Probabilistic Constraint Construction for Network-Safe Load Coordination

Sunho Jang, *Student Member, IEEE*, Necmiye Ozay, *Senior Member, IEEE*, and Johanna L. Mathieu, *Senior Member, IEEE*

***Abstract*—Distributed Energy Resources (DERs) can provide balancing services to the grid, but their power variations might cause voltage and current constraint violations in the distribution network, compromising network safety. This could be avoided by including network constraints within DER control formulations, but the entities coordinating DERs (e.g., aggregators) may not have access to network information, which typically is known only to the utility. Therefore, it is challenging to develop network-safe DER control algorithms when the aggregator is not the utility; it requires these entities to coordinate with each other. In this article, we develop an aggregator-utility coordination framework that enables network-safe control of thermostatically-controlled loads to provide frequency regulation. In our framework, the utility sends a network-safe constraint set on the aggregator's input without directly sharing any network information. We propose a constraint set construction algorithm that guarantees satisfaction of a chance constraint on network safety. Assuming monotonicity of the probability of network safety with respect to the aggregator's input, we leverage the bisection method to find the largest possible constraint set, providing maximum flexibility to the aggregator. Simulations show that, compared to two benchmark algorithms, the proposed approach provides a good balance between service quality and network safety.**

***Index Terms*—Chance constraints, distributed energy resources, load control, network safety, thermostatically-controlled loads.**

I. Introduction

As THE amount of intermittent renewable generation is rapidly growing, it is becoming more difficult to rely solely on the conventional ways of balancing power systems. One emerging solution is to leverage Distributed Energy Resources (DERs), such as thermostatically-controlled loads (TCLs), batteries, and electric vehicles, to provide grid services, which can improve grid reliability and reduce grid operating costs and environmental impacts. However, DERs coordinated to provide balancing services might cause issues in the distribution network, such as under/over-voltages, over-current violations, and transformer overheating, compromising network safety.

Manuscript received 2 February 2023; revised 9 May 2023 and 4 August 2023; accepted 21 August 2023. Date of publication 28 August 2023; date of current version 21 February 2024. This work was supported by US NSF under Grant CNS-1837680. Paper no. TPWRS-00153-2023. *(Corresponding author: Sunho Jang.)*

The authors are with the Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor MI 48109 USA (e-mail: sunhoj@umich.edu; necmiye@umich.edu; jlmath@umich.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPWRS.2023.3309131.

Digital Object Identifier 10.1109/TPWRS.2023.3309131

When the distribution network operator (i.e., the utility) coordinates DERs to provide grid services it can adopt a centralized algorithm that explicitly manages distribution network constraints, e.g., the algorithms provided in [1], [2], [3]. However, in competitive U.S. electricity markets it is becoming more likely that third-party (i.e., non-utility) DER aggregators will take on this role. Unfortunately, the aggregator does not have access to detailed distribution network information typically known only to the utility, and so it is unable to directly determine how its actions would affect the distribution network. This challenge has already been recognized by the US Federal Energy Regulatory Commission (FERC) [4].

Thus, there is a need for coordination between the aggregator and the utility to ensure network-safe operation of DERs. The recent FERC Order No. 2222 [5] provided some guidance on the development of operational coordination architectures between DER aggregators, utilities, and market coordinators; however, it is still unclear how these architectures will evolve and which architecture is "best." Beyond ensuring network safety, coordination architectures should also 1) ensure that each entity's private information (e.g., sensitive network information held by the utility, proprietary DER coordination strategies held by the aggregator, and private DER state information held by the DERs' end-users) is not shared with the other entities and 2) communication between the entities is minimal for compatibility with current communications infrastructure and/or to reduce the cost of any newly required infrastructure. Furthermore, architectures need to specify coordination protocols on different timescales, for example, 1) for operational planning such that the aggregator can determine its offer for balancing services, and 2) for real-time control in case network conditions differ significantly from forecasts and aggregator actions need to be curtailed.

In this article, we propose an aggregator-utility coordination framework for a collection of TCLs to provide balancing services like frequency regulation while ensuring distribution network-safety with high probability. We focus on real-time coordination, specifically, a setting in which an aggregator is already committed to provide a certain amount of balancing services, but real-time distribution network conditions require curtailment of those services. In our framework, the utility sends the aggregator a one-step ahead constraint set (i.e., a set computed at time step t to be implemented at time step $t+1$) on the aggregator's control input. Choosing an input from this set guarantees the satisfaction of a chance constraint on

network safety with a certain confidence level. This method leverages estimation from Monte Carlo simulation and the bi-section method to provide the largest possible constraint set to maximize the network-safe TCL flexibility. To achieve light communication requirements, the aggregator control algorithm assumes the TCLs all respond to the same scalar control input. This constrains the aggregator's degrees-of-freedom but also makes it possible for the utility to define a simple constraint set on the control input.

Previous work, e.g. [6], [7], [8], [9], has proposed strategies to control aggregations of TCLs, such as air conditioners and water heaters, to provide balancing services in ways that are non-disruptive to end-users. TCLs have inherent thermal energy storage capacity and non-disruptiveness can be achieved, e.g., by keeping internal temperatures inside a narrow temperature dead-band. However, network safety was not considered in the above papers. Some work has developed network-safe control algorithms for TCLs coordinated by third-party aggregators. Ref. [10] proposes both a utility-centric and an aggregator-centric coordination framework, differentiated by which entity ultimately sends control inputs to the TCLs. That paper and [11] develop utility-centric strategies wherein the utility blocks aggregator's inputs that would otherwise cause network constraint violations. In contrast, our proposed approach would be considered aggregator-centric.

Aggregator-centric network-safe DER coordination could be achieved through (convex) inner approximation of safe operating regions [12], [13], [14], which could be computed by the utility and sent to the aggregator as constraints on the net DER power deviations at each node. Research from Australia refers to these nodal constraints as operating envelopes [15], [16], [17]. Ref. [18] proposes an optimization problem to obtain a hyper-rectangular constraint set on the net DER power consumption at each node in order to satisfy chance constraints on the voltage at each node. However, these approaches all require constraints to be applied at each node, rather than applying a constraint on aggregate power deviations by DERs located across a network. Ref. [19] proposes a method to constrain the norm of the power deviations across all nodes in the network, but requires significant computation to compute the constraint. Assuming an aggregate power deviation constraint exists, our previous work [20] develops an aggregator-centric TCL coordination algorithm using formal methods, but does not develop an approach to obtain the constraint, and the solutions are very conservative.

In contrast to this previous work, this article makes the following contributions: 1) we develop a new aggregator-centric approach to enable network-safe control of TCLs for balancing services; 2) assuming a simple control scheme that leverages a scalar control input to coordinate TCLs to provide balancing services (the aggregator's algorithm), we develop an approach to constrain the control input to satisfy a chance constraint on network safety (the utility's algorithm); and 3) we demonstrate our approach in simulation and compare its performance to two benchmark approaches. In contrast to past work on network-safe control that assumes the system is deterministic, e.g., [19], here we consider uncertainty in the power consumption of non-participating loads. Furthermore, in contrast to [20], we assume the aggregator has incomplete information about the TCLs to reduce communication requirements and preserve some level of privacy. Lastly, though some past work leveraged chance constraints to develop network-safe DER coordination approaches, e.g., [18], [21], [22], [23], [24], [25], [26], these papers all assume that the controller has detailed distribution network information (enabling the formulation of a chance-constrained optimal power flow problem), which is inconsistent with our utility-aggregator coordination framework. Several approaches attempt to tackle the lack of detailed network information within aggregator-centric frameworks, including via sensitivity analysis as demonstrated in [27] and via reinforcement learning as proposed in [28]. However, these approaches do not formally guarantee network safety nor do they satisfy a chance constraint on network safety.

This article is organized as follows. Section II introduces the problem of interest. Section III explains the aggregator's control approach and Section IV details the proposed constraint construction algorithm used by the utility to achieve network safety at a high level of probability. Section V presents the results of a case study comparing the proposed approach to two benchmarks. The appendix includes proofs of two of the theorems.

Notation: $\mathbb{N}$ is the set of natural numbers. $[N]$ denotes $\{1,\dots,N\}$ and $[N]_0$ denotes $\{0,1,\dots,N\}$. The jth element of vector $\boldsymbol{y}$ is y_j. Binomial distribution $\mathcal{B}(n_\mathrm{s},\nu)$ has n_s trials, each with success probability ν, and cumulative density function (cdf) $\mathcal{F}_\mathrm{B}(x;n_\mathrm{s},\nu)$. $\mathcal{N}(\mu,\sigma^2)$ is the normal distribution with mean μ and variance σ^2. Function $\mathbb{1}(A)$ is 1 if A is true, and 0 otherwise. All random variables are capitalized English letters, e.g., X, with realizations denoted $\tilde{x}$ and estimates/approximates denoted $\hat{x}$. All other variables are denoted by symbols other than capitalized English letters. Vectors and matrices are bolded.

II. Framework & Problem of Interest

We consider a framework in which an aggregator and utility coordinate to provide network-safe grid balancing services, e.g., frequency regulation, by aggregations of TCLs. TCLs switch ON/OFF to maintain temperature within a dead-band. We focus on real-time coordination, i.e., we assume that the aggregator has already participated in the ancillary services market and committed balancing service capacity to the independent system operator (ISO). The amount of balancing service capacity offered by the aggregator was based on forecasts of the capabilities of the TCLs and the network state. However, the real-time network state differs significantly from its forecasts and so the committed balancing service capacity must be curtailed to avoid distribution network constraint violations. This could happen when load consumption and/or renewable power injections are significantly different from forecasts and the network is operating close to its limits.

We assume that the following coordination steps occur at each discrete time step t, with a sampling time Δt. The coordination scheme is shown in Fig. 1.

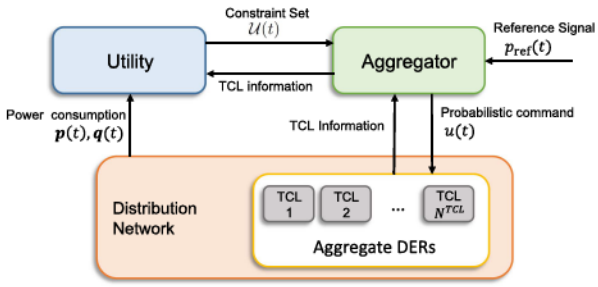

Fig. 1. Aggregator-utility coordination scheme.

1) The aggregator receives a constraint set $\mathcal{U}(t)$ from the utility and a reference signal $p_{\text{ref}}(t)$ (e.g., a scaled and shifted frequency regulation signal) from the ISO.
2) The aggregator determines the input $u(t) \in \mathcal{U}(t)$ and broadcasts the same input to all TCLs.
3) Each TCL maintains or switches its ON/OFF mode based on its temperature and the aggregator's input $u(t)$.
4) The utility observes the active and reactive power consumption at each network node $p(t)$ and $q(t)$, and obtains some information from the aggregator (described below). Then, it constructs a one-step ahead constraint set $\mathcal{U}(t+1)$ and sends it to the aggregator. (And go back to step 1.)

The aggregator's goal is to select $u(t)$ to maximize the quality of grid balancing services. This means that the aggregator should choose an input $u(t)$ that is likely to adjust the aggregate power of the TCLs to match the reference signal $p_{\text{ref}}(t)$ as closely as possible. Here, we assume the aggregator's input $u(t)$ is a real scalar in the range $[-1, 1]$ and is interpreted by each TCL as the probability it should switch modes; the details of how it switches are given in Section III. TCL coordination through probabilistic switching has been considered in previous work e.g., [6], [10]. An advantage of this type of input is that it only needs simple broadcast communication infrastructure. However, it does not allow the aggregator to directly adjust the power consumption of individual TCLs, which means that the aggregator has a low degree-of-freedom in control.

