

Optimal Load Ensemble Control in Chance-Constrained Optimal Power Flow

Ali Hassan, Robert Mieth, Michael Chertkov, Deepjyoti Deka, Yury Dvorkin

Abstract—Distribution system operators (DSOs) world-wide foresee a rapid roll-out of distributed energy resources. From the system perspective, their reliable and cost effective integration requires accounting for their physical properties in operating tools used by the DSO. This paper describes an decomposable approach to leverage the dispatch flexibility of thermostatically controlled loads (TCLs) for operating distribution systems with a high penetration level of photovoltaic resources. Each TCL ensemble is modeled using the Markov Decision Process (MDP). The MDP model is then integrated with a chance constrained optimal power flow that accounts for the uncertainty of PV resources. Since the integrated optimization model cannot be solved efficiently by existing dynamic programming methods or off-the-shelf solvers, this paper proposes an iterative Spatio-Temporal Dual Decomposition algorithm (ST-D2). We demonstrate the merits of the proposed integrated optimization and ST-D2 algorithm on the IEEE 33-bus test system.

I. INTRODUCTION

Distributed energy resources (DERs) are viewed as a technologically viable alternative to conventional generation resources and, in some cases, have been shown to provide cost-competitive system support services, including peak shaving, ancillary services, emergency and disaster response, and investment deferral [1]. E.g., the State of New York estimates a total technical potential of roof-top photovoltaic (PV) resources (roughly, 80% of DERs) at 2,615 MW of the cumulative peak capacity and 8,223 GWh production by 2030 [2]. Under such PV penetration levels, distribution system operators (DSO) are likely to exhaust existing means for compensating PV intermittency, as well as for distribution power flow and voltage control. In turn, lack of such means may limit the ability to further integrate DER resources. One way to deal with such challenges and overcome existing barriers for DER integration is to fully realize the potential of behind-the-meter demand response capabilities [2]. This paper proposes an approach to leverage the flexibility of behind-the-meter thermostatically controlled loads (TCLs) for operating PV-dominant distribution systems.

Previously, demand response capabilities have been considered at the system-level for centralized, hierarchical, and distributed control architectures [3]. The common element of these architectures is their reliance on load aggregators as mediators between the DSO and behind-the-meter DERs that are not observable by the DSO. Each aggregator can continuously refine knowledge of each individual resource, e.g. via machine learning [4], and use this knowledge to

accurately quantify their dispatch capabilities. In turn, this flexibility whether in the form of load curtailment or load increase can be offered by the aggregator to the DSO in exchange for a certain compensation. The primary difficulty associated with such schemes is two-fold. First, the aggregator needs to model an ensemble of individual TCLs that may vary in a wide range. Second, these resources are driven by the activity of their customers and therefore the aggregator needs to accurately predict their behavior. In case of TCL ensembles, these two difficulties can simultaneously be overcome by modeling each ensemble as a Markov Decision Process (MDP). In [5]–[9], each TCL ensemble is modeled as a discrete-time, discrete-space MDP that is well suited for capturing stochastic dynamics of individual TCLs and is computationally scalable to accommodate hundreds of TCLs in each ensemble. The models [7]–[9] exploit naive economic dispatch frameworks that co-optimize the flexibility of TCL ensembles and distribution system operations. The common caveat of [7]–[9] is that network constraints are neglected and, as a result, these models do not ensure compliance with power flow and voltage limits. Furthermore, [7]–[9] do not account for the uncertainty of PV injections and treat these resources in a deterministic manner. The former caveat is addressed in [10], where TCL ensembles are operated by the aggregator and network constraints are included using the *LinDistFlow* power flow model [11]. The resulting problem in [10] is solved using an iterative solution technique. Similarly to [7]–[9], the model in [10] treats PV injections deterministically and therefore is vulnerable to the effects of their intermittency.

Recent efforts to account for the intermittency of PV resources in decision-support tools for DSOs include the use of stochastic programming, especially chance constraints [12]–[15]. Chance constraints naturally fit distribution system operations as they impose a desired tolerance (probability level) to violations of technical constraints (e.g. power flow and voltage limits) so that DSOs can adjust their tolerance based on their reliability preferences and standards. Additionally, the use of chance constraints is motivated as follows. First, as in [16], they can be reformulated as second-order conic (SOC) constraints that are computationally tractable. Such reformulations exist for multiple probability distributions that are shown to accurately represent the uncertainty of PV resources [17]. Second, chance constraints make it possible to trade-off solution cost and robustness by adjusting the desired tolerance to constraint violations. Finally, chance constraints have a well-established connection to data-driven optimization methods, [18], that can be leveraged to overcome limitations of assuming a particular probability distribution. Dall’Anese

The authors acknowledge support from NSF # CMMI-1825212 and DOE through the Grid Modernization Lab Consortium, and the Center for Non Linear Studies (CNLS) at Los Alamos.

et al. [12] present a chance-constrained (CC) optimal power flow (CC-OPF) model with AC power flow constraints based on the *LinDistFlow* power flow model. This work is extended in [15] by introducing new power-flow-based control policies for PV resources that enhance their ability to participate in voltage regulation and power loss minimization. Reference [13] extends the chance constraints derived in [12], [15] under the Gaussian assumption into a data-robust form. Reference [19] optimizes the TCL dispatch in a centralized manner using the chance-constrained framework. However, solving this optimization from the perspective of the centralized controller increases computational complexity of the problem and, therefore, it may not scale well due to a significant communication overhead.

