudes and power flows adequately satisfy the network constraints. For example, one can require that *each* scenario in Ω respects the constraints in (6). Such a stringent enforcement of network constraints is generally considered overly conservative for a planning problem such as HCA. Instead, several authors (e.g., [19], [23]) have considered probabilistic constraint enforcement, i.e., they call the network constraints adequately satisfied, if

$$\mathbb{P}[W_i \le \underline{W}_i] \le \nu, \; \mathbb{P}[W_i \ge \overline{W}_i] \le \nu, \mathbb{P}[P_i^2 + Q_i^2 \ge \overline{S}_i^2] \le \gamma$$

for each $i = 1, \ldots, n-1$. Probabilistic constraint enforcement suffers from two major drawbacks. First, it simply restricts the probability of violation, but not the severity of that violation. Large violations can have significant implications, especially in the context of line flow capacity limits, which can typically be temporarily exceeded by small amounts. Second, the sets of feasible solutions that satisfy probabilistic constraints are often non-convex. Non-convexity typically makes it difficult to design efficient algorithms that come with feasibility and optimality guarantees. And, such algorithms often require careful parameter tuning. In the following section, we propose a risk-sensitive approach to HCA that is able to regulate both the probability and the severity of violations, while retaining convexity, allowing us to leverage existing techniques for solving convex optimization problems with provable guarantees.

III. Conditional Value At Risk (CVaR) in Hosting Capacity Analysis

Our key modeling innovation to HCA is the use of the *conditional value at risk* (CVaR) measure to process the statistics of constraint violations across power injection scenarios, as depicted in Fig. 1. We begin by explaining this risk measure

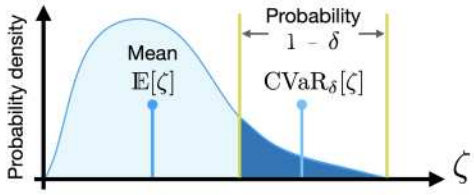

Fig. 2. Example of relationship between the mean and conditional value at risk (CVaR_δ) with parameter δ for a random variable ζ.

for a general scalar random variable ζ with a smooth probability distribution as in Fig. 2. Interpret ζ as a random loss that a decision-maker might face. Parameterized by $\delta \in [0,1)$, CVaR measures the average over the $1-\delta$ fraction of worst-case scenarios. Stated precisely, letting F_ζ be the cumulative distribution function,

$$\mathrm{CVaR}_\delta[\zeta] := \mathbb{E}[\zeta | \zeta \geq F_\zeta^{-1}(\delta)],$$

where $\mathbb{E}[\cdot]$ is the expectation with respect to the probability distribution over ζ. Selection of δ allows expression of a tolerance towards high potential values. For example, setting $\delta = 0$, $\mathrm{CVaR}_0[\zeta] = \mathbb{E}[\zeta]$, while as $\delta \uparrow 1$, $\mathrm{CVaR}_\delta[\zeta]$ approaches the highest value (essential supremum) that ζ can take. For general distributions, CVaR can be written as

$$\mathrm{CVaR}_\delta[\zeta] := \underset{t\in\mathbb{R}}{\text{minimum}} \left\{ t + \frac{1}{1-\delta}\mathbb{E}[[\zeta - t]_+] \right\}, \quad (7)$$

where $[z]_+ = \max\{0, z\}$ for any scalar z. CVaR also bears a close relationship with probabilistic constraint enforcement. In fact one can show, as Fig. 2 suggests,

$$\mathrm{CVaR}_\delta[\zeta] \leq 0 \Longrightarrow \mathbb{P}[\zeta \geq 0] \leq 1-\delta. \quad (8)$$

Thus, CVaR-based constraint enforcement automatically limits the violation probability of network constraints. Furthermore, in computing the expectation over all constraint violations, it regulates the severity of said violations as well.

Adopting the CVaR measure, we now present the method of assessing whether network constraints are sufficiently satisfied over Ω. To that end, consider the violation function of the voltage upper limit, $\boldsymbol{W}(\omega) - \overline{\boldsymbol{W}}$. We impose the constraint

$$\mathrm{CVaR}_\nu[\boldsymbol{W}(\omega) - \overline{\boldsymbol{W}}] \leq 0, \quad (9)$$

where the constraint is interpreted element-wise. Per (8),

$$\mathrm{CVaR}_\nu[\boldsymbol{W}(\omega) - \overline{\boldsymbol{W}}] \leq 0 \Longrightarrow \mathbb{P}\{W_i(\omega) \geq \overline{W}_i\} \leq 1-\delta \quad (10)$$

for $i = 1, \ldots, n$. CVaR is translation-invariant, implying that (9) holds, if and only if

$$\mathrm{CVaR}_\nu[\boldsymbol{W}(\omega)] \leq \overline{\boldsymbol{W}}. \quad (11)$$

Repeating the above exercise for the remaining network constraints in (6), we say that the vector of installed solar capacities, ψ, is *acceptable* if

$$\mathrm{CVaR}_\nu[\boldsymbol{W}(\omega)] \leq \overline{\boldsymbol{W}}, \quad (12a)$$

$$\mathrm{CVaR}_\nu[-\boldsymbol{W}(\omega)] \leq -\underline{\boldsymbol{W}}, \quad (12b)$$

$$\mathrm{CVaR}_\gamma[\boldsymbol{P}^2(\omega) + \boldsymbol{Q}^2(\omega)] \leq \overline{\boldsymbol{S}}^2, \quad (12c)$$

where almost surely over Ω,

$$\boldsymbol{p}_{\mathrm{inj}}(\omega) = \boldsymbol{\alpha}(\omega) \odot \boldsymbol{\psi} - \boldsymbol{p}_{\mathrm{D}}, \quad (13a)$$

$$\boldsymbol{q}_{\mathrm{inj}}(\omega) = \boldsymbol{\eta}_{\mathrm{G}} \odot \boldsymbol{\alpha}(\omega) \odot \boldsymbol{\psi} - \boldsymbol{q}_{\mathrm{D}}, \quad (13b)$$

$$(\boldsymbol{P}(\omega), \boldsymbol{Q}(\omega), \boldsymbol{W}(\omega), \boldsymbol{L}(\omega)) \in \mathscr{D}(\boldsymbol{p}_{\mathrm{inj}}(\omega), \boldsymbol{q}_{\mathrm{inj}}(\omega)). \quad (13c)$$

The set of ψ that satisfy (12) is non-convex, due to the quadratic equality constraint (1d) in (1). It is common practice to *convexify* the power flow model, either via linearization or convex relaxation. For example, linearization is used in [24]. We adopt a convex relaxation of the set of feasible injections, similar to that in [25], replacing (1d) in the definition of $\mathscr{D}$ with

$$[\boldsymbol{B}_+ \boldsymbol{W}(\omega)] \odot \boldsymbol{L}(\omega) \geq [\boldsymbol{P}(\omega)]^2 + [\boldsymbol{Q}(\omega)]^2. \quad (14)$$

Using this, we define the convex DistFlow model as

$$(\boldsymbol{P}(\omega), \boldsymbol{Q}(\omega), \boldsymbol{W}(\omega), \boldsymbol{L}(\omega)) \in \mathscr{D}^{\mathrm{C}}(\boldsymbol{p}_{\mathrm{inj}}(\omega), \boldsymbol{q}_{\mathrm{inj}}(\omega)). \quad (15)$$

In effect, $\mathscr{D}^{\mathrm{C}}$ represents the convex set of all $(\boldsymbol{P}(\omega), \boldsymbol{Q}(\omega), \boldsymbol{W}(\omega), \boldsymbol{L}(\omega))$ that satisfy (1a)–(1c), (14) with power injections $\boldsymbol{p}_{\mathrm{inj}}(\omega)$, $\boldsymbol{q}_{\mathrm{inj}}(\omega)$.