Since the aggregator does not have detailed distribution network information and cannot evaluate how its input would affect the network, the utility sends a constraint set $\mathcal{U}(t+1)$ on the aggregator's input $u(t+1)$. This set $\mathcal{U}(t+1)$ is designed such that, if $u(t+1) \in \mathcal{U}(t+1)$, then probability of network safety is over a desired value $1-\epsilon$. We propose a method for the utility to compute $\mathcal{U}(t+1)$ in Section IV, which is the main contribution of this work. To do this, the utility leverages:

1) Real-time data from household smart meters to obtain the active and reactive power consumption at each node, $p(t)$ and $q(t)$. We recognize that in practice most utilities do not currently gather smart meter data in real-time, but this is possible with most existing smart meters and could be enabled through reconfiguration of their settings.

2) Forecasts of the probability distributions of the one-step ahead active and reactive power consumption of non-participating loads at each node, $\boldsymbol{P}^{\text{L}}(t+1)$ and $\boldsymbol{Q}^{\text{L}}(t+1)$. We assume that these distributions are estimated using historical and real-time data from household smart meters, and leveraging a disaggregation technique [29] to separate the power consumption of the TCLs from that of the non-participating loads. We assume that $\boldsymbol{P}^{\text{L}}(t)$ and $\boldsymbol{Q}^{\text{L}}(t)$ are correlated and $f_{\boldsymbol{P}^{\text{L}},\boldsymbol{Q}^{\text{L}}}(t)$ is their joint probability density function (pdf), which we can also estimate from historical and real-time data. We note that energy disaggregation is imperfect and that estimating $f_{\boldsymbol{P}^{\text{L}},\boldsymbol{Q}^{\text{L}}}$ accurately requires a large amount of data, particularly for large-scale distribution networks. In this article, we assume that the disaggregation process is accurate and we have sufficient data to estimate $f_{\boldsymbol{P}^{\text{L}},\boldsymbol{Q}^{\text{L}}}$ with high precision.

3) Some real-time TCL information from the aggregator that is necessary for constraint set computation. This should be minimal to protect end-user privacy. In our framework, the aggregator provides the one-step ahead estimated fractions of TCLs that will be outside of their temperature dead-band and switched OFF-to-ON and ON-to-OFF by their thermostats, $\hat{w}^{\text{ON}}(t+1)$ and $\hat{w}^{\text{OFF}}(t+1)$. Details on how this information is used are provided in Section IV-A.

In this article, for the sake of simplicity, we define network safety in terms of under-voltage violations. Specifically, we say that the network is safe if there are no under-voltage violations, and unsafe if there are any violations. The approach can be easily extended to include over-voltage violations and other distribution network constraint violations. The problem can be formally stated as follows.

Problem 1: Given the desired safety probability $1-\epsilon$, the real-time active and reactive power consumption at each node $p(t)$ and $q(t)$, the joint pdfs of the uncontrollable loads $f_{\boldsymbol{P}^{\text{L}},\boldsymbol{Q}^{\text{L}}}(t)$, $f_{\boldsymbol{P}^{\text{L}},\boldsymbol{Q}^{\text{L}}}(t+1)$, and the fractions of TCLs that are outside of their dead-band $w^{\text{ON}}(t+1)$, $w^{\text{OFF}}(t+1)$, find a constraint set $\mathcal{U}(t+1)$ such that the following chance constraint holds if $u(t+1) \in \mathcal{U}(t+1)$,

$$\Pr\left(\min_{j\in[n]} V_j(t+1) \geq \underline{v}\right) \geq 1-\epsilon, \tag{1}$$

where $\underline{v}$ is the lower bound on each of the nodal voltages V_j and n is the number of nodes other than the substation.

To solve this problem, we define the one-time step ahead voltage at each node $V_j(t+1)$ as a random variable whose distribution depends on the input $u(t+1)$; the details are explained in Section IV. It is difficult to obtain a closed-form expression for the probability distribution of each $V_j(t+1)$. Therefore, our approach leverages Monte Carlo simulation to estimate the left side of (1) given a one-step ahead input $u(t+1)$. Since estimation from sampling leads to error, we find a constraint set $\mathcal{U}(t+1)$ with a confidence level over a desired level $1-\beta$ rather than giving an exact solution.

III. Aggregator's Control Approach

In this section, we explain how the TCLs operate under the aggregator's input $u(t)$. For simplicity, we assume that all participating TCLs are cooling TCLs (e.g., air conditioners), though the approach also applies to heating TCLs. We denote by $\boldsymbol{n}^{\text{TCL}}$ the vector whose element n_j^{TCL} is the number of participating

TCLs at node j. The total number of participating TCLs is denoted by

$$n^{\text{TCL}} := \mathbf{1}^\top \boldsymbol{n}^{\text{TCL}} = \sum_{j=1}^{n} n_j^{\text{TCL}}.$$

The internal temperature of the ith TCL at time t is denoted by $\theta^i(t)$ and its mode is denoted by $m^i(t)$, which is 0 when it is OFF, and 1 when it is ON. The temperature dynamics of the ith TCL follow the affine model from [30],

$$\theta^i(t+1) = a_{\text{th}}^i \theta^i(t) + \left(1 - a_{\text{th}}^i\right)\left(\theta_a^i(t) + r_{\text{th}}^i p_{\text{tr}}^i m^i(t)\right), \quad (2)$$

where $\theta_a^i(t)$ is the ambient temperature and

$$a_{\text{th}}^i = \exp\left(-\frac{\Delta t}{r_{\text{th}}^i c_{\text{th}}^i}\right),$$

where r_{th}^i is the thermal resistance and c_{th}^i is the thermal capacitance of the ith TCL. Also, p_{tr}^i is the energy transfer rate of the ith TCL, which is negative for a cooling TCL. The power consumption of the ith TCL in the ON mode is

$$p^i := p_{\text{tr}}^i / \zeta^i,$$

where ζ^i is the coefficient of performance; the power consumption in the OFF mode is 0. We assume that the reactive power consumption of the ith TCL is

$$q^i := \omega^i p^i,$$

where ω^i is a positive constant. The aggregate active power consumption of the TCLs is

$$p_{\text{agg}}(t) := \sum_{i=1}^{n^{\text{TCL}}} p^i m^i(t).$$

Each TCL has a temperature range $[\underline{\theta}^i, \overline{\theta}^i]$ within which its internal temperature should always be; this range is called the temperature dead-band. The temperature set-point

$$\theta_{\text{s}}^i := \frac{\underline{\theta}^i + \overline{\theta}^i}{2},$$

which is set by its end-user, is the midpoint of the dead-band. Whenever a TCL's internal temperature reaches or goes beyond the boundary of its dead-band it switches its mode to go back into the dead-band.

At each time step t, the aggregator determines its input $u(t)$ and broadcasts it to all participating TCLs. TCLs within their dead-bands interpret this input as the desired probability of OFF TCLs to switch ON when $u(t) > 0$, and the desired probability of ON TCLs to switch OFF when $u(t) < 0$. To determine whether or not to switch, each TCL draws a random number $z^i(t)$ from the uniform distribution on the interval $[0, 1)$ and compares it to the input $u(t)$. If it is OFF and $z^i(t) \le u(t)$, then it switches ON. If it is ON and $z^i(t) \le -u(t)$, then it switches OFF.

In summary, the mode of the ith TCL is

$$m^i(t) = \begin{cases} 1 & \text{if } \theta^i(t) \ge \overline{\theta}^i \\ 0 & \text{if } \theta^i(t) \le \underline{\theta}^i \\ m_c(z^i(t), u(t)) & \text{otherwise,} \end{cases} \quad (3)$$

where $m_c(z^i(t), u(t))$ is equal to

$$\begin{cases} 1 & \text{if } m^i(t-1) = 0 \text{ and } z^i(t) \le u(t) \\ 0 & \text{if } m^i(t-1) = 1 \text{ and } z^i(t) \le -u(t) \\ m^i(t-1) & \text{otherwise.} \end{cases}$$

Note that, when positive (negative) $u(t)$ is broadcast to the TCLs, the fraction of the OFF (ON) TCLs within their dead-bands that are switched is approximately $u(t)$ ($-u(t)$). Thus, $|u(t)|$ can be interpreted by the aggregator as the *ratio* of the power consumption increase (decrease) compared to the maximal increase (decrease). Therefore, even though the power consumption of each TCL is not directly controlled by the aggregator, the aggregator can manipulate the aggregate power consumption of the TCLs $p_{\text{agg}}(t)$ by selecting the input $u(t) \in \mathcal{U}(t)$ that is likely to adjust $p_{\text{agg}}(t)$ to match the reference signal $p_{\text{ref}}(t)$ as closely as possible, i.e., the optimal input,

$$u_{\text{opt}}(t) = \arg\min_{u \in \mathcal{U}(t)} \left|\mathbb{E}\left[P_{\text{agg}}(t)\right] - p_{\text{ref}}(t)\right|, \quad (4)$$

where $\mathcal{U}(t)$ is provided by the utility.

Since the aggregate power consumption of TCLs under command $u(t)$ is stochastic, there may be deviations from the expected value. However, when a large number of TCLs participate, by the central limit theorem, the aggregate power consumption of the TCLs is usually close to the expected value and, in turn, the reference signal. This claim is empirically verified through the case study in Section V.

IV. Utility's Constraint Construction Method

As mentioned in Section II, the utility computes a constraint set $\mathcal{U}(t+1)$, which should be a solution to Problem 1. This requires the utility to be able to evaluate how the input $u(t+1)$ would affect the probability of network safety. In this section, we first show how the voltage at each node is modeled as a random variable. For ease of exposition, we consider only balanced radial distribution networks. Then, we derive the probability of network safety (i.e., the probability that no under-voltage violations happen) as a function of the input $u(t+1) = u$.

Next, we show how to verify whether or not the chance constraint (1) is satisfied under $u(t+1) = u$ with a desired confidence level, and how the utility can construct $\mathcal{U}(t+1)$ to ensure (1) is satisfied. We introduce a theorem establishing a confidence interval for the success probability of a Bernoulli random variable using Monte Carlo simulations. Using this result, we leverage the bisection method to find the largest upper bound on $u(t+1)$ that guarantees (1) with a desired confidence level. The largest upper bound gives the aggregator the greatest possible flexibility in determining its input.

A. Modeling the Probability of Network Safety

We denote the active and reactive power consumption of participating TCLs across all nodes by $\boldsymbol{P}^{\text{T}}(t)$ and $\boldsymbol{Q}^{\text{T}}(t) \in \mathbb{R}^n$. The utility approximates the nodal values as

$$P_j^{\text{T}}(t) \approx \overline{p}_j N_j^{\text{ON}}(t), \;\; Q_j^{\text{T}}(t) \approx \overline{q}_j N_j^{\text{ON}}(t) \;\; \forall j \in [n], \quad (5)$$

where $N_j^{\mathrm{ON}}(t)$ and $N_j^{\mathrm{OFF}}(t)$ are the number of ON and OFF TCLs at node j, and $\overline{p}_j$ and $\overline{q}_j$ are the average active and reactive power rating (i.e., the ON-mode consumption) of the TCLs at node j. We additionally define diagonal matrices Ξ_p and $\Xi_q \in \mathbb{R}^{n\times n}$ whose jth diagonal elements are $\overline{p}_j$ and $\overline{q}_j$, respectively. Then, $\boldsymbol{P}^{\mathrm{T}}(t) = \Xi_p \boldsymbol{N}^{\mathrm{ON}}(t)$ and $\boldsymbol{Q}^{\mathrm{T}}(t) = \Xi_q \boldsymbol{N}^{\mathrm{ON}}(t)$, and the total active and reactive power consumption across all nodes is $\boldsymbol{P}(t) = \Xi_p \boldsymbol{N}^{\mathrm{ON}}(t) + \boldsymbol{P}^{\mathrm{L}}(t)$ and $\boldsymbol{Q}(t) = \Xi_q \boldsymbol{N}^{\mathrm{ON}}(t) + \boldsymbol{Q}^{\mathrm{L}}(t)$.