IV. Formulating Acceptable and Maximal Hosting Capacity Problems With Scenarios

We now cast CVaR-sensitive HCA as optimization programs with a collection of samples defining Ω, which characterizes the variety of operational conditions that the distribution network might face. Consider K independent and identically distributed samples $\omega^1, \ldots, \omega^K$ from Ω. Each of these K samples encode solar irradiance, real and reactive power demands across the buses in $\overline{\mathcal{V}}$, denoted by the tuple $(\boldsymbol{\alpha}^k, \boldsymbol{p}_{\mathrm{D}}^k, \boldsymbol{q}_{\mathrm{D}}^k)$ for $k = 1, \ldots, K$. With these samples, the convexified DistFlow model becomes

$$\boldsymbol{p}_{\mathrm{inj}}^k = \boldsymbol{\alpha}^k \odot \boldsymbol{\psi} - \boldsymbol{p}_{\mathrm{D}}^k, \quad (16a)$$

$$\boldsymbol{q}_{\mathrm{inj}}^k = \boldsymbol{\eta}_{\mathrm{G}} \odot \boldsymbol{\alpha}^k \odot \boldsymbol{\psi} - \boldsymbol{q}_{\mathrm{D}}^k, \quad (16b)$$

$$(\boldsymbol{P}^k, \boldsymbol{Q}^k, \boldsymbol{W}^k, \boldsymbol{L}^k) \in \mathscr{D}^C(\boldsymbol{p}_{\mathrm{inj}}^k, \boldsymbol{q}_{\mathrm{inj}}^k), \quad (16c)$$

for each $k = 1, \ldots, K$. The risk-sensitive acceptability test remains (12), now taken uniformly over the samples, e.g., using (7), the scenario-based counterpart of (12a) becomes

$$\underset{\overline{\boldsymbol{w}}}{\text{minimize}} \left\{ \overline{\boldsymbol{w}} + \frac{1}{1-\nu}\frac{1}{K}\sum_{k=1}^{K}[\boldsymbol{W}^k - \overline{\boldsymbol{w}}]^+ \right\} \leq \overline{\boldsymbol{W}}. \quad (17)$$

Imposing that the minimum over $\overline{\boldsymbol{w}}$ must be less than $\overline{\boldsymbol{W}}$ amounts to imposing the existence of a feasible $\overline{\boldsymbol{w}}$ such that the expression being minimized is less than $\overline{\boldsymbol{W}}$, i.e., the above inequality is equivalent to the same with the minimization dropped. This constraint can further be written as a combination of the following constraints,

$$\overline{\boldsymbol{w}} + \frac{1}{K(1-\nu)}\sum_{k=1}^{K}\boldsymbol{t}_w^k \leq \overline{\boldsymbol{W}},\ \boldsymbol{t}_w^k \geq \boldsymbol{W}^k - \overline{\boldsymbol{w}},\ \boldsymbol{t}_w^k \geq 0, \quad (18)$$

for some $\overline{\boldsymbol{w}}, \boldsymbol{t}_w^k, k = 1, \ldots, K$. Thus, the CVaR constraints with samples can be equivalently enforced via the above deterministic

constraints (18). Repeating the same procedure for all constraints in (12) results in the following sampled acceptability test, defined by the feasibility of the following inequalities,

$$\overline{w} + \frac{1}{1-\nu}\frac{1}{K}\sum_{k=1}^{K} \overline{t}_w^k \le \overline{W},\ \overline{t}_w^k \ge W^k - \overline{w}, \tag{19a}$$

$$\underline{w} + \frac{1}{1-\nu}\frac{1}{K}\sum_{k=1}^{K} \underline{t}_w^k \le -\underline{W},\ \underline{t}_w^k \ge -W^k - \underline{w}, \tag{19b}$$

$$s + \frac{1}{1-\gamma}\frac{1}{K}\sum_{k=1}^{K} t_s^k \le \overline{S}^2,\ t_s^k \ge [P^k]^2 + [Q^k]^2 - s, \tag{19c}$$

$$\overline{t}_w^k \ge 0,\ \underline{t}_w^k \ge 0,\ t_s^k \ge 0,\ \text{for } k = 1, \ldots, K, \tag{19d}$$

where (P^k, Q^k, W^k, L^k) is related to $(p_{\text{inj}}^k, q_{\text{inj}}^k)$ via (16) for each $k = 1, \ldots, K$. For a given solar installation configuration ψ, its acceptability can therefore be assessed by solving the optimization problem,

$$\mathsf{ACC}(\psi): \text{maximize } 0, \text{ subject to (16), (19)}, \tag{20}$$

over P^k, Q^k, W^k, L^k, $\overline{w}$, $\underline{w}$, $\overline{t}_w^k$, $\underline{t}_w^k$, t_s^k for $k = 1, \ldots, K$. Furthermore, $\mathsf{ACC}(\psi)$ can be written as a second-order cone program (SOCP),

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{maximize}} \quad & 0, \\ \text{subject to} \quad & C_\psi \psi = D_\psi \boldsymbol{x} + E_\psi, && \text{(21a)} \\ & D\boldsymbol{x} \le E, && \text{(21b)} \\ & \begin{pmatrix} F_i \boldsymbol{x} + G_i \\ f_i^{\mathsf{T}} \boldsymbol{x} + g_i \end{pmatrix} \in \mathcal{K}_{\text{SO}}, \\ & i = 1, \ldots, I, && \text{(21c)} \end{aligned}$$

where $\boldsymbol{x}$ is the vectorized concatenation of all variables in (20). Here, $\mathcal{K}_{\text{SO}}$ is the second-order cone, defined as

$$\mathcal{K}_{\text{SO}} = \{(\boldsymbol{y}^{\mathsf{T}}, v)^{\mathsf{T}} \,|\, \|\boldsymbol{y}\|_2 \le v\}, \tag{22}$$

where $\|\cdot\|_2$ computes the 2-norm of a vector. The equalities (16a) and (16b) are written compactly as (21a). The equality in (1c) can be written as two inequalities. These, together with (12a), (12b) and (19d) become examples of (21b). The nonlinear constraints in (14) and (19c) are written as I second order cone constraints (21c) using the procedure outlined in Appendix A. Formulated as in (21), the acceptability assessment problem is an instance of the widely studied class of convex optimization problems, for which efficient off-the-shelf solvers exist. Beyond testing candidate solar installation configurations, another interesting question is to solve for the *maximal* solar hosting capacity. That is, one seeks the maximal total installed solar capacity over all acceptable configurations. Written mathematically, this amounts to solving

$$\mathsf{OPT}: \text{maximize } \mathbb{1}^{\mathsf{T}}\psi, \text{ subject to (16), (19)}, \tag{23}$$

over ψ, P^k, Q^k, W^k, L^k, $\overline{w}$, $\underline{w}$, $\overline{t}_w^k$, $\underline{t}_w^k$, t_s^k for $k = 1, \ldots, K$. The procedure used to derive (21) can be applied to OPT to derive an SOCP over x, ψ. Again, existing SOCP solvers can then tackle OPT.

The scenario-based CVaR constraints provide a structured framework to solve $\mathsf{ACC}(\psi)$ and OPT as SOCPs for which efficient algorithms come with provable guarantees. While there are parallels between our approach and the BayesOpt framework with probabilistic constraints in [26], there are also important differences. To contrast the two methods, consider the optimality problem with probabilistic constraint enforcement, i.e., replace $\text{CVaR}_\nu[W] \le \overline{W}$ with $\mathbb{P}\{W - \overline{W} > 0\} \le 1 - \nu$. Under a scenario based approach, the probabilistic constraint amounts to imposing

$$\frac{1}{K}\sum_{k=1}^{K} \mathbb{I}_{\{W^k - \overline{W} > 0\}} \le 1 - \nu, \tag{24}$$

where $\mathbb{I}$ is the indicator function that evaluates to unity when its argument is true, and becomes zero otherwise. Tackling such constraints with indicator functions within our framework requires the solution of a nonconvex mixed-integer SOCP, which is typically more computationally challenging than solving the SOCP problem with CVaR-based constraints. The primary difficulty lies in the lack of convexity in the left-hand-side of (24). This makes it challenging to find natural directions to update ψ, as feasibility information cannot be extrapolated to additional points. Thus, each time the candidate installation profile ψ is updated to ψ', the computed $\mathscr{D}^C(p_{\text{inj}}^k, q_{\text{inj}}^k)$ for each scenario ω^k and installation profile ψ must be discarded entirely and recomputed from scratch for installation profile $\psi' \neq \psi$. Essentially, lacking a convex structure for the set of W^k that satisfy (24) together with (16), it becomes difficult to recycle computations across candidate installation profiles. The Bayesian optimization framework in [26] tackles this challenge through a combination of two steps. First, the constraint in (24) is included through a penalty in the objective function to construct an unconstrained optimization problem. Then, the resulting objective is minimized via kernel smoothing of the objective function. Running such algorithms requires careful choice of the penalty factor and the kernel.

Our CVaR-based formulation is inherently convex, and thus obviates the need for such manual tuning techniques. Furthermore, our proposed approach never requires explicit evaluation of CVaR over scenarios for any ψ. Rather, we leverage the variational characterization of CVaR in (7) to formulate the optimality problem over ψ and the other optimization variables in (23), where a solver can exploit the convexity of the problem to iteratively update these variables together, ensuring that the solution guarantees acceptability.

Our formulation utilizes the SOCP-based relaxation of the set of feasible injections, characterized by Kirchhoff's laws. The setup is flexible in that one can utilize any "convexified" power flow model to obtain a convex programming formulation of the CVaR-sensitive hosting capacity problem. This approach offers a sharp contrast to those that rely heavily on the choice of a particular power flow model. For example, the sample reduction technique in [24] crucially depends on a linear power flow model, where constraint violations are processed

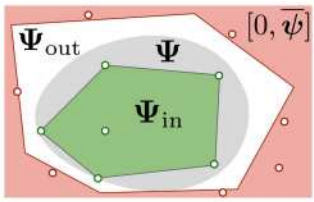

Fig. 3. Inner (green) and outer (white) sets portray Ψ_{in} and Ψ_{out}, respectively, that approximate the set of acceptable installations Ψ (gray). The outer frame encodes the set $[0, \overline{\psi}]$. The sets $\mathbb{A}_\psi$ and $\mathbb{U}_\psi$ comprise the points with green and red borders, respectively.

through a quadratic penalty and optimized, permitting the use of multi-parametric quadratic programming theory to assess the feasibility of an installation configuration ψ. Such a method does not naturally generalize to SOCP-based power flow models.

V. An Incremental Approach to Assess Acceptability

For any given vector of solar hosting capacities ψ, acceptability can be assessed by evaluating the feasibility of (20). However, as the number of samples grow large, this can lead to exceedingly large problem dimension, requiring significant computational time. Instead, it would be beneficial to exploit existing information regarding previously evaluated acceptable and unacceptable hosting capacities to certify the acceptability of ψ. In the sequel, we develop an *incremental algorithm* to certify acceptability. That is, we first attempt to leverage the knowledge of a list of acceptable and unacceptable configurations to certify whether a new candidate configuration is acceptable.

Our design is such that this check is often much faster than solving (20). If unsuccessful, we then solve (20) and use the new configuration to update the knowledge about acceptable and unacceptable configurations. As will be evident, such a method drastically reduces the instances of ψ's for which (20) must be solved to certify acceptability with more tests. Thus, the more we test, the less we require to solve (20). Such a testing paradigm is useful, especially to test acceptability of installation configurations that are close to each other, e.g., those that lie on likely solar adoption paths. Let Ψ denote the set of all acceptable configurations of physically realizable installed solar capacities, i.e., $\Psi \subseteq [0, \overline{\psi}]$. Thanks to the convexity of CVaR, Ψ is a convex set, shaded in gray in Fig. 3. We maintain polyhedral inner and outer approximations Ψ_{in} and Ψ_{out}, respectively, of Ψ satisfying

$$\Psi_{\text{in}} \subseteq \Psi \subseteq \Psi_{\text{out}} \subseteq [0, \overline{\psi}]. \tag{25}$$

These sets are visualized in Fig. 3. We construct Ψ_{in} and Ψ_{out} incrementally from prior acceptability tests. For a new candidate ψ, if $\psi \in \Psi_{\text{in}}$, then ψ is acceptable. On the other hand, if $\psi \notin \Psi_{\text{out}}$, then ψ is unacceptable. If neither of these tests certify the acceptability of ψ, then we run (20). The result of the optimization program either grows Ψ_{in} or shrinks Ψ_{out}, thus expanding the space of ψ's where acceptability can be certified without running (20). In the rest of this section, we outline the incremental construction of Ψ_{in} and Ψ_{out}, and how we test $\psi \in \Psi_{\text{in}}$ or $\psi \notin \Psi_{\text{out}}$. The method for testing acceptability of ψ is summarized in Algorithm 1.