We first show how the one-step ahead number of ON TCLs $\boldsymbol{N}^{\mathrm{ON}}(t+1) \in \mathbb{R}^n$ is modeled as a random variable under the input $u(t+1) = u$. The number $\boldsymbol{N}^{\mathrm{ON}}(t+1)$ depends upon how many TCLs are switched both by their thermostat (i.e., the first and second cases of (3)) and by the aggregator's input (i.e., the third case of (3)). The number of TCLs at each node j that will be switched ON, OFF by their thermostats is

$$\begin{aligned} S_j^{\mathrm{ON}}(t+1) &= w_j^{\mathrm{ON}}(t+1) N_j^{\mathrm{OFF}}(t), \\ S_j^{\mathrm{OFF}}(t+1) &= w_j^{\mathrm{OFF}}(t+1) N_j^{\mathrm{ON}}(t), \end{aligned} \tag{6}$$

where, as defined in Section II, $w_j^{\mathrm{ON}}(t+1)$ is the one-step ahead fraction of OFF TCLs that will be switched ON and $w_j^{\mathrm{OFF}}(t+1)$ is the one-step ahead fraction of ON TCLs that will be switched OFF by their thermostats at bus j. We assume that the aggregator estimates $w_j^{\mathrm{ON}}(t+1)$ and $w_j^{\mathrm{OFF}}(t+1)$ using a model of the aggregate TCL dynamics and sends the estimated values $\hat{w}_j^{\mathrm{ON}}(t+1)$ and $\hat{w}_j^{\mathrm{OFF}}(t+1)$ to the utility, which corresponds to the TCL information illustrated in Fig. 1. The utility uses these estimates to obtain realizations of $S_j^{\mathrm{ON}}(t+1)$ and $S_j^{\mathrm{OFF}}(t+1)$ via Monte Carlo simulation, which will be explained in Section IV-B.

According to (3), the numbers of TCLs at each node j that will be switched ON and OFF by the aggregator's input follow binomial distributions,

$$\begin{aligned} C_{u,j}^{\mathrm{ON}}(t+1) &\sim \mathcal{B}\left(N_j^{\mathrm{OFF}}(t) - S_j^{\mathrm{ON}}(t+1), u^+\right), \\ C_{u,j}^{\mathrm{OFF}}(t+1) &\sim \mathcal{B}\left(N_j^{\mathrm{ON}}(t) - S_j^{\mathrm{OFF}}(t+1), u^-\right), \end{aligned} \tag{7}$$

where $u^+ := \max(u, 0)$ and $u^- = \max(-u, 0)$. Therefore, the number of ON TCLs given the input $u(t+1) = u$ is

$$\begin{aligned} \boldsymbol{N}_u^{\mathrm{ON}}(t+1) = \; & \boldsymbol{N}^{\mathrm{ON}}(t) + \boldsymbol{S}^{\mathrm{ON}}(t+1) \\ & - \boldsymbol{S}^{\mathrm{OFF}}(t+1) + \boldsymbol{C}_u^{\mathrm{ON}}(t+1) - \boldsymbol{C}_u^{\mathrm{OFF}}(t+1). \end{aligned} \tag{8}$$

Since the distributions of $C_{u,j}^{\mathrm{ON}}(t+1)$ and $C_{u,j}^{\mathrm{OFF}}(t+1)$ depend on u, the active and reactive power consumption across all nodes $\boldsymbol{P}(t+1)$ and $\boldsymbol{Q}(t+1)$ also depend on u. Therefore, from now on, we denote these random variables under the input $u(t+1) = u$ as $\boldsymbol{P}_u(t+1)$ and $\boldsymbol{Q}_u(t+1)$.

The next step is to model the one-step ahead voltage $V_j(t+1)$ at each node j as a random variable. Suppose that v_j is the voltage magnitude at node j; p_j^{b} and q_j^{b} are the active and reactive power flowing through the branch whose receiving end is node j; and the resistance and reactance of the branch are $r_j > 0$ and $x_j > 0$, respectively. Then, the DistFlow equations [31] corresponding to a single-phase equivalent model of a radial three-phase balanced network are

$$\begin{aligned} p_j^{\mathrm{b}} &= \sum_{k\in c(j)} p_k^{\mathrm{b}} + p_j + r_j |i_j^{\mathrm{b}}| \\ q_j^{\mathrm{b}} &= \sum_{k\in c(j)} q_k^{\mathrm{b}} + q_j + x_j |i_j^{\mathrm{b}}| \\ v_j^2 &= v_{e(j)}^2 - 2(r_j p_j^{\mathrm{b}} + x_j q_j^{\mathrm{b}}) + (r_j^2 + x_j^2)|i_j^{\mathrm{b}}|, \end{aligned} \tag{9}$$

where $e(j)$ and $c(j)$ are the parent node and set of child nodes of node j, respectively, and $|i_j^{\mathrm{b}}| = ((p_j^{\mathrm{b}})^2 + (q_j^{\mathrm{b}})^2)/v_{e(j)}^2$ is the square of the magnitude of the current flowing through the branch whose receiving end is node j. Given active and reactive power consumption $\boldsymbol{p}$ and $\boldsymbol{q} \in \mathbb{R}^n$ and substation voltage v_0, we let $f_{v_j}(\boldsymbol{p}, \boldsymbol{q}, v_0)$ be the voltage solution of (9), which can be obtained by various algorithms such as Backward-Forward Sweep [32]. Then, the one-step ahead voltage at node j under the input $u(t+1) = u$ is $V_{u,j}(t+1) = f_{v_j}(\boldsymbol{P}_u(t+1), \boldsymbol{Q}_u(t+1), v_0)$. Note that we cannot obtain an explicit pdf of $V_{u,j}(t+1)$ since there is no closed-form solution of f_{v_j}. Instead, we can obtain a realization of $V_{u,j}(t+1)$ by solving (9) for a set of realizations $\tilde{p}$ and $\tilde{q}$ of $\boldsymbol{P}_u(t+1)$ and $\boldsymbol{Q}_u(t+1)$.

Finally, we define a Bernoulli random variable that indicates whether or not an under-voltage violation exists,

$$X_u(t+1) = \mathbb{1}\left(\min_{j\in[n]} V_{u,j}(t+1) \geq \underline{v}\right), \tag{10}$$

whose success probability $\nu_u(t+1) = \Pr(X_u(t+1) = 1)$ corresponds to the one-step ahead probability of network safety under input $u(t+1) = u$. Thus, the utility's problem is to find a set $\mathcal{U}(t+1)$ such that, for any $u \in \mathcal{U}(t+1)$, $\nu_u(t+1)$ is larger than $1-\epsilon$ with confidence level over $1-\beta$. The solution to this problem is explained in the next section.

B. Probabilistically-Safe Set Construction

In this section, we first present a theorem on computing a confidence interval for the success probability of a Bernoulli random variable via a Monte Carlo simulation. Based on this theorem, we then show how the utility can test whether an input $u(t+1) = u$ is probabilistically safe and how this test procedure can be used to construct the set $\mathcal{U}(t+1)$ of all commands that satisfy the chance constraint.

Theorem 1: Suppose that $X^{(1)}, \ldots, X^{(n_{\mathrm{s}})}$ are i.i.d. samples of a random variable X following Bernoulli distribution $B(1, \nu)$ for a positive ν (i.e. $\Pr(X^{(i)} = 1) = \nu$, $\Pr(X^{(i)} = 0) = 1 - \nu$ for any $i \in [n_{\mathrm{s}}]$). Let $M_{n_{\mathrm{s}}} := \sum_{i=1}^{n_{\mathrm{s}}} X^{(i)}/n_{\mathrm{s}}$ be the estimator of ν, and $\tilde{m}_{n_{\mathrm{s}}}$ a realization of $M_{n_{\mathrm{s}}}$. If the following inequalities hold,

$$\tilde{m}_{n_{\mathrm{s}}} > 1 - \epsilon \tag{11}$$

$$n_{\mathrm{s}} > \ln\left(\frac{1}{\beta}\right) \frac{1}{(\tilde{m}_{n_{\mathrm{s}}} + \epsilon)\ln(\tilde{m}_{n_{\mathrm{s}}} + \epsilon) - (\tilde{m}_{n_{\mathrm{s}}} + \epsilon - 1)}, \tag{12}$$

then $[1-\epsilon, 1]$ is a confidence interval for the success probability ν of X with the confidence level over $1-\beta$.

The proof is given in Appendix A.. In our problem, $\tilde{m}_{n_{\mathrm{s}}}$ is a realization of an estimator of the success probability $\nu_u(t+1)$ obtained from realizations of $X_u(t+1)$. This theorem implies that, if both $\tilde{m}_{n_{\mathrm{s}}}$ and the number of samples n_{s} are sufficiently

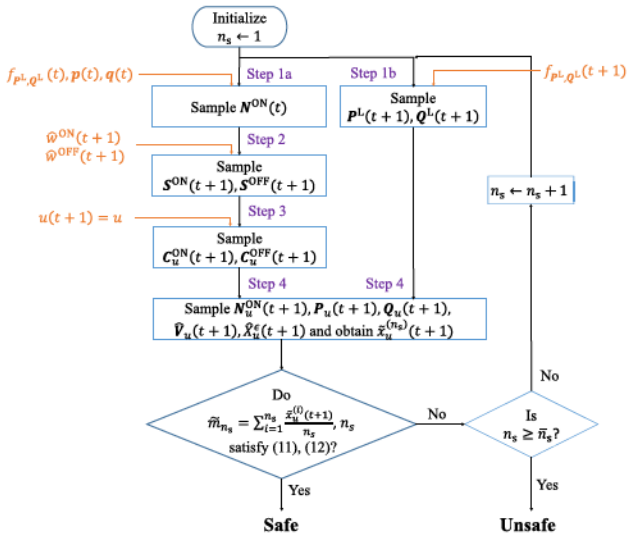

Fig. 2. Flowchart of the test procedure to check if an input $u(t+1) = u$ satisfies the chance constraint. The information required for each step is in orange.

large, then $\nu_u(t+1)$ is larger than $1-\epsilon$. Thus, to verify whether or not $\nu_u(t+1)$ is larger than $1-\epsilon$, the utility can obtain a number of realizations of $X_u(t+1)$ and check if inequalities (11) and (12) hold.

Now, we introduce the procedure the utility uses to obtain realizations of $X_u(t+1)$ given some $u \in [-1, 1]$. The utility first computes the probability mass function (pmf) of $\boldsymbol{N}^{\text{ON}}(t)$ given the observed $\boldsymbol{p}(t)$ and $\boldsymbol{q}(t)$ as follows,

$$\begin{aligned}
&\Pr\left(\boldsymbol{N}^{\text{ON}}(t) = \boldsymbol{n}^{\text{ON}} \mid (\boldsymbol{P}(t) = \boldsymbol{p}(t)) \cap (\boldsymbol{Q}(t) = \boldsymbol{q}(t))\right) \\
&= \Pr\left(\left(\boldsymbol{P}^{\text{L}}(t) = \boldsymbol{p}(t) - \Xi_p \boldsymbol{n}^{\text{ON}}\right) \cap \left(\boldsymbol{Q}^{\text{L}}(t) = \boldsymbol{q}(t) - \Xi_q \boldsymbol{n}^{\text{ON}}\right)\right. \\
&\qquad \left. \mid (\boldsymbol{P}(t) = \boldsymbol{p}(t)) \cap (\boldsymbol{Q}(t) = \boldsymbol{q}(t))\right) \\
&= \frac{f_{\boldsymbol{P}^{\text{L}},\boldsymbol{Q}^{\text{L}}}\left(\boldsymbol{p}(t) - \Xi_p \boldsymbol{n}^{\text{ON}}, \boldsymbol{q}(t) - \Xi_q \boldsymbol{n}^{\text{ON}}\right)}{\sum_{\boldsymbol{n} \in \mathbb{N}^{\text{ON}}} f_{\boldsymbol{P}^{\text{L}},\boldsymbol{Q}^{\text{L}}}\left(\boldsymbol{p}(t) - \Xi_p \boldsymbol{n}, \boldsymbol{q}(t) - \Xi_q \boldsymbol{n}\right)},
\end{aligned} \tag{13}$$

where $\mathbb{N}^{\text{ON}} := \{\boldsymbol{n}^{\text{ON}} \mid n_j^{\text{ON}} \in [n_j^{\text{TCL}}]_0 \ \forall j \in [n]\}$ is the set of all possible vectors for $\boldsymbol{N}^{\text{ON}}(t)$. Then, the utility obtains a realization $\tilde{x}_u(t+1)$ of $X_u(t+1)$ through the following sampling procedure, illustrated in Fig. 2.