With a list $\mathbb{A}_\psi := \{\psi^a[1], \ldots, \psi^a[M]\}$ of acceptable configurations, we construct Ψ_{in} as their *convex hull*,

$$\Psi_{\text{in}} = \text{conv}(\mathbb{A}_\psi) = \left\{ \sum_{i=1}^{M} \beta_i \psi^a[i] : \boldsymbol{\beta} \in \mathbb{R}_+^M, \mathbb{1}^\mathsf{T}\boldsymbol{\beta} = 1 \right\}. \tag{26}$$

Testing if $\psi \in \Psi_{\text{in}}$ can be solved as a linear program over $\boldsymbol{\beta} \in \mathbb{R}^M$. If ψ is certified acceptable from running (20), then $\mathbb{A}_\psi \leftarrow \mathbb{A}_\psi \cup \psi$ and define again $\Psi_{\text{in}} = \text{conv}(\mathbb{A}_\psi)$. We grow Ψ_{in} incrementally, starting from the null set.

Next, we describe how we construct Ψ_{out}, starting with $[0, \overline{\psi}]$. In the event that (20) finds ψ unacceptable, we construct a *feasibility cut*, borrowing the idea from the generalized Benders' decomposition in [27]. The cut generation technique makes use of the following result. We use $I = 1$ in (21) for notational simplicity. The proof is deferred to Appendix B.

Proposition 1: If there exists $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2)$ that satisfy

$$\boldsymbol{D}^\mathsf{T}\boldsymbol{\mu} = \boldsymbol{D}_\psi^\mathsf{T}\boldsymbol{\lambda} + \boldsymbol{F}^\mathsf{T}\boldsymbol{\mu}_1 + \boldsymbol{f}\mu_2,\ \boldsymbol{\mu} \geq 0,\ (\boldsymbol{\mu}_1^\mathsf{T}, \mu_2)^\mathsf{T} \in \mathcal{K}_{\text{SO}}, \tag{27a}$$

$$(\boldsymbol{C}_\psi \psi - \boldsymbol{E}_\psi)^\mathsf{T}\boldsymbol{\lambda} - \boldsymbol{E}^\mathsf{T}\boldsymbol{\mu} - \boldsymbol{G}^\mathsf{T}\boldsymbol{\mu}_1 - \mu_2 g > 0, \tag{27b}$$

then $\text{ACC}(\psi)$ is not feasible, i.e., ψ is not an acceptable installation configuration.

Existence of a point satisfying the conditions delineated in (27) certifies the unacceptability of ψ. As the proof of Proposition 1 in Appendix B reveals, $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2)$ satisfying (27a) is a feasible point in the dual program of $\text{ACC}(\psi)$. Corresponding to such a point, the *ray* characterized by $(\kappa\boldsymbol{\lambda}, \kappa\boldsymbol{\mu}, \kappa\boldsymbol{\mu}_1, \kappa\mu_2)$ for all $\kappa \geq 0$ is also dual-feasible. When this ray additionally satisfies (27b), it provides a direction along which the objective function of the dual program increases to infinity. Assuming that (21) is strongly infeasible (see [28] for a definition) for all unacceptable ψ, each such ψ must admit such a dual improving ray. Interior point solvers for (20) typically produce a dual improving ray, when infeasible.[2] Consider a list of unacceptable configurations $\mathbb{U}_\psi := \{\psi^u[1], \ldots, \psi^u[M']\}$. For an unacceptable $\psi^u[i] \in \mathbb{U}_\psi$, let the dual improving ray be characterized by $(\boldsymbol{\lambda}[i], \boldsymbol{\mu}[i], \boldsymbol{\mu}_1[i], \mu_2[i])$. Then, using $\mathcal{J}(\boldsymbol{\lambda}[i], \boldsymbol{\mu}[i], \boldsymbol{\mu}_1[i], \mu_2[i]; \psi)$ to denote the left-hand-side of (27b) at that point, define

$$\Psi_{\text{out}} = \{\psi : \mathcal{J}(\boldsymbol{\lambda}[i], \boldsymbol{\mu}[i], \boldsymbol{\mu}_1[i], \mu_2[i]; \psi) \leq 0, \quad i = 1, \ldots, M', \psi \in [0, \overline{\psi}]\}. \tag{28}$$

If ψ violates at least one among the M' affine constraints, then $\psi \notin \Psi_{\text{out}}$. When (20) returns a new unacceptable ψ, then $\mathbb{U}_\psi \leftarrow \mathbb{U}_\psi \cup \psi$, and the cut defined via its dual improving ray is added to the description of Ψ_{out} via (28). Addition of this cut shrinks Ψ_{out}.

Algorithm 1 crucially relies on the convex nature of the set Ψ–a property that follows from properties of the CVaR measure. With a probabilistic constraint enforcement as in [26], this set

[2] See [29] for the mechanics of producing such a ray.

Algorithm 1: Incremental Scenario-Based Acceptability Test.

Input: ψ, $(\mathbb{A}_\psi, \mathbf{\Psi}_{\text{in}}), (\mathbb{U}_\psi, \mathbf{\Psi}_{\text{out}})$.
if $\psi \notin \mathbf{\Psi}_{\text{out}}$ **then return** unacceptable.
else if $\psi \in \mathbf{\Psi}_{\text{in}}$ **then return** acceptable.
else
  Solve (20).
  if infeasible and dual improving ray $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ exists **then**
    $\mathbb{U}_\psi \leftarrow \mathbb{U}_\psi \cup \psi$, update $\mathbf{\Psi}_{\text{out}}$ via (28).
    return unacceptable.
  else
    $\mathbb{A}_\psi \leftarrow \mathbb{A}_\psi \cup \psi$, update $\mathbf{\Psi}_{\text{in}}$ via (26).
    return acceptable.
  end if
end if

TABLE I
SUMMARY OF RESULTS FROM ACCEPTABILITY TESTS OVER THE 3-BUS DISTRIBUTION NETWORK EXAMPLE

Test	Mean Runtime (s)	Frequency
$\psi \in \mathbf{\Psi}_{\text{in}}$ test	9.678×10^{-4}	451
$\psi \notin \mathbf{\Psi}_{\text{out}}$ test	1.195×10^{-6}	489
Solving (20), when $\psi \in \mathbf{\Psi}_{\text{in}}$	46.288	47
Solving (20), when $\psi \notin \mathbf{\Psi}_{\text{out}}$	133.107	13

TABLE II
SUMMARY OF RESULTS FROM ACCEPTABILITY TESTS OVER THE 56-BUS DISTRIBUTION NETWORK EXAMPLE

Test	Mean Runtime (s)	Frequency
$\psi \in \mathbf{\Psi}_{\text{in}}$ test	1.473×10^{-3}	6
$\psi \notin \mathbf{\Psi}_{\text{out}}$ test	1.419×10^{-6}	79
Solving (20), when $\psi \in \mathbf{\Psi}_{\text{in}}$	978.764	49
Solving (20), when $\psi \notin \mathbf{\Psi}_{\text{out}}$	1297.369	16

may *not* be convex, rendering the construction of polyhedral inner/outer approximations of $\mathbf{\Psi}$ challenging.

VI. Numerical Experiments

We now study these acceptability and optimality evaluations for two power network examples: a 3-bus network for illustrative purposes and the Southern California Edison's 56-bus network from [30], studied previously in [26], to demonstrate scalability. The considered 3-bus network is a simple modification of the MATPOWER 4-bus distribution network [31], with the 4th bus removed. Simulations on the 3-bus network were performed on an i5-4690 k machine with 16 GB RAM, while the 56-bus simulations were performed on an AWS m4.10xlarge instance. All results were generated using MOSEK version 9.3. Load data and DER energy profiles were taken at 10 minute intervals over a 365-day period (52.56 K intervals) using data from Southern California Edison. For all experiments, DER installation power factors were set to 0.97 ($\eta_{\text{G}} = 0.251$), and the feeder voltage was fixed at $w_1 = 1$ p.u. Data and code are available on GitHub at https://github.com/amadavan/RSHC.

A. Numerical Experiments for Acceptability Assessment

1) On the Three-Bus Network: We tested acceptability of uniformly randomly generated installation configurations $(\psi_2, \psi_3) \in [0, 4]^2$ under risk parameters $\nu = \gamma = 0.8$. At the beginning of the testing process, $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$ do not contain relevant information to judge acceptability of candidate configurations, and so (20) is solved for these configurations. However, as more configurations are tested, $\mathbf{\Psi}_{\text{in}}$ grows and $\mathbf{\Psi}_{\text{out}}$ shrinks, as Fig. 4 reveals. As a result, more configurations are quickly certified as acceptable/unacceptable using $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$. As Table I shows, the run-time differences between solving (20) and testing via $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$ are substantial.

Only a small fraction (6%) of the 1000 tested configurations, per Table I, required the solution of (20) for acceptability certification. The sets $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$ therefore effectively capture the knowledge of past tests, highlighting the efficacy of our incremental approach in solving $\mathsf{ACC}(\psi)$.

Fig. 4 portrays how the sets $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$ change with tests. As one might expect, the set of acceptable configurations with lower risk parameters ($\nu = \gamma = 0.6$) in the top half of Fig. 4 is larger than that with higher risk parameters ($\nu = \gamma = 0.8$) in the bottom half of Fig. 4. The two sequences of plots were generated with the same set of test configurations, and hence, illustrate how they contribute to the construction of the sets $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$ under different risk parameters.

In practice, installation configurations chosen for certification will likely not be generated uniformly randomly. Rather, these configurations will lie along projected solar adoption paths, starting from an initial acceptable configuration. These capacities being nearby, we expect $\mathbf{\Psi}_{\text{in}}$ to be effective in quickly declaring acceptability for most points. Again, as the adoption path leaves $\mathbf{\Psi}$, a few tests near the boundary should construct cutting planes, so that later points on the trajectory can be quickly ruled out using $\mathbf{\Psi}_{\text{out}}$.

2) On the 56-Bus Distribution Network: We ran our acceptability test with 15 K scenarios over capacity configurations that are uniformly distributed within a square with side length 0.4 centered at $\psi^\star$, obtained with $\nu = 0.9$ and $\gamma = 0.8$. We first ran 32 tests corresponding to the corners of this 5-dimensional hypercube. The corner configurations of the hypercube provide valuable information to the construction of $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$. After these 32 corners, we considered an additional 118 data points sampled uniformly over the aforementioned hypercube. Table II records the results of our numerical experiments. Checking acceptability via $\mathbf{\Psi}_{\text{in}}$ or $\mathbf{\Psi}_{\text{out}}$ takes less than a second, while running (20) takes substantially longer than for the 3-bus network example, given the increased network size. Even with a small number (118) of tests, the majority of the points are certified to be acceptable or unacceptable using $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$, without requiring explicit solution of (20). The run-time differences and the frequency of certification using $\mathbf{\Psi}_{\text{in}}$ and $\mathbf{\Psi}_{\text{out}}$ illustrate the value of our incremental testing procedure for acceptability of installation configurations for large distribution networks. The

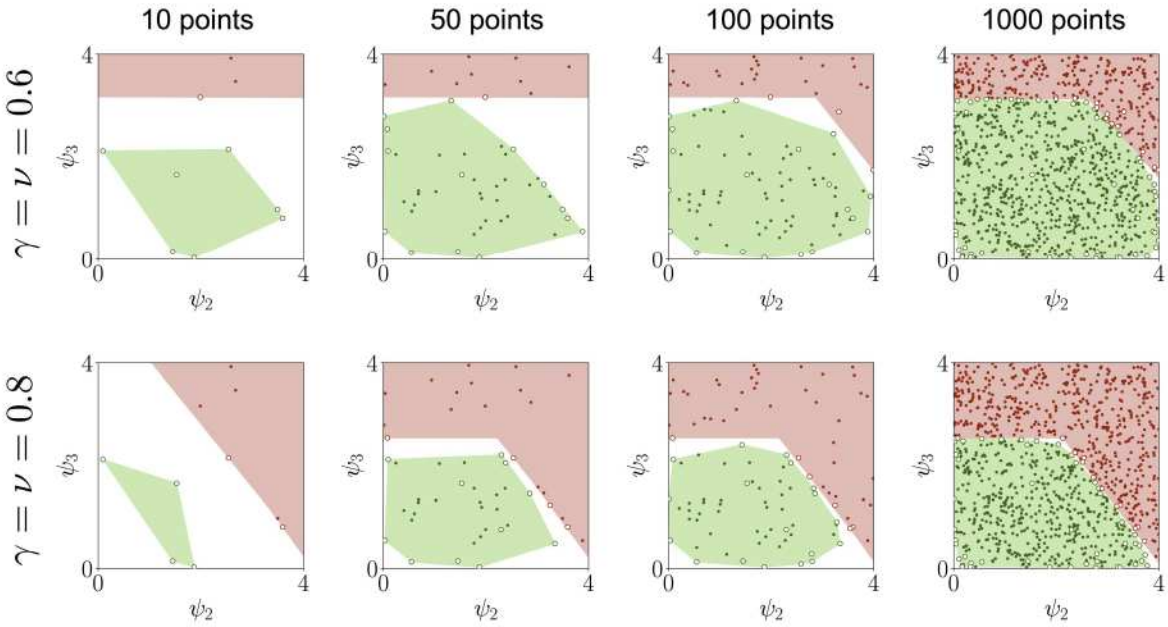

Fig. 4. Visualization of $\mathbf{\Psi}_{\text{in}}$ (green) and $[0, \overline{\psi}] \setminus \mathbf{\Psi}_{\text{out}}$ (red) regions after testing for acceptability of 10, 50, 100 and 1000 configurations on the 3-bus network example with $\nu = \gamma = 0.6$ on the top and $\nu = \gamma = 0.8$ on the bottom. Hollow points required solving (21), while solid points could be certified using $\mathbf{\Psi}_{\text{in}}/\mathbf{\Psi}_{\text{out}}$.