1) a. Obtain a realization $\tilde{\boldsymbol{n}}^{\text{ON}}(t)$ of $\boldsymbol{N}^{\text{ON}}(t)$ by sampling from its pmf derived through (13), and compute $\tilde{\boldsymbol{n}}^{\text{OFF}}(t) = \boldsymbol{n}^{\text{TCL}} - \tilde{\boldsymbol{n}}^{\text{ON}}(t)$. b. Obtain realizations $\tilde{\boldsymbol{p}}^{\text{L}}(t+1)$ and $\tilde{\boldsymbol{q}}^{\text{L}}(t+1)$ of $\boldsymbol{P}^{\text{L}}(t+1)$ and $\boldsymbol{Q}^{\text{L}}(t+1)$ by sampling from $f_{\boldsymbol{P}^{\text{L}},\boldsymbol{Q}^{\text{L}}}(t+1)$.
2) Obtain realizations $\tilde{\boldsymbol{s}}^{\text{ON}}$ and $\tilde{\boldsymbol{s}}^{\text{OFF}}$ of $\boldsymbol{S}^{\text{ON}}(t+1)$ and $\boldsymbol{S}^{\text{OFF}}(t+1)$ by computing their elements per (6) as

$$\begin{aligned}
\tilde{s}_j^{\text{ON}}(t+1) &= \hat{w}_j^{\text{ON}}(t+1)\tilde{n}_j^{\text{OFF}}(t) \ \ \forall j \in [n] \\
\tilde{s}_j^{\text{OFF}}(t+1) &= \hat{w}_j^{\text{OFF}}(t+1)\tilde{n}_j^{\text{ON}}(t) \ \ \forall j \in [n].
\end{aligned}$$

3) Obtain realizations $\tilde{\boldsymbol{c}}_u^{\text{ON}}(t+1)$ and $\tilde{\boldsymbol{c}}_u^{\text{OFF}}(t+1)$ of $\boldsymbol{C}_u^{\text{ON}}(t+1)$ and $\boldsymbol{C}_u^{\text{OFF}}(t+1)$ by sampling their elements per (7) from the binomial distributions $\mathcal{B}(\tilde{n}_j^{\text{OFF}}(t) - \tilde{s}_j^{\text{ON}}(t+1), u^+)$ and $\mathcal{B}(\tilde{n}_j^{\text{ON}}(t) - \tilde{s}_j^{\text{OFF}}(t+1), u^-)$.
4) Obtain realizations of $\boldsymbol{N}_u^{\text{ON}}(t+1)$, $\boldsymbol{P}_u(t+1)$, $\boldsymbol{Q}_u(t+1)$, $\boldsymbol{V}_u(t+1)$, and $X_u(t+1)$ as

$$\begin{aligned}
\tilde{\boldsymbol{n}}_u^{\text{ON}}(t+1) &= \tilde{\boldsymbol{n}}^{\text{ON}}(t) + \tilde{\boldsymbol{s}}^{\text{ON}}(t+1) - \tilde{\boldsymbol{s}}^{\text{OFF}}(t+1) \\
&\quad + \tilde{\boldsymbol{c}}_u^{\text{ON}}(t+1) - \tilde{\boldsymbol{c}}_u^{\text{OFF}}(t+1) \\
\tilde{\boldsymbol{p}}_u(t+1) &= \tilde{\boldsymbol{p}}^{\text{L}}(t+1) + \Xi_p \tilde{\boldsymbol{n}}_u^{\text{ON}}(t+1) \\
\tilde{\boldsymbol{q}}_u(t+1) &= \tilde{\boldsymbol{q}}^{\text{L}}(t+1) + \Xi_q \tilde{\boldsymbol{n}}_u^{\text{ON}}(t+1) \\
\tilde{v}_{u,j}(t+1) &= f_{v_j}(\tilde{\boldsymbol{p}}_u(t+1), \tilde{\boldsymbol{q}}_u(t+1), v_0) \ \ \forall j \in [n] \\
\tilde{x}_u(t+1) &= \mathbb{1}(\min_{j \in [n]} \tilde{v}_{u,j}(t+1) \geq \underline{v}).
\end{aligned}$$

The utility can obtain multiple realizations of $X_u(t+1)$ by iterating this sampling procedure. Denote each realization i of $X_u(t+1)$ as $\tilde{x}_u^{(i)}(t+1)$, where $i \in [n_s]$. In each iteration, the utility updates the realization of the estimator $\tilde{m}_{n_s} = \sum_{i=1}^{n_s} \tilde{x}_u^{(i)}(t+1)/n_s$ and checks if the inequalities (11), (12) hold. If they do, $u(t+1) = u$ satisfies the chance constraint with confidence level over $1-\beta$; otherwise, the utility continues to iterate until n_s reaches some pre-determined upper bound $\overline{n}_s$, as shown in Fig. 2. We note that an input $u(t+1) = u$ that passes the test procedure ensures satisfaction of chance-constraint (1) with a confidence level over $1-\beta$ even in the presence of random responses from TCLs under input u, since those random responses are also sampled in Step 3.

Next, we construct a one-step ahead constraint set $\mathcal{U}(t+1)$. We first make an assumption on the monotonicity of $\nu_u(t+1)$.

Assumption 1: The one-step ahead probability of network safety $\nu_u(t+1)$ monotonically decreases with respect to u.

The intuition behind this assumption is that the active and reactive power consumption at each node is likely to increase as u increases, which is also likely to lead to a voltage decrease at every node. We acknowledge that Assumption 1 may not hold for every distribution network under every loading condition. In Appendix B., we justify this assumption by showing that an approximation of $\nu_u(t+1)$ (derived from LinDistFlow [33]) is a monotonically decreasing function with respect to u. We also demonstrate monotonicity empirically using DistFlow. Furthermore, in Appendix C., we propose an alternative method that does not rely on Assumption 1, but is overly conservative.

Under Assumption 1, the following holds.

Theorem 2: Suppose that Assumption 1 holds and let $\tilde{x}_{\overline{u}}^{(1)}(t+1), \ldots, \tilde{x}_{\overline{u}}^{(n_s)}(t+1)$ be n_s realizations of $X_{\overline{u}}(t+1)$ for an input $\overline{u} \in [-1, 1]$. If n_s and $\tilde{m}_{n_s} = \sum_{i=1}^{n_s} \tilde{x}_{\overline{u}}^{(i)}(t+1)/n_s$ satisfy (11) and (12), then $\mathcal{U}(t+1) = [-1, \overline{u}]$ is a solution to the Problem 1 with confidence level over $1-\beta$.

Proof: By Theorem 1, the interval $[1-\epsilon, 1]$ is a confidence interval for $\nu_{\overline{u}}(t+1)$ with confidence level over $1-\beta$. Also, under Assumption 1, $\nu_u(t+1) \geq \nu_{\overline{u}}(t+1)$ holds for any $u \in \mathcal{U}(t+1) = [-1, \overline{u}]$. Thus, $\nu_u(t+1)$ is greater than or equal to $1-\epsilon$ for any $u \in \mathcal{U}(t+1)$ with confidence level over $1-\beta$. □

This theorem means that, if the probability of network safety under the input $u(t+1) = \overline{u}$ is greater than or equal to the desired safety probability, then any less aggressive input, i.e.,

one in the range $[-1, \overline{u}]$, also satisfies the chance constraint. Therefore, a solution to Problem 1 is the interval $[-1, \overline{u}]$, where $\overline{u}$ passes the test procedure in Fig. 2.

The choice of probabilistically-safe set $\mathcal{U}(t+1)$ is not unique. A larger $\mathcal{U}(t+1)$ gives more flexibility to the aggregator, potentially improves the quality of balancing services, and reduces the conservativeness of our approach. Therefore, the utility should find the largest possible $\overline{u}$ that passes the test procedure. This can be achieved using the bisection method [34], starting with $\overline{u} = 1$.

Remark 1: To restrict the probability of over-voltage violations, we can also apply the monotonicity assumption; the probability of over-voltage violations increases as the input u decreases. In this case, we can use the bisection method to obtain a lower bound on $u(t+1)$. Then, the utility can send both a lower and upper bound on $u(t+1)$ to restrict the probability of over- and under-voltage violations.

Remark 2: Since the utility approximates P^{T} and Q^{T} in (5) and uses estimates of w^{ON} and w^{OFF} in Step 2 of the sampling procedure, Theorem 2 holds only if those approximations/estimates are accurate. We will justify the use of these approximations/estimations through simulation in Section V.

V. Case Study

We next present the result of a case study in which we compare the proposed approach with two benchmark approaches. We examine the reference tracking and safety performance of each algorithm under scenarios in which the balancing service capacity committed by the aggregator must be reduced to preserve network safety during a period with large differences between real-time network conditions and their forecasts.

The proposed algorithm is benchmarked against two algorithms: a tracking controller benchmark and a robust OPF benchmark. The tracking controller benchmark does not take into account network safety. It picks the optimal input $u_{\mathrm{opt}}(t)$ using (4) with $\mathcal{U}(t) = [-1, 1]$, where $\mathbb{E}[P_{\mathrm{agg}}(t)]$ is the expected aggregate power of the TCLs under $u(t) = u$, which is computed with the same approximate aggregate TCL model. This benchmark aims to verify the occurrence of significant under-voltage violations when the aggregate power consumption of the TCLs closely tracks the regulation signal, which demonstrates the need for a network-safe approach.

The robust OPF benchmark approximately enforces network safety assuming linearized power flow. It solves the following mixed integer linear program at each time step to compute the optimal one-step ahead mode of each TCL,

$$\min_{m^i} \ |p_{\mathrm{agg}} - p_{\mathrm{ref}}| \tag{14a}$$

$$\text{s.t. } p_{\mathrm{agg}} = \sum_{i=1}^{n^{\mathrm{TCL}}} p^i m^i \tag{14b}$$

$$p_j^{\mathrm{T}} = \sum_{i \in \mathcal{I}_j} p^i m^i, \ \ q_j^{\mathrm{T}} = \sum_{i \in \mathcal{I}_j} q^i m^i, \qquad \forall j \in [n] \tag{14c}$$

$$\text{TCL temperature dynamics (2)}, \qquad \forall i \in [n^{\mathrm{TCL}}] \tag{14d}$$

$$\theta^i \in [\underline{\theta^i}, \overline{\theta^i}], \qquad \forall i \in [n^{\mathrm{TCL}}] \tag{14e}$$

$$v = \Phi_p(p^{\mathrm{T}} + p^{\mathrm{Lmax}}) + \Phi_q(q^{\mathrm{T}} + q^{\mathrm{Lmax}}) + \Phi_c \tag{14f}$$

$$\underline{v} \le v, \tag{14g}$$

where $\mathcal{I}_j$ is the set of indices of TCLs connected to j and (14f) is the linearized power flow developed in [1] under the maximum active and reactive power consumption of the non-participating loads, which makes its solution robust to load forecast error. This benchmark aims to produce "best-case" results but is inconsistent with the coordination scheme in Fig. 1; it requires the optimizing entity to have full information from both the utility and the aggregator (e.g., this would be possible if the utility were the aggregator). Incorporating accurate one-step ahead power consumption forecasts from the non-participating loads and using the nonlinear AC power flow equations could potentially yield less conservative and more accurate "best-case" solutions. However, this is unrealistic as load forecasts are imperfect and using the nonlinear AC power flow equations leads to a non-convex mixed-integer nonlinear programming problem that we have found exceedingly challenging to solve with modern solvers.

The experiments were carried out on two radial distribution networks:

1) A 56-bus balanced distribution network from [35] derived from the IEEE 123-node test feeder [36]
2) A 116-bus balanced distribution network from [37] that is an electrically-equivalent model of the IEEE European Low Voltage Test Feeder [38].

Section V-A presents the results for the 56-node network and Section V-B presents the results for the 116-node network.

A. Results for the 56-Node Network

For the 56-node network, we assume the aggregator commits ± 0.91 MW of frequency regulation capacity to the ISO anticipating nominal network loading (2.62 MW, which would enable TCLs to deliver their full frequency regulation capacity); however, real-time network loading is much higher than usual (3.78 MW on average) and if the aggregation were to closely track the frequency regulation signal it would cause under-voltage violations. We first describe our simulation setup and detail the benchmark approaches. Then, we present our results.