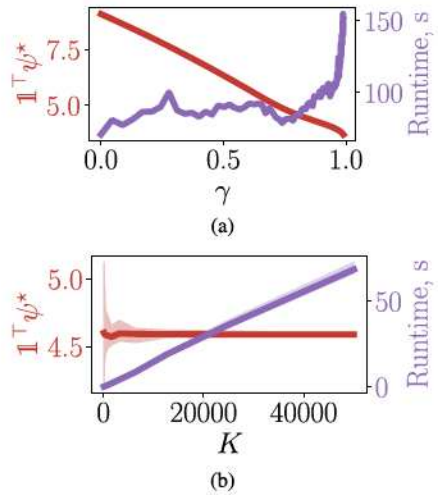

Fig. 5. 3-bus network results: (a) the effect of the CVaR parameter $\gamma = \nu$ on the maximum capacity $\mathbb{1}^\top\psi^\star$ and optimization runtime, and (b) the standard deviation in solutions with increasing number of samples K for $\gamma = \nu = 0.8$.

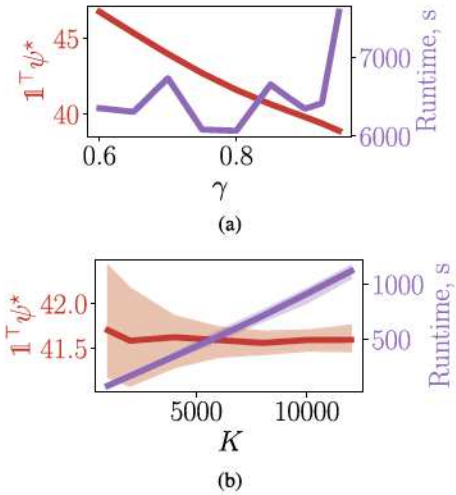

Fig. 6. 56-bus network results with $\nu = 0.9$: (a) the effect of the CVaR parameter γ on the maximum capacity and runtime, and (b) the deviation in solutions with increasing number of samples K for $\gamma = 0.8$.

savings in run-times can only increase with larger number of scenarios, network sizes, and number of tests conducted.

B. Numerical Experiments for Maximal Hosting Capacity

1) On a Three-Bus Network: In Fig. 5(a), we examine the effect of the risk parameters ($\gamma = \nu$) on both the solution and the run-time. The optimal aggregate capacity $\mathbb{1}^\top\psi^\star$ decreases with risk aversion. This is anticipated as a higher risk parameter leads to a tighter risk-sensitive constraint, shrinking the feasible set. In turn, this leads to a decreased maximum total installed capacity. We believe that increasing run-times are a consequence of the tightening constraints.

The problem dimension scales linearly with the number of scenarios K. Each scenario introduces $9n - 5$ constraints and $7n - 4$ variables, where n is the number of buses. In Fig. 5(b), we evaluate the effect of sub-sampling on the resulting solution, plotting the mean optimal solution with deviations over 20 runs per sample size. Fewer samples result in greater variance in the solution, but reduced run-times.

2) On a 56-Bus Distribution Network: Fig. 6 shows the effect of the risk parameter γ (holding ν constant) on both run-times

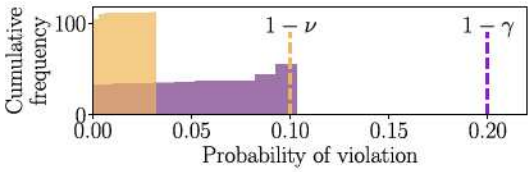

Fig. 7. Cumulative frequency histogram of probability of voltage and power flow constraint violation across 52,560 scenarios, under the optimal solution $\psi^\star$ obtained with $\nu = 0.9$, $\gamma = 0.8$ in the 56-bus network.

and optimal objective values attained with $K = 52.56$K scenarios. As in the 3-bus example, the objective value decreases with ν. The algorithm run-times are 1-2 hours for each value of risk parameter considered, and increase significantly for γ close to 1.

As previously noted, for a random variable ζ, the constraint $\text{CVaR}_\delta[\zeta] \leq 0$ is a convex inner approximation of the constraint $\mathbb{P}[\zeta \geq 0] \leq 1 - \delta$. In other words, a configuration ψ satisfying the CVaR of voltage variations and power flows in OPT, will also satisfy the corresponding probabilistic constraints. To illustrate this property empirically with the optimal solution $\psi^\star$, consider the solution obtained with $\nu = 0.9$ and $\gamma = 0.8$. For this solution, we evaluate the proportion of the 52.56 K scenarios for which each of the 110 upper and lower voltage limits and 55 power flow limits are violated. Fig. 7 shows the cumulative distribution histogram of the probability of voltage violations across the buses and power flow violations across the lines. The support of voltage violation proportions lie fully to the left of $1 - \nu$ and the power flow violation proportions to the left of $1 - \gamma$. That is, our CVaR-based constraint enforcement automatically guarantees probabilistic constraint enforcement.

While our reported solutions used MOSEK, we have compared our results with an implementation in Gurobi. For high risk-aversion parameters, e.g., $\nu = \gamma = 0.9$, Gurobi took almost twice as long to produce an optimal solution, compared to MOSEK, even with a small number of scenarios. However, in performing our simulations, MOSEK tended to yield a stable optimal solution (approximately matching Gurobi) the fastest, when the objective function was scaled by a large parameter, e.g., 10^9. Run-times reduced by as much as a factor of two with such scaling. Objective function scaling did not have an impact on the Gurobi implementation.

Compared to the BayesOpt algorithm in [26], our implementation is *centralized*, and naturally requires more memory than a solution architecture that solves scenario-wise power flow problems. A side-by-side comparison with earlier works remains difficult, given that optimization with probabilistic constraint enforcement has no optimality guarantees and its outcome depends strongly on implementation details.

One might surmise that our formulation of OPT can be parallelized using an algorithmic architecture such as generalized Benders' decomposition in [27]. Indeed, one can devise such a method that processes scenarios ω^k's in parallel. However, our implementation of generalized Benders' decomposition for OPT led to increasingly larger master problems, resulting from feasibility cuts corresponding to individual scenarios, that proved computationally prohibitive. Given that the centralized implementation ran in reasonable time, we did not further pursue the parallelized approach. A more careful implementation with lazy constraint generation protocols is left for future endeavors.

The proposed approach evaluates solar hosting capacity based on the statistics of line flows and voltages across a large set of load and solar irradiance scenarios. One may be tempted to rely on heuristics to identify "important" samples and then ensure that the engineering constraints are not violated with a given solar installation configuration across these chosen scenarios. Such a *robust* enforcement of constraints with heuristically chosen scenarios may yield conservative solutions, or actually turn out to be *infeasible*. For example, for our 56-bus network, the voltage and line flow limits are tight enough that no installed solar configuration (not even zero) guarantees feasibility with all scenarios.

In order to provide a meaningful comparison between robust optimization and our CVaR-based approach, we implemented another experiment. Specifically, we considered a problem similar to OPT that optimized over ψ, but enforced voltage and line capacity limits on each sample with a slack as follows. Instead of (19a), we imposed

$$\boldsymbol{W}^k \leq \overline{\boldsymbol{W}} + \zeta^k, \tag{29}$$

where the slack $\zeta^k \geq 0$ was penalized by adding an additional term $C\zeta^k$ to the objective function, with $C > 0$. That is, the optimization problem allows, but penalizes constraint violations. Other constraints were similarly altered, with analogous terms added to the objective function. With such an optimization formulation, we then considered a limited set of 8 scenarios, corresponding to the lowest and highest aggregate net load for various levels of renewable penetration. In other words, we sought the maximal hosting capacity that led to the least constraint violations with specific load/irradiance scenarios. The resulting maximal capacity total was found to be over 25% lower than that obtained from OPT with $\nu = 0.9, \gamma = 0.8$. That is, the hosting capacity with heuristically chosen scenarios is conservative, compared to that obtained with the risk-sensitive counterpart. The configuration obtained with heuristic scenario selection led to constraint violations in 27.7% of the scenarios, when evaluated across all 56.5 K scenarios. The optimal configuration from the CVaR-based approach, on the other hand, led to violations in 30.3% of the scenarios. While instances of constraint violation are more frequent in the risk-sensitive approach, the maximum expected violation magnitudes do not increase. The CVaR-based approach processes the entire statistics of possible violations and avoids conservative installations reached with heuristic scenario selection, thereby automating risk handling.

VII. Conclusion

Having already experienced several years of sustained growth, the continued adoption of distributed and utility scale solar generation is in many jurisdictions not only expected, but now

mandated by legislation. The primary question facing distribution grid operators and planners has therefore become not if, but how and when solar will be integrated at scale. HCA assesses the limits of a network for safely accommodating solar generation capacity. Given the highly intermittent nature of solar generation, together with changing consumption patterns, moving forward, notions of safety must account for the statistics of both present and forecasted operating conditions. In order to address the scalability issues arising from consideration of large numbers of scenarios, in this work we proposed an HCA formulation based on the use of CVaR as a network constraint risk measure. CVaR provides guarantees as to the frequency and extent of constraint violations, and moreover possesses desirable mathematical properties yielding a tractable approach to HCA. Further leveraging these properties, we provided an expedited method for assessing the acceptability of a given installation configuration, using prior knowledge of installations certified as either acceptable or unacceptable.

There are a number of interesting research directions we aim to pursue. First, while our optimization approach requires no specialized algorithm design or tuning, the resulting problems are still large and thus, require considerable memory resources. It is desirable to develop decomposition techniques that alleviate such difficulties, without sacrificing quality of the solutions obtained. Second, we aim to explore the potential of scenario reduction, e.g., via clustering, to help relax the computational burden. Additionally, we want to model and evaluate the effect of active control of distributed energy resources in CVaR-sensitive HCA with possibly unbalanced three-phase distribution networks. Thinking forward, many important questions remain as to how the information provided by our assessment tools might be used in prioritization and pricing of proposed capacity expansions or grid enhancements. For instance, bids on the remaining capacity in a network could be assessed on the basis of the combined benefit and risk introduced.

APPENDIX

A. Quadratic Inequality Constraints in SOC Form

For any scalars Z_1, Z_2, z_1, z_2, we have

$$\begin{aligned} & z_1 z_2 \geq Z_1^2 + Z_2^2 \\ \iff & \left\| \left(2Z_1, 2Z_2, z_1 - z_2\right)^{\mathsf{T}} \right\|_2 \leq z_1 + z_2 \\ \iff & \left(2Z_1, 2Z_2, z_1 - z_2, z_1 + z_2\right)^{\mathsf{T}} \in \mathcal{K}_{\mathrm{SO}}. \end{aligned} \tag{30}$$

B. Proof of Proposition 1

Consider the following Lagrangian function for (21),

$$\begin{aligned} & \mathcal{L}(x, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2) \\ & = \boldsymbol{\lambda}^{\mathsf{T}}(\boldsymbol{C}_\psi \psi - \boldsymbol{D}_\psi \boldsymbol{x} - \boldsymbol{E}_\psi) + \boldsymbol{\mu}^{\mathsf{T}}(\boldsymbol{D}\boldsymbol{x} - \boldsymbol{E}) \\ & \quad - \boldsymbol{\mu}_1^{\mathsf{T}}(\boldsymbol{F}\boldsymbol{x} + \boldsymbol{G}) - \mu_2(\boldsymbol{f}^{\mathsf{T}}\boldsymbol{x} + g) \end{aligned} \tag{31}$$

with Lagrange multipliers $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2$ of compatible dimensions. Then, (21) can be written as

$$-\mathsf{ACC}^\star(\psi) = \min_{\boldsymbol{x}} \max_{\substack{\boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0, \\ (\boldsymbol{\mu}_1, \mu_2) \in \mathcal{K}_{\mathrm{SO}}}} \mathcal{L}(x, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2), \tag{32}$$

where $\mathsf{ACC}^\star(\psi)$ denotes the optimal objective function value of $\mathsf{ACC}(\psi)$. The above representation leverages the fact that $\mathcal{K}_{\mathrm{SO}}$ is a self-dual convex cone (see [32] for details). Weak duality implies

$$-\mathsf{ACC}^\star(\psi) \geq \max_{\substack{\boldsymbol{\lambda}, \boldsymbol{\mu} \geq 0, \\ (\boldsymbol{\mu}_1, \mu_2) \in \mathcal{K}_{\mathrm{SO}}}} \min_{\boldsymbol{x}} \mathcal{L}(x, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2), \tag{33}$$

where the right hand side of the above inequality equals the optimal cost of the dual problem. Solving the inner minimization over x in the above relation, yields

$$\begin{aligned} -\mathsf{ACC}^\star(\psi) \geq \ & \underset{\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2}{\text{maximum}} \quad \mathcal{J}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2; \psi), \\ & \text{subject to} \quad \boldsymbol{D}^{\mathsf{T}}\boldsymbol{\mu} = \boldsymbol{D}_\psi^{\mathsf{T}}\boldsymbol{\lambda} + \boldsymbol{F}^{\mathsf{T}}\boldsymbol{\mu}_1 + \boldsymbol{f}\mu_2, \\ & \qquad\qquad \boldsymbol{\mu} \geq 0, \ (\boldsymbol{\mu}_1^{\mathsf{T}}, \mu_2)^{\mathsf{T}} \in \mathcal{K}_{\mathrm{SO}}, \end{aligned} \tag{34}$$

recalling that

$$\begin{aligned} & \mathcal{J}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2; \psi) \\ & \quad = (\boldsymbol{C}_\psi \psi - \boldsymbol{E}_\psi)^{\mathsf{T}}\boldsymbol{\lambda} - \boldsymbol{E}^{\mathsf{T}}\boldsymbol{\mu} - \boldsymbol{G}^{\mathsf{T}}\boldsymbol{\mu}_1 - \mu_2 g. \end{aligned} \tag{35}$$

The right-hand-side of (34) is the dual program for (21)–an SOCP for which the origin is always feasible. The feasible set for this dual problem is a convex cone.[3] For any feasible $(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2)$ yielding a positive dual objective, we have

$$\begin{aligned} & \mathcal{J}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\mu}_1, \mu_2) > 0 \\ & \Rightarrow \lim_{\kappa \to \infty} \mathcal{J}(\kappa\boldsymbol{\lambda}, \kappa\boldsymbol{\mu}, \kappa\boldsymbol{\mu}_1, \kappa\mu_2; \psi) = \infty \\ & \Rightarrow \mathsf{ACC}^\star(\psi) = -\infty, \end{aligned} \tag{36}$$

in turn implying $\mathsf{ACC}(\psi)$ is infeasible and ψ is unacceptable.