We denote the nominal active and reactive power consumption at node j by p_j^{Ln} and q_j^{Ln}, respectively. We set the safe lower bound on the voltage to $\underline{v} = 0.95$ pu. TCL parameters are randomly sampled,[1] and the TCLs are distributed throughout the network so that the aggregate TCLs' nominal active power consumption at node j is approximately $0.25 p_j^{\mathrm{Ln}}$. For simplicity, we assume that the active and reactive power consumption of the non-participating

[1]Each parameter is sampled from uniform distributions with intervals: $\theta_a^i \in [29, 31]$ °C, $c_{\mathrm{th}}^i \in [1.5, 2.5]$kWh/°C, $r_{\mathrm{th}}^i \in [1.2, 2.5]$°C/kW, $p_{\mathrm{tr}}^i \in [-18, -14]$ kW, $\zeta^i \in [2.3, 2.7]$, $\theta_s^i \in [20, 25]$°C, $\overline{\theta}^i - \underline{\theta}^i \in [1.5, 2]$°C, and $\omega^i = \tan(\arccos(\phi^i))$, where $\phi^i \in [0.95, 0.99]$.

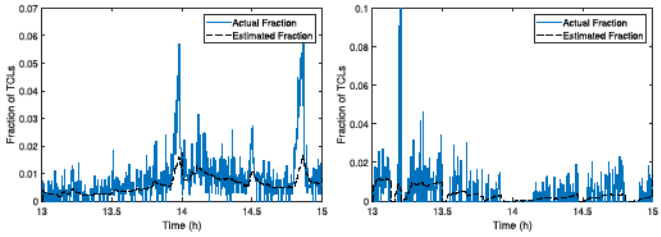

Fig. 3. Actual and estimated fractions of TCLs switched ON (left) and OFF (right) by their thermostats, in the proposed approach ($\epsilon = 0.02$), on the 56-node network.

TABLE I
REFERENCE TRACKING AND SAFETY PERFORMANCE OF EACH ALGORITHM ON THE 56-NODE NETWORK

	Track Ctrl Benchmark	Robust OPF Benchmark	Proposed Approach $\epsilon = 0.05$	Proposed Approach $\epsilon = 0.02$
RMSE (kW)	77.05	168.3	102.8	118.8
Safety Probability	0.908	1.00	0.981	0.986

loads at each node $P_j^{\mathrm{L}}(t)$ and $Q_j^{\mathrm{L}}(t)$ follow normal distributions $\mathcal{N}(\overline{p}_j^{\mathrm{Ln}}(t), (0.15p_j^{\mathrm{Ln}})^2)$ and $\mathcal{N}(\overline{q}_j^{\mathrm{Ln}}(t), (0.15q_j^{\mathrm{Ln}})^2)$ truncated by the intervals $[p_j^{\mathrm{Lmin}}, p_j^{\mathrm{Lmax}}] = [-0.25p_j^{\mathrm{Ln}}, 0.675p_j^{\mathrm{Ln}}]$ and $[q_j^{\mathrm{Lmin}}, q_j^{\mathrm{Lmax}}] = [-0.25q_j^{\mathrm{Ln}}, 0.675q_j^{\mathrm{Ln}}]$, respectively. We conduct 2 h simulations (13h-15 h) and let $\overline{p}_j^{\mathrm{Ln}}(t)$ and $\overline{q}_j^{\mathrm{Ln}}(t)$ linearly increase from 0.5 to 0.65 of their nominal values from 13.0 h to 13.9 h, stay constant from 13.9 h to 14.1 h, and linearly decrease to 0.5 of their nominal values from 14.1 h to 15.0 h. The reference signal $p_{\mathrm{ref}}(t)$ is a scaled and shifted 2 h segment of a PJM RegD signal [39]. We use the desired safety probabilities $1 - \epsilon = 0.95$ and 0.98, the desired confidence level $1 - \beta = 0.999$, and the upper bound on the number of samples $\overline{n}_{\mathrm{s}} = 140,000$. The aggregator obtains the estimates $\hat{w}_j^{\mathrm{ON}}(t+1)$ and $\hat{w}_j^{\mathrm{OFF}}(t+1)$ for each node leveraging an approximate model of the dynamics of the TCL aggregation. The model was developed in past work, e.g., [6], and so not detailed here. While we could identify different models for each node, here we use the same model for each node $j \in [n]$ and so $\hat{w}_j^{\mathrm{ON}}(t+1)$ and $\hat{w}_j^{\mathrm{OFF}}(t+1)$ are identical across nodes. Fig. 3 demonstrates the model's estimation performance, showing the actual and estimated fractions of TCLs outside of their dead-bands. Although the estimates do not perfectly track the actual values, they capture the overall trends.

Fig. 4 illustrates the results of the comparison between the two benchmarks and our proposed approach with $\epsilon = 0.05$ and 0.02. Table I shows the root mean squared error (RMSE) of the aggregate power from the reference signal, along with the empirical safety probability computed as the fraction of time steps in which under-voltage violations (computed with the nonlinear power flow equations) do not happen. The tracking controller benchmark has the best tracking performance, but frequently causes under-voltage violations. This demonstrates the need to employ network-safe DER control strategies. In contrast, the OPF benchmark avoids under-voltage violations, but has the worst tracking performance, demonstrating that approaches that (approximately) enforce network safety will at times have poor balancing service performance.

TABLE II
REFERENCE TRACKING AND SAFETY PERFORMANCE OF EACH ALGORITHM ON THE 116-NODE NETWORK

	Track Ctrl Benchmark	Robust OPF Benchmark	Proposed Approach $\epsilon = 0.05$
RMSE (kW)	28.51	71.14	48.26
Safety Probability	0.891	1.00	0.991

Our approach achieves a better trade-off between tracking performance and network safety; specifically, it achieves better tracking performance than the OPF benchmark and satisfies the chance constraint on network safety, resulting in fewer under-voltage violations than the tracking controller benchmark. As shown in Table I, the empirical safety probabilities are over the target values $1 - \epsilon$. The RMSE increases as ϵ decreases, which is expected since higher $1 - \epsilon$ results in more conservative bounds on the input commands. The results also verify that the model error shown in Fig. 4 does not degrade the tracking results in a significant way.

The average computation time for constraint set construction was 33 s, which is longer than the length of each time step $\Delta t = 10$ s. However, the algorithm implementation was not optimized for computational efficiency, which means the computation time could be reduced. In addition, the sampling procedure could be efficiently executed in parallel across multiple processors, allowing the approach to achieve the required computation time. For larger networks, the computational requirements of Step 4 of the test procedure (i.e., solving the DistFlow equations via e.g., Forward-Backward Sweep) would increase proportionally to the number of nodes in the network. This means that parallelization would be key to ensuring applicability in real distribution networks.

B. Results for the 116-Node Network

For the 116-node network, we assume the aggregator commits ± 0.20 MW of capacity. The real-time network loading is 1.29 MW on average exceeding the expected network loading of 0.98 MW. We set the lower bound on the voltage to $\underline{v} = 0.97$ pu and assume that $P_j^{\mathrm{L}}(t)$ and $Q_j^{\mathrm{L}}(t)$ follow normal distributions $\mathcal{N}(p_j^{\mathrm{Ln}}, (0.1p_j^{\mathrm{Ln}})^2)$ and $\mathcal{N}(q_j^{\mathrm{Ln}}, (0.1q_j^{\mathrm{Ln}})^2)$ truncated by the intervals $[p_j^{\mathrm{Lmin}}, p_j^{\mathrm{Lmax}}] = [0.3p_j^{\mathrm{Ln}}, 1.6p_j^{\mathrm{Ln}}]$ and $[q_j^{\mathrm{Lmin}}, q_j^{\mathrm{Lmax}}] = [0.3q_j^{\mathrm{Ln}}, 1.6q_j^{\mathrm{Ln}}]$, respectively. The desired safety probability is $1 - \epsilon = 0.95$ with desired confidence level $1 - \beta = 0.999$. The other settings including the TCLs' parameters and regulation signal are the same as in Section V-A.

Table II shows the result of the experiment. Similar to the results of the 56-node network, these results show that our approach satisfies the chance constraint, unlikely the tracking controller benchmark, while achieving better tracking performance than the robust OPF benchmark. The average computation time for constraint set construction was 90 s.

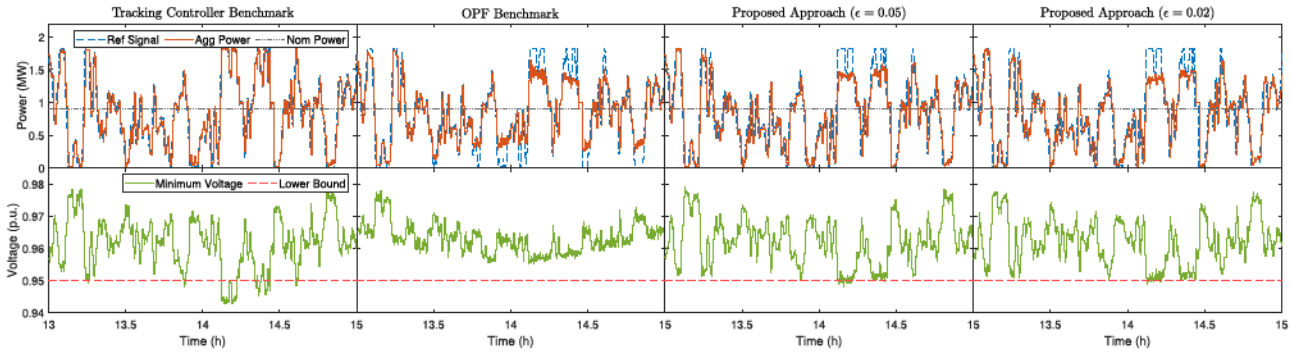

Fig. 4. Reference signal and the TCLs' aggregate power (top), and the minimum network voltage and the safe lower bound (bottom) for each algorithm on the 56-node network.

VI. Conclusion

This article proposed an approach to coordinate a collection of TCLs to provide balancing services while guaranteeing network safety with high probability. In particular, we proposed a constraint construction method that would allow the utility to constrain the input commands of an aggregator providing balancing services like frequency regulation. The approach imposes a chance constraint on network safety, wherein both the violation probability and confidence level are design parameters that can be selected by the utility. We used the bisection method to compute the largest possible constraint set, which provides the most flexibility to the aggregator.

Future work will extend the proposed approach to incorporate different types of DERs, such as stationary batteries, electric vehicles, and curtailable solar photovoltaics, into the framework. We already have some preliminary work along this direction; specifically, [40] proposes a method to extend the proposed approach to include battery energy storage. In addition, we will also explore methods to improve the computational performance of our approach and test the approach on realistic/real networks. Furthermore, we plan to compare our proposed utility-aggregator coordination strategy with other coordination approaches that achieve network awareness such as the use of operating envelopes [41].

Appendix

A. Proof of Theorem 1

Proof: By Theorem 4.1 in [42], the following inequality is derived from the Chernoff bound for any $0 < \delta \leq \frac{1-\nu}{\nu}$,

$$\Pr(M_{n_s} \geq (1+\delta)\nu) \leq \left(\frac{1}{1+\delta}\right)^{(1+\delta)n_s\nu} e^{\delta n_s \nu} = e^{n_s\nu(\delta-(1+\delta)\ln(1+\delta))}. \quad (15)$$

We substitute c/ν, with $c \in [0, 1-\nu]$, for δ and obtain

$$\Pr\left(M_{n_s} - \nu \geq c\right) \leq e^{n_s\left(c-(\nu+c)\ln\left(1+\frac{c}{\nu}\right)\right)}$$
$$\iff \Pr\left(\nu \geq M_{n_s} - c\right) \geq 1 - e^{n_s\left(c-(\nu+c)\ln\left(1+\frac{c}{\nu}\right)\right)}. \quad (16)$$

Hence, $[\tilde{m}_{n_s} - c, 1]$ is a confidence interval for ν with confidence level over $1 - e^{n_s(c-(\nu+c)\ln(1+\frac{c}{\nu}))}$. Thus, if there exists $c > 0$ that satisfies $\tilde{m}_{n_s} - c \geq 1 - \epsilon$ and $1 - e^{n_s(c-(\nu+c)\ln(1+\frac{c}{\nu}))} > 1 - \beta$, then $[1-\epsilon, 1]$ is a confidence interval for ν with confidence level over $1-\beta$. Next, we show that such a c exists. First, we derive a lower bound on $1 - e^{n_s(c-(\nu+c)\ln(1+\frac{c}{\nu}))}$. Focusing on the exponent, observe that

$$\frac{\partial}{\partial \nu}\left(c - (\nu+c)\ln\left(1+\frac{c}{\nu}\right)\right) = -\ln\left(1+\frac{c}{\nu}\right) + \frac{c}{\nu}. \quad (17)$$

If we let $h_1(x) := -\ln(1+x) + x$, the right side of (17) is equal to $h_1(c/\nu)$. From $h_1(0) = 0$ and $\partial h_1(x)/\partial x = -1/(1+x) + 1 \geq 0 \quad \forall x \in [0, \infty)$, we have $h_1(x) \geq 0$ for all $x \in [0, \infty)$, which means $h_1(c/\nu)$ is non-negative. Hence, the exponent is increasing with respect to ν, and thus achieves its maximum at $\nu = 1$. Therefore,

$$\Pr\left(\nu \geq M_{n_s} - c\right) \geq 1 - e^{n_s(c-(1+c)\ln(1+c))}. \quad (18)$$

Since $\nu \leq 1$, (18) implies that $[\tilde{m}_{n_s} - c, 1]$ is a confidence interval for ν with confidence level over $1 - e^{n_s(c-(1+c)\ln(1+c))}$.

Now, suppose that (11), (12) hold and define $h_2(x) := x - (1+x)\ln(1+x)$; the exponent on the right side of (18) is $n_s h_2(c)$. From $h_2(0) = 0$ and $\partial h_2(x)/\partial x < 0$ for all $x \in (0, \infty)$, we have $h_2(x) < 0$ for all $x \in (0, \infty)$. Since $\tilde{m}_{n_s} - (1-\epsilon) > 0$ by (11), $(\tilde{m}_{n_s} + \epsilon - 1) - (\tilde{m}_{n_s} + \epsilon)\ln(\tilde{m}_{n_s} + \epsilon) = h_2(\tilde{m}_{n_s} - (1-\epsilon))$ is negative. Also, substituting c with $\tilde{m}_{n_s} - (1-\epsilon)$, the right side of (18) becomes $1 - e^{n_s((\tilde{m}_{n_s}+\epsilon-1)-(\tilde{m}_{n_s}+\epsilon)\ln(\tilde{m}_{n_s}+\epsilon))}$ which is less than $1 - e^{-\ln(\frac{1}{\beta})} = 1 - \beta$, per (12). Hence, $1 - e^{n_s(c-(1+c)\ln(1+c))} \geq 1 - \beta$ and, thus, the interval $[1-\epsilon, 1] = [\tilde{m}_{n_s} - c, 1]$ is a confidence interval for ν with confidence level over $1-\beta$. □

B. Justification of Assumption 1

In this section, we justify Assumption 1 by showing that an approximation of $\nu_u(t+1)$ is a monotonically decreasing function with respect to u. We consider the LinDistFlow equations [33],

which drop the nonlinear terms of (9), i.e.,

$$\hat{p}_j^{\mathrm{b}} = \sum_{k\in c(j)} \hat{p}_k^{\mathrm{b}} + p_j, \quad \hat{q}_j^{\mathrm{b}} = \sum_{k\in c(j)} \hat{q}_k^{\mathrm{b}} + q_j$$
$$\hat{v}_j^2 = \hat{v}_{e(j)}^2 - 2(r_j\hat{p}_j^{\mathrm{b}} + x_j\hat{q}_j^{\mathrm{b}}), \tag{19}$$

where variables with hats correspond to approximations of the original DistFlow variables. Let $\hat{f}_{v_j}(\boldsymbol{p},\boldsymbol{q},v_0)$ be the voltage solution of (19), i.e., $\hat{V}_{u,j}(t+1) := \hat{f}_{v_j}(\boldsymbol{P}_u(t+1), \boldsymbol{Q}_u(t+1), v_0)$ is the approximate voltage at node j. Also, let

$$\hat{X}_u(t+1) = \mathbb{1}\left(\min_{j\in[n]} \hat{V}_{u,j}(t+1) \geq \underline{v}\right) \tag{20}$$

whose success probability $\hat{\nu}_u(t+1) := \Pr(\hat{X}_u(t+1)=1)$ approximates $\nu_u(t+1)$. To show that $\hat{\nu}_u(t+1)$ is decreasing with respect to u, we start with a proposition.

Proposition 1: Suppose that $\boldsymbol{p}^{(1)}, \boldsymbol{p}^{(2)} \in \mathbb{R}^n$ and $\boldsymbol{q}^{(1)}, \boldsymbol{q}^{(2)} \in \mathbb{R}^n$ are different instances of active and reactive power consumption where $p_j^{(1)} \leq p_j^{(2)}$ and $q_j^{(1)} \leq q_j^{(2)}\, \forall\, j \in [n]$. Then, $\hat{f}_{v_j}(\boldsymbol{p}^{(1)},\boldsymbol{q}^{(1)},v_0) \geq \hat{f}_{v_j}(\boldsymbol{p}^{(2)},\boldsymbol{q}^{(2)},v_0)$ for all $j \in [n]$.

Proof: First let $\hat{f}_{p_j^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0)$ and $\hat{f}_{q_j^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0)$ be the solutions of (19) corresponding to p_j^{b} and q_j^{b} when the active and reactive power consumption at each node are $\boldsymbol{p}$ and $\boldsymbol{q}$, and the substation voltage is v_0. Then, for all $j \in [n]$ [33]

$$\hat{f}_{p_j^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0) = \sum_{k\in d(j)} p_k, \quad \hat{f}_{q_j^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0) = \sum_{k\in d(j)} q_k$$
$$\hat{f}_{v_j}^2(\boldsymbol{p},\boldsymbol{q},v_0) = v_0^2 - 2\sum_{k\in a(j)} \Big(r_k \hat{f}_{p_k^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0) + x_k \hat{f}_{q_k^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0)\Big), \tag{21}$$

where $d(j) := c(j) \cup \{j\}$ is the set of indices of all descendants of node j including itself, and $a(j)$ is the set of indices of all ancestors of node j including itself. Hence, $\hat{f}_{p_j^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0)$ and $\hat{f}_{q_j^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0)$ are increasing as p_k and q_k increase for all $k \in [n]$, and $\hat{f}_{p_j^{\mathrm{b}}}(\boldsymbol{p}^{(1)},\boldsymbol{q}^{(1)},v_0) \leq \hat{f}_{p_j^{\mathrm{b}}}(\boldsymbol{p}^{(2)},\boldsymbol{q}^{(2)},v_0)$ and $\hat{f}_{q_j^{\mathrm{b}}}(\boldsymbol{p}^{(1)},\boldsymbol{q}^{(1)},v_0) \leq \hat{f}_{q_j^{\mathrm{b}}}(\boldsymbol{p}^{(2)},\boldsymbol{q}^{(2)},v_0)$ for all $j \in [n]$. Also, since all r_k and x_k are positive, $\hat{f}_{v_j}(\boldsymbol{p},\boldsymbol{q},v_0)$ is decreasing as $\hat{f}_{p_k^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0)$ and $\hat{f}_{q_k^{\mathrm{b}}}(\boldsymbol{p},\boldsymbol{q},v_0)$ increase for all $k \in [n]$. Therefore, $\hat{f}_{v_j}(\boldsymbol{p}^{(1)},\boldsymbol{q}^{(1)},v_0) \geq \hat{f}_{v_j}(\boldsymbol{p}^{(2)},\boldsymbol{q}^{(2)},v_0)\, \forall\, j \in [n]$. □

This proposition states that $\hat{f}_{v_j}(\boldsymbol{p},\boldsymbol{q},v_0)$ monotonically decreases as the active and reactive power consumption p_j and q_j at every node increase for all $j \in [n]$. Since the one-step ahead active and reactive power consumption of the TCLs at each node are likely to increase as u increases (recall that in Section III we made the realistic assumption that TCLs have constant lagging power factors, and so their active and reactive power consumption change in the same direction), this proposition implies that the probability of under-voltage violations increases as u increases. This is stated in the following theorem.

Theorem 3: The approximate probability of network safety $\hat{\nu}_u(t+1)$ under the one-step ahead input u is a monotonically decreasing function of u.

To prove Theorem 3, we first introduce and prove Lemma 1, which is required for the proof of Lemma 2. Then, we prove Lemma 2, which is used in the proof of Theorem 3. Finally, we prove Theorem 3.

Lemma 1: Suppose that $a_w(x), b_w(x) : \mathcal{X} \to \mathbb{R}^+$ are non-negative functions with parameter $w \in \mathbb{R}$, and $\{\tilde{x}_1,\ldots,\tilde{x}_N\}$ $(\tilde{x}_1 \leq \ldots \leq \tilde{x}_N)$ is a finite subset of the domain $\mathcal{X}$. Also, assume that the following two conditions hold: 1) $\sum_{k=1}^j a_w(\tilde{x}_k)$ is a decreasing function with respect to w for any $j \in \{1,\ldots,N\}$, and 2) $b_w(x)$ is decreasing function with respect to both x and w. Then, $g(w) := \sum_{k=1}^N a_w(\tilde{x}_k) b_w(\tilde{x}_k)$ is a decreasing function with respect to w.

Proof: We prove the lemma by showing that, for $w \leq \overline{w}$, $\sum_{k=1}^j a_w(\tilde{x}_k) b_w(\tilde{x}_k) \geq \sum_{k=1}^j a_{\overline{w}}(\tilde{x}_k) b_{\overline{w}}(\tilde{x}_k)$ for any $j \in [N]$ and $w_1, w_2 \in \mathbb{R}$ as follows:

$$\sum_{k=1}^j a_w(\tilde{x}_k) b_w(\tilde{x}_k) \geq \sum_{k=1}^j a_w(\tilde{x}_k) b_{\overline{w}}(\tilde{x}_k) \tag{22a}$$
$$= b_{\overline{w}}(\tilde{x}_j)\sum_{k=1}^j a_w(\tilde{x}_k) + \sum_{k=1}^{j-1} \Delta b_{\overline{w}}(\tilde{x}_k)\sum_{l=1}^k a_w(\tilde{x}_l) \tag{22b}$$
$$\geq b_{\overline{w}}(\tilde{x}_j)\sum_{k=1}^j a_{\overline{w}}(\tilde{x}_k) + \sum_{k=1}^{j-1} \Delta b_{\overline{w}}(\tilde{x}_k)\sum_{l=1}^k a_{\overline{w}}(\tilde{x}_l) \tag{22c}$$
$$= \sum_{k=1}^j a_{\overline{w}}(\tilde{x}_k) b_{\overline{w}}(\tilde{x}_k) \tag{22d}$$

where $\Delta b_{\overline{w}}(\tilde{x}_k) := (b_{\overline{w}}(\tilde{x}_k) - b_{\overline{w}}(\tilde{x}_{k+1}))$, (22a) holds by condition 2 and (22c) holds by condition 1. □

Lemma 2: Suppose that $Y_w^{(j)}$ $(j \in [n])$ is a discrete random variable with the finite sample space $\mathcal{Y}^{(j)} = \{\acute{y}_1^j,\ldots,\acute{y}_{\kappa_j}^j\}$ $(\acute{y}_1^j \leq \ldots \leq \acute{y}_{\kappa_j}^j)$ with parameter $w \in \mathbb{R}$ having the following properties: 1) $Y_w^{(1)},\ldots,Y_w^{(n)}$ are independent of each other, and 2) the cdf $F_{Y^{(j)}}(y;w)$ of $Y_w^{(j)}$ is a decreasing function with respect to w for any $y \in \mathcal{Y}^{(j)}$. Then, for any $\overline{z}^{(i)} \in \mathbb{R}\,(i \in [n_c])$ and non-negative coefficients $a_{ij} \in \mathbb{R}^+$, $\Pr(\bigwedge_{i=1}^{n_c}(\sum_{j=1}^n a_{ij} Y_w^{(j)} \leq \overline{z}^{(i)}))$ monotonically decreases as w increases.