REFERENCES

[1] S. Joshi, S. Mittal, P. Holloway, P. R. Shukla, B. Ó Gallachóir, and J. Glynn, "High resolution global spatiotemporal assessment of rooftop solar photovoltaics potential for renewable electricity generation," *Nature Commun.*, vol. 12, no. 1, pp. 1–15, 2021.

[2] J. Spector, "Hawaii is ahead of schedule for renewable power adoption," Feb. 2021, Accessed: Feb. 22, 2022. [Online]. Available: https://www.greentechmedia.com/articles/read/hawaii-is-ahead-of-schedule-for-renewable-power-adoption

[3] B. Azibek, A. Abukhan, H. K. Nunna, B. Mukatov, S. Kamalasadan, and S. Doolla, "Hosting capacity enhancement in low voltage distribution networks: Challenges and solutions," in *Proc. IEEE Int. Conf. Power Electron., Smart Grid Renewable Energy*, 2020, pp. 1–6.

[4] U. E. I. Administration, "California's curtailments of solar electricity generation continue to increase," Accessed: Feb. 22, 2022. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=49276

[5] E. Mulenga, M. H. Bollen, and N. Etherden, "A review of hosting capacity quantification methods for photovoltaics in low-voltage distribution grids," *Int. J. Elect. Power Energy Syst.*, vol. 115, 2020, Art. no. 105445.

[6] N. Baldenko and S. Behzadirafi, "Determination of photovoltaic hosting capacity on radial electric distribution feeders," in *Proc. IEEE Int. Conf. Power Syst. Technol.*, 2016, pp. 1–4.

[3] $\mathcal{S}$ is a convex cone, if $s_1, s_2 \in \mathcal{S}, \kappa_1, \kappa_2 \geq 0$ implies $\kappa_1 s_1 + \kappa_2 s_2 \in \mathcal{S}$.

[7] F. Capitanescu, L. F. Ochoa, H. Margossian, and N. D. Hatziargyriou, "Assessing the potential of network reconfiguration to improve distributed generation hosting capacity in active distribution systems," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 346–356, Jan. 2015.

[8] FERC, "Small generator interconnection procedures (SGIP)," Tech. Rep. RM16-1-000, 2020.

[9] F. A. Viawan, F. Vuinovich, and A. Sannino, "Probabilistic approach to the design of photovoltaic distributed generation in low voltage feeder," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst.*, 2006, pp. 1–7.

[10] M. Rossi, G. Viganò, and D. Moneta, "Hosting capacity of distribution networks: Evaluation of the network congestion risk due to distributed generation," in *Proc. IEEE Int. Conf. Clean Elect. Power*, 2015, pp. 716–722.

[11] F. Golestaneh, P. Pinson, and H. B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generation–with application to solar energy," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3850–3863, Sep. 2016.

[12] G. A. Hanasusanto, D. Kuhn, and W. Wiesemann, "A comment on "computational complexity of stochastic programming problems," *Math. Program.*, vol. 159, no. 1, pp. 557–569, 2016.

[13] J. Kisiala, "Conditional value-at-risk: Theory and applications," 2015, *arXiv:1511.00140*.

[14] A. E. Lim, J. G. Shanthikumar, and G.-Y. Vahn, "Conditional value-at-risk in portfolio optimization: Coherent but fragile," *Operations Res. Lett.*, vol. 39, no. 3, pp. 163–171, 2011.

[15] S. Bruno, S. Ahmed, A. Shapiro, and A. Street, "Risk neutral and risk averse approaches to multistage renewable investment planning under uncertainty," *Eur. J. Oper. Res.*, vol. 250, no. 3, pp. 979–989, 2016.

[16] A. N. Madavan, S. Bose, Y. Guo, and L. Tong, "Risk-sensitive security-constrained economic dispatch via critical region exploration," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2019, pp. 1–5.

[17] J. Smith and M. Rylander, "Stochastic analysis to determine feeder hosting capacity for distributed solar PV," Electric Power Research Inst., Palo Alto, CA, USA, Tech. Rep. 1026640, 2012.

[18] F. Ding and B. Mather, "On distributed PV hosting capacity estimation, sensitivity study, and improvement," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 1010–1020, Jul. 2017.

[19] M. S. S. Abad, J. Ma, D. Zhang, A. S. Ahmadyar, and H. Marzooghi, "Probabilistic assessment of hosting capacity in radial distribution systems," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1935–1947, Oct. 2018.

[20] S. F. Santos, D. Z. Fitiwi, M. Shafie-Khah, A. W. Bizuayehu, C. M. Cabrita, and J. P. Catalão, "New multistage and stochastic mathematical model for maximizing res hosting capacity–Part I: Problem formulation," *IEEE Trans. Sustain. Energy*, vol. 8, no. 1, pp. 304–319, Jan. 2017.

[21] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 725–734, Jan. 1989.

[22] A. H. Arab et al., "Maximum power output performance modeling of solar photovoltaic modules," *Energy Rep.*, vol. 6, pp. 680–686, 2020.

[23] H. Al-Saadi, R. Zivanovic, and S. F. Al-Sarawi, "Probabilistic hosting capacity for active distribution networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 5, pp. 2519–2532, Oct. 2017.

[24] S. Taheri, M. Jalali, V. Kekatos, and L. Tong, "Fast probabilistic hosting capacity analysis for active distribution systems," *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2000–2012, May 2021.

[25] S. H. Low, "Convex relaxation of optimal power flow–Part I: Formulations and equivalence," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 15–27, Mar. 2014.

[26] X. Geng, L. Tong, A. Bhattacharya, B. Mallick, and L. Xie, "Probabilistic hosting capacity analysis via Bayesian optimization," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2021, pp. 1–5.

[27] A. M. Geoffrion, "Generalized benders decomposition," *J. Optim. Theory Appl.*, vol. 10, no. 4, pp. 237–260, 1972.

[28] B. F. Lourenço, M. Muramatsu, and T. Tsuchiya, "Weak infeasibility in second order cone programming," *Optim. Lett.*, vol. 10, no. 8, pp. 1743–1755, 2016.

[29] B. O'donoghue, E. Chu, N. Parikh, and S. Boyd, "Conic optimization via operator splitting and homogeneous self-dual embedding," *J. Optim. Theory Appl.*, vol. 169, no. 3, pp. 1042–1068, 2016.

[30] L. Gan, N. Li, U. Topcu, and S. H. Low, "Exact convex relaxation of optimal power flow in radial networks," *IEEE Trans. Autom. Control*, vol. 60, no. 1, pp. 72–87, Jan. 2015.

[31] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[32] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.