Proof: Let $\boldsymbol{Y}_w = (Y_w^{(1)},\ldots,Y_w^{(n)})^\top$ be a multivariate random variable with elements $Y_w^{(j)}$ and $\mathcal{P} = \{\boldsymbol{y} \mid A\boldsymbol{y} \leq \overline{\boldsymbol{z}}\}$ be a polyhedron with elements a_{ij}. Then,

$$\Pr\left(\bigwedge_{i=1}^{n_c}\left(\sum_{j=1}^n a_{ij} Y_w^{(j)} \leq \overline{z}^{(i)}\right)\right) = \Pr\left(\boldsymbol{Y}_w \in \mathcal{P}\right).$$

Note that $\mathcal{P}$ is a lower polyhedron in $\Pi_{j=1}^n[\acute{y}_1^j, \acute{y}_{\kappa_j}^j]$; if $\boldsymbol{y} \in \mathcal{P}$, then $\boldsymbol{y}' \in \mathcal{P}$ also holds for any $\boldsymbol{y}' \leq \boldsymbol{y}$. Thus, it is sufficient to show that $\Pr(\boldsymbol{Y}_{w_1} \in \mathcal{P}') \geq \Pr(\boldsymbol{Y}_{w_2} \in \mathcal{P}')\, \forall\, w_1 \geq w_2$ and any lower polyhedron $\mathcal{P}'$, which we do as follows:

1) Let $n = 1$ and $\mathcal{P}_1' \subset [\acute{y}_1^1, \acute{y}_{\kappa_1}^1]$ be a 1-dimensional lower polyhedron. Then, there exists $\overline{y}$ such that $\mathcal{P}_1' = [\acute{y}_1^1, \overline{y}]$, and $\Pr(Y_{w_1}^{(1)} \in \mathcal{P}_1') = F_{Y^{(1)}}(\overline{y};w_1) \geq F_{Y^{(1)}}(\overline{y};w_2) = \Pr(Y_{w_2}^{(1)} \in \mathcal{P}_1')$, which proves the statement for $n = 1$.

2) Let $n = k$ and suppose $\Pr(\boldsymbol{Y}_{w_1}^{(1:k)} \in \mathcal{P}'_k) \geq \Pr(\boldsymbol{Y}_{w_2}^{(1:k)} \in \mathcal{P}'_k)$ holds $\forall w_1 \geq w_2$ and for any k-dimensional lower polyhedron $\mathcal{P}'_k \subset \Pi_{j=1}^{k}[\acute{y}_1^j, \acute{y}_{\kappa_j}^j]$. Define $\mathcal{P}_k^-(y_{k+1}) = \{(y_1, \ldots, y_k)^\top \mid (y_1, \ldots, y_k, y_{k+1})^\top \in \mathcal{P}'_{k+1}\}$. Then, $\mathcal{P}_k^-(y_{k+1})$ is a lower polyhedron for any $y_{k+1} \in [\acute{y}_1^{k+1}, \acute{y}_{\kappa_{k+1}}^{k+1}]$. Therefore, $\Pr(\boldsymbol{Y}_{w_1}^{(1:k+1)} \in \mathcal{P}'_{k+1}) = \sum_{j=1}^{\kappa_{k+1}} \Pr(Y_{w_1}^{(k+1)} = \acute{y}_j^{k+1}) \Pr(\boldsymbol{Y}_{w_1}^{(1:k)} \in \mathcal{P}_k^-(\acute{y}_j^{k+1}))$ for any $k+1$-dimensional lower polyhedron $\mathcal{P}'_{k+1} \subset \Pi_{j=1}^{k+1}[\acute{y}_1^j, \acute{y}_{\kappa_j}^j]$. This is greater than or equal to $\sum_{j=1}^{\kappa_{k+1}} \Pr(Y_{w_2}^{(k+1)} = \acute{y}_j^{k+1}) \Pr(\boldsymbol{Y}_{w_2}^{(1:k)} \in \mathcal{P}_k^-(\acute{y}_j^{k+1}))$ by Lemma 1, which in turn equals $\Pr(\boldsymbol{Y}_{w_2}^{(1:k+1)} \in \mathcal{P}'_{k+1})$. This proves the statement for $n = k+1$.

Therefore, by mathematical induction, $\Pr(\boldsymbol{Y}_{w_1} \in \mathcal{P}') \geq \Pr(\boldsymbol{Y}_{w_2} \in \mathcal{P}')$ holds for any lower polyhedron $\mathcal{P}'$.□

Proof of Theorem 3: From (21), we obtain

$$\begin{aligned}\hat{V}_{u,j}^2(t+1) &= \hat{f}_{v_j}^2(\boldsymbol{P}_u(t+1), \boldsymbol{Q}_u(t+1), v_0) \\ &= v_0^2 - 2\sum_{k\in a(j)} \Big(r_k \hat{f}_{p_k^{\mathrm{b}}}(\boldsymbol{P}_u(t+1), \boldsymbol{Q}_u(t+1), v_0) \\ &\quad + x_k \hat{f}_{q_k^{\mathrm{b}}}(\boldsymbol{P}_u(t+1), \boldsymbol{Q}_u(t+1), v_0) \Big) \\ &= v_0^2 - 2\sum_{k\in a(j)} \sum_{l\in d(k)} (r_k P_{u,l}(t+1) + x_k Q_{u,l}(t+1)).\end{aligned} \tag{23}$$

Substituting $P_{u,l}(t+1)$ with $P_l^{\mathrm{L}}(t+1) + \overline{p}_l N_{u,l}^{\mathrm{ON}}(t+1)$, $Q_{u,l}(t+1)$ with $Q_l^{\mathrm{L}}(t+1) + \overline{q}_l N_{u,l}^{\mathrm{ON}}(t+1)$, $N_{u,l}^{\mathrm{ON}}(t+1)$ with the right side of (8), and leveraging (23) we obtain

$$\begin{aligned}\hat{V}_{u,j}(t+1) \geq \underline{v} &\iff \hat{V}_{u,j}^2(t+1) \geq \underline{v}^2 \\ &\iff g_j(\boldsymbol{C}_u(t+1)) \leq h_j(\boldsymbol{R}),\end{aligned}$$

where vector $\boldsymbol{R} := (\boldsymbol{N}^{\mathrm{ON}}(t)^\top, \boldsymbol{P}^{\mathrm{L}}(t+1)^\top, \boldsymbol{Q}^{\mathrm{L}}(t+1)^\top, \boldsymbol{S}^{\mathrm{ON}}(t+1)^\top, \boldsymbol{S}^{\mathrm{OFF}}(t+1)^\top)^\top$ collects random variables, $\boldsymbol{C}_u(t+1) := \boldsymbol{C}_u^{\mathrm{ON}}(t+1) - \boldsymbol{C}_u^{\mathrm{OFF}}(t+1)$ is the net number of TCL OFF to ON switches by the aggregator's input, and the functions g_j and h_j are

$$\begin{aligned}g_j(\boldsymbol{C}_u(t+1)) &:= 2\sum_{k\in a(j)} \sum_{l\in d(k)} (r_k \overline{p}_l + x_k \overline{q}_l) C_{u,l}(t+1), \\ h_j(\boldsymbol{R}) &:= v_0^2 - \underline{v}^2 \\ &\quad - 2\sum_{k\in a(j)} \sum_{l\in d(k)} \big(r_k P_l^{\mathrm{L}}(t+1) + x_k Q_l^{\mathrm{L}}(t+1) \\ &\quad + (r_k \overline{p}_l + x_k \overline{q}_l)\big(N_l^{\mathrm{ON}}(t) + S_l^{\mathrm{ON}}(t+1) - S_l^{\mathrm{OFF}}(t+1)\big)\big).\end{aligned}$$

Note that g_j is a non-negative linear combination of $C_{u,l}(t+1)$ for all $j \in [n]$, i.e., there exist $a_{jl} \geq 0$ for any $j, l \in [n]$ such that $g_j(\boldsymbol{C}_u(t+1))$ is equal to $\sum_{l=1}^{n} a_{jl} C_{u,l}(t+1)$.

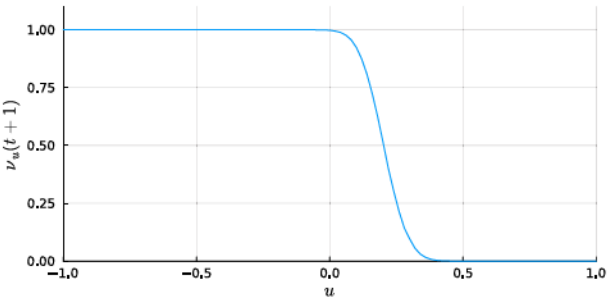

Fig. 5. Demonstration of the monotonicity of $\nu_u(t+1)$ with respect to u.

Let $\mathcal{R}$ be the sample space of $\boldsymbol{R}$ and $f_{\boldsymbol{R}}$ be the joint probability density function of $\boldsymbol{R}$. Then, we have

$$\begin{aligned}\hat{\nu}_u(t+1) &= \Pr\left(\bigwedge_{j=1}^{n} \left(\hat{V}_{u,j}(t+1) \geq \underline{v}\right)\right) \\ &= \int_{\tilde{r}\in\mathcal{R}} \Pr\left(\bigwedge_{j=1}^{n} (g_j(\boldsymbol{C}_u(t+1)) \leq h_j(\tilde{r})) \,\middle|\, \boldsymbol{R} = \tilde{r}\right) f_{\boldsymbol{R}}(\tilde{r}) d\tilde{r}.\end{aligned} \tag{24}$$

For any realization, $\tilde{r} := (\tilde{n}^{\mathrm{ON}}(t)^\top, \tilde{p}^{\mathrm{L}}(t+1)^\top, \tilde{q}^{\mathrm{L}}(t+1)^\top, \tilde{s}^{\mathrm{ON}}(t+1)^\top, \tilde{s}^{\mathrm{OFF}}(t+1)^\top)^\top \in \mathcal{R}$, $C_{u,l}(t+1) = C_u^{\mathrm{ON}}(t+1)$ when $u \geq 0$, and $C_{u,l}(t+1) = -C_u^{\mathrm{OFF}}(t+1)$ when $u < 0$. Thus, by (7), the conditional cdf of $C_{u,l}(t+1)$ is computed as $\Pr(C_{u,l}(t+1) \leq k | \boldsymbol{R} = \tilde{r}) = \mathcal{F}_{\mathrm{B}}(k; \tilde{n}_l^{\mathrm{OFF}}(t) - \tilde{s}_l^{\mathrm{ON}}(t+1), u)$ when $u \geq 0$, and $\Pr(C_{u,l}(t+1) \leq k | \boldsymbol{R} = \tilde{r}) = 1 - \mathcal{F}_{\mathrm{B}}(-k; \tilde{n}_l^{\mathrm{ON}}(t) - \tilde{s}_l^{\mathrm{OFF}}(t+1), -u)$ when $u < 0$. In addition, from [43], the cdf of a binomial random variable $\mathcal{B}(n; \nu)$ is

$$\mathcal{F}_{\mathrm{B}}(k; n, \nu) = (n-k)\binom{n}{k} \int_0^{1-\nu} t^{n-k-1}(1-t)^k dt,$$

which is a monotonically decreasing function with respect to ν. Thus, $\Pr(C_{u,l}(t+1) \leq k | \boldsymbol{R} = \tilde{r})$ monotonically decreases as u increases, and $C_{u,1}(t+1)|\tilde{r}, \ldots, C_{u,n}(t+1)|\tilde{r}$ for any $\tilde{r} \in \mathcal{R}$ satisfies the conditions on the random variables in Lemma 2. Thus, $\Pr(\bigwedge_{j=1}^{n}(g_j(\boldsymbol{C}_u(t+1)) \leq h_j(\tilde{r}) \,|\, \boldsymbol{R} = \tilde{r}))$ is a decreasing function with respect to u. Therefore, by (24), $\hat{\nu}_u(t+1)$ is also a decreasing function with respect to u. □

While Theorem 3 justifies Assumption 1 for the approximation $\hat{\nu}_u(t+1)$, we also empirically validate that $\nu_u(t+1)$ is a monotonically decreasing function of u for both the 56-node network and the 116-node network. We first randomly chose 50 combinations of non-participating load power consumption and TCL modes ($\boldsymbol{P}^{\mathrm{L}}(t)$, $\boldsymbol{Q}^{\mathrm{L}}(t)$, $\boldsymbol{N}^{\mathrm{ON}}(t)$). For each combination, we generate 30,000 realizations of $X_u(t+1)$ for each of 101 uniformly spaced points of u from -1 to 1 and empirically estimated $\nu_u(t+1)$. Through the experiment, we verified that $\nu_u(t+1)$ is monotonically decreasing for all of the combinations. Fig. 5 demonstrates the empirically-measured $\nu_u(t+1)$ of one combination on the 56-node network.

C. Alternative to Assumption 1

Leveraging the monotonicity of $\hat{\nu}_u(t+1)$, we can modify the test procedure in Fig. 2 to remove the need for Assumption 1 as follows. We first assume that the error between LinDistFlow and DistFlow solutions is bounded; similar assumptions have been used in prior work [1], [44]. We then modify the sampling procedure to obtain an estimator of the approximate safety probability $\hat{\nu}_u(t+1)$ with an increased voltage lower bound, $\underline{v}+\gamma$, where γ is the bound on the error between LinDistFlow and DistFlow solutions, and test whether $\hat{\nu}_u(t+1)$ is greater than $1-\epsilon$. For a sufficiently large γ, we can demonstrate that $\hat{\nu}_u(t+1)$ is smaller than the actual safety probability $\nu_u(t+1)$ so that the constraint set obtained from the bisection method using the modified test procedure is a solution of Problem 1. However, in numerical experiments, we found that this modified approach can lead to excessively conservative constraint sets $\mathcal{U}(t+1)$. Given that the original approach is already generally conservative as evidenced by the case study results, we do not include further details on the modified approach and associated results here.

Acknowledgments

The authors would like to thank Zexiang Liu for providing core ideas to simplify the proofs of Lemma 1 and Lemma 2.

References

[1] E. Dall'Anese, S. Guggilam, A. Simonetto, Y. C. Chen, and S. Dhople, "Optimal regulation of virtual power plants," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1868–1881, Mar. 2018.

[2] A. Bernstein and E. Dall'Anese, "Real-time feedback-based optimization of distribution grids: A unified approach," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 3, pp. 1197–1209, Sep. 2019.

[3] E. Vrettos and G. Andersson, "Combined load frequency control and active distribution network management with thermostatically controlled loads," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, 2013, pp. 247–252.

[4] Federal Energy Regulatory Commission, "Notice inviting post-technical conference comments," FERC, Washington, DC, USA, Tech. Rep. RM18-9-000, Apr. 2018. [Online]. Available : https://cms.ferc.gov/sites/default/files/2020-09/Notice-for-Comments-on-NOPR-RM18-9.pdf

[5] Federal Energy Regulatory Commission, "FERC order no. 2222: Participation of distributed energy resource aggregations in markets operated by regional transmission organizations and independent system operators," FERC, Washington, DC, USA, Sep. 2020. [Online]. Available: https://www.ferc.gov/sites/default/files/2020-09/E-1_0.pdf

[6] J. Mathieu, S. Koch, and D. S. Callaway, "State estimation and control of electric loads to manage real-time energy imbalance," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 430–440, Feb. 2013.

[7] S. Bashash and H. K. Fathy, "Modeling and control of aggregate air conditioning loads for robust renewable power management," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 4, pp. 1318–1327, Jul. 2013.

[8] W. Zhang, J. Lian, C.-Y. Chang, and K. Kalsi, "Aggregated modeling and control of air conditioning loads for demand response," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4655–4664, Nov. 2013.

[9] S. Tindemans, V. Trovato, and G. Strbac, "Decentralized control of thermostatic loads for flexible demand response," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 5, pp. 1685–1700, Sep. 2015.

[10] S. C. Ross, N. Ozay, and J. L. Mathieu, "Coordination between an aggregator and distribution operator to achieve network-aware load control," in *Proc. PowerTech*, 2019, pp. 1–6.

[11] S. Ross and J. Mathieu, "Strategies for network-safe load control with a third-party aggregator and a distribution operator," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3329–3339, Jul. 2021.

[12] D. Lee, K. Turitsyn, D. K. Molzahn, and L. A. Roald, "Robust AC optimal power flow with robust convex restriction," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 4953–4966, Nov. 2021.

[13] H. D. Nguyen, K. Dvijotham, and K. Turitsyn, "Constructing convex inner approximations of steady-state security regions," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 257–267, Jan. 2019.

[14] N. Nazir and M. Almassalkhi, "Grid-aware aggregation and realtime disaggregation of distributed energy resources in radial networks," *IEEE Trans. Power Syst.*, vol. 37, no. 3, pp. 1706–1717, May 2022.

[15] K. Petrou, M. Z. Liu, A. T. Procopiou, L. F. Ochoa, J. Theunissen, and J. Harding, "Operating envelopes for prosumers in LV networks: A weighted proportional fairness approach," in *Proc. IEEE ISGT Europe*, 2020, pp. 579–583.

[16] Y. Yi and G. Verbič, "Fair operating envelopes under uncertainty using chance constrained optimal power flow," *Electr. Power Syst. Res.*, vol. 213, 2022, Art. no. 108465.

[17] J. S. Russell, P. Scott, and A. Attarha, "Stochastic shaping of aggregator energy and reserve bids to ensure network security," *Electr. Power Syst. Res.*, vol. 212, 2022, Art. no. 108418.

[18] J. Comden, A. S. Zamzam, and A. Bernstein, "Secure control regions for distributed stochastic systems with application to distributed energy resource dispatch," in *Proc. IEEE Amer. Control Conf.*, 2022, pp. 2208–2213.

[19] S. C. Ross and J. L. Mathieu, "A method for ensuring a load aggregator's power deviations are safe for distribution networks," *Electr. Power Syst. Res.*, vol. 189, 2020, Art. no. 106781.

[20] S. Jang, N. Ozay, and J. L. Mathieu, "Large-scale invariant sets for safe coordination of thermostatic loads," in *Proc. IEEE Amer. Control Conf.*, 2021, pp. 4163–4170.

[21] K. Baker, E. Dall'Anese, and T. Summers, "Distribution-agnostic stochastic optimal power flow for distribution grids," in *Proc. IEEE NAPS*, 2016, pp. 1–6.

[22] E. Dall'Anese, K. Baker, and T. Summers, "Chance-constrained AC optimal power flow for distribution systems with renewables," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3427–3438, Sep. 2017.

[23] A. Hassan, Y. Dvorkin, D. Deka, and M. Chertkov, "Chance-constrained ADMM approach for decentralized control of distributed energy resources," in *Proc. IEEE Power Syst. Comput. Conf.*, 2018, pp. 1–7.

[24] K. S. Ayyagari, N. Gatsis, and A. F. Taha, "Chance constrained optimization of distributed energy resources via affine policies," in *Proc. IEEE Glob. Conf. Signal Inf. Process.*, 2017, pp. 1050–1054.

[25] A. Hassan, R. Mieth, M. Chertkov, D. Deka, and Y. Dvorkin, "Optimal load ensemble control in chance-constrained optimal power flow," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5186–5195, Sep. 2019.

[26] Y. Chen and Y. Lin, "Combining model-based and model-free methods for stochastic control of distributed energy resources," *Appl. Energy*, vol. 283, 2021, Art. no. 116204.

[27] V. Rigoni, D. Flynn, and A. Keane, "Coordinating demand response aggregation with LV network operational constraints," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 979–990, Mar. 2021.

[28] A. Lesage-Landry and D. S. Callaway, "Batch reinforcement learning for network-safe demand response in unknown electric grids," *Electr. Power Syst. Res.*, vol. 212, 2022, Art. no. 108375.

[29] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992.

[30] R. C. Sonderegger, "Dynamic models of house heating based on equivalent thermal parameters," Ph.D. dissertation, Princeton Univ., Princeton, NJ, USA, 1978.

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

[32] W. H. Kersting, "Distribution system modeling and analysis," in *Electric Power Generation, Transmission, and Distribution: The Electric Power Engineering Handbook*. Boca Raton, FL, USA: CRC, 2018, pp. 26–1.

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

[34] R. L. Burden, J. D. Faires, and A. M. Burden, *Numerical Analysis*. Boston, MA, USA: Cengage Learn., 2015.

[35] S. Bolognani and S. Zampieri, "On the existence and linear approximation of the power flow solution in power distribution networks," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 163–172, Jan. 2016.

[36] W. H. Kersting, "Radial distribution test feeders," *IEEE Trans. Power Syst.*, vol. 6, no. 3, pp. 975–985, Aug. 1991.

[37] M. A. Khan and B. P. Hayes, "A reduced electrically-equivalent model of the IEEE European low voltage test feeder," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2022, pp. 1–5.

[38] IEEE PES AMPS DSAS Test Feeder Working Group, "PES AMPS DSAS test feeder working group, " IEEE PES test feeders," 2015. [Online]. Available : https://cmte.ieee.org/pes-testfeeders/resources/
[39] PJM, "RTO regulation signal data for 7.19.2019 & 7.20.2019.xls," 2019. Accessed: Oct. 22, 2019. [Online]. Available: https://www.pjm.com/markets-and-operations/ancillary-services.aspx
[40] S. Jang, N. Ozay, and J. L. Mathieu, "Data-driven estimation of probabilistic constraints for network-safe distributed energy resource control," in *Proc. 58th Annu. Allerton Conf. Commun., Control Comput.*, 2022, pp. 1–8.
[41] M. Z. Liu et al., "Grid and market services from the edge: Using operating envelopes to unlock network-aware bottom-up flexibility," *IEEE Power Energy Mag.*, vol. 19, no. 4, pp. 52–62, Jul./Aug. 2021.
[42] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2017.
[43] G. P. Wadsworth, J. G. Bryan, and A. C. Eringen, "Introduction to probability and random variables," *J. Appl. Mech.*, vol. 28, no. 2, 1961, Art. no. 319.
[44] L. Gan, N. Li, U. Topcu, and S. H. Low, "Exact convex relaxation of optimal power flow in radial networks," *IEEE Trans. Automat. Contr.*, vol. 60, no. 1, pp. 72–87, Jan. 2015.

Sunho Jang (Student Member, IEEE) received the B.S. degree in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2019. He is currently working toward the Ph.D. degree in electrical and computer engineering with the University of Michigan, Ann Arbor, MI, USA. His research interests include safe control for cyber-physical systems, especially with applications to network-safe control of distributed energy resources.

Necmiye Ozay (Senior Member, IEEE) received the B.S. degree from Bogazici University, Istanbul, Türkiye, in 2004, the M.S. degree from the Pennsylvania State University, State College, PA, USA, in 2006, and the Ph.D. degree from Northeastern University, Boston, MA, USA, in 2010, all in electrical engineering. She was a Postdoctoral Scholar with the California Institute of Technology, Pasadena, CA, USA, between 2010 and 2013. She joined the University of Michigan, Ann Arbor, MI, USA, in 2013, where she is currently an Associate Professor of electrical engineering and computer science, and robotics. Her research interests include hybrid dynamical systems, control, optimization and formal methods with applications in cyber-physical systems, system identification, verification, validation, autonomy, and dynamic data analysis.

Johanna L. Mathieu (Senior Member, IEEE) received the B.S. degree in ocean engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2004 and the M.S. and Ph.D. degrees in mechanical engineering from the University of California, Berkeley, CA, USA, in 2008 and 2012, respectively. She is an Associate Professor with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA. Prior to joining the University of Michigan, she was a Postdoctoral Researcher with the Swiss Federal Institute of Technology (ETH) Zurich, Switzerland. Her research interests include modeling, estimation, control, and optimization of distributed energy resources.