With the exception of our previous work in [10], the MDP based model for TCLs and CC-OPF optimization are always performed separately. Since TCL ensembles and PV resources are best modeled by the MDP and chance constrained frameworks respectively, this paper seeks to bridge the gap between the MDP approaches to model TCL ensembles from the aggregator perspective, [7]–[9], and the CC-OPF literature that operates the distribution system from the centralized DSO perspective, [12]–[15]. Based on [10], we propose a decomposition-based algorithm that divides the optimization tasks between the DSO and the TCL ensembles, while minimizing communication needs among them. This paper makes the following contributions:

- 1) It formulates an integrated optimization problem that includes both the MDP optimization of TCL ensembles and the CC-OPF optimization of the distribution system. Relative to the previous work in [7]–[9], [12]–[15], the integrated model not only accounts for the TCL dispatch, but also ensures compliance with distribution system limits and internalizes the PV uncertainty via chance constraints. Furthermore, the CC-OPF optimization is extended to account for the expected value of the quadratic real power losses in the objective function.
- 2) To efficiently solve the proposed integrated model, we develop a Spatio-Temporal Dual Decomposition (ST-D2) algorithm, which is based on the traditional dual decomposition [20]. This novel application of the dual decomposition makes it possible to co-optimize the MDP and CC-OPF subproblems iteratively using dynamic programming and SOC programming methods, respectively.

The proposed integrated model and ST-D2 algorithm are tested on the IEEE 33-bus test system presented in [21] to demonstrate the efficiency of our model. To assess the performance of the proposed algorithm and demonstrate its scalability, additional simulations are performed on the 37-, 123- and 8500-bus IEEE systems [22].

The rest of the paper is organized as follows. Section II presents an MDP model for optimizing the dispatch of TCL ensembles operated by the aggregator and then integrates this model with the distribution CC-OPF model. Section III describes the proposed algorithm to solve the integrated model described in Section II. Section IV presents the case study

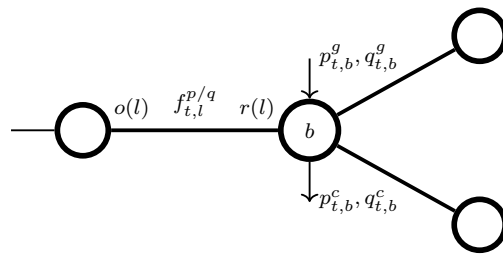


Figure 1. A radial distribution system with notations.

to validate the the proposed model and algorithm. Section V concludes the paper.

II. MATHEMATICAL FORMULATION

A. Preliminaries

We consider a radial distribution system represented by graph $\Gamma = (\mathcal{E}, \mathcal{N})$, where \mathcal{E} and \mathcal{N} are the sets of lines (edges) and buses (nodes), see Fig. 1. The set of operating time intervals is represented by \mathcal{T} , indexed by t . The set of buses where controllable generation resources are located is denoted as $\mathcal{N}^g \subseteq \mathcal{N}$, the set of buses where PV resources are located is denoted as $\mathcal{N}^{PV} \subseteq \mathcal{N}$ and the set of buses where TCL ensembles are located is denoted as $\mathcal{N}^T \subseteq \mathcal{N}$. Each node can be characterized by its active and reactive load ($p_{t,b}^c$ and $q_{t,b}^c$, $\forall t, b \in \mathcal{N}$), active and reactive power output of controllable generation resources ($p_{t,b}^g$ and $q_{t,b}^g$, $\forall t, b \in \mathcal{N}^g$), active and reactive PV generation ($p_{t,b}^{PV}$ and $q_{t,b}^{PV}$, $\forall t, b \in \mathcal{N}^{PV}$), active and reactive injections ($p_{t,b}$ and $q_{t,b}$, $\forall t, b \in \mathcal{N}^T$) and voltage magnitude $v_{t,b} \in [\bar{V}_b, \underline{V}_b]$, $\forall t, b \in \mathcal{N}$, where \bar{V}_b and \underline{V}_b are the upper and lower nodal voltage limits respectively. The square of the nodal voltage is denoted as $u_{t,b} = v_{t,b}^2$, $\forall t, b \in \mathcal{N}$, with limits as $\bar{U}_b = (\bar{V}_b)^2$ and $\underline{U}_b = (\underline{V}_b)^2$. Each line is characterized by its active and reactive power flows ($f_{t,l}^p$ and $f_{t,l}^q$, $\forall t, l \in \mathcal{E}$), its resistance and reactance (R_l and X_l , $l \in \mathcal{E}$). The origin and receiving buses for each line are indexed as $o(l)$ and $r(l)$. The bold font will denote the uncertain quantities.

The main notations are defined as follows:

Variables

$f_{t,l}^{p/q}$	Active/reactive power flows in line l during time interval t
$\mathcal{P}_{t,b}^{\alpha\beta}$	Transition probability from state β to state α for the TCL ensemble at bus b during time interval t
$p_{t,b}/q_{t,b}$	Active/reactive power injections at bus b during time interval t
$p_{t,b}^g/q_{t,b}^g$	Active/reactive power output of controllable generation resources at bus b during time interval t
$p_{t,b}^{PV}/q_{t,b}^{PV}$	Active/reactive PV generation at bus b during time interval t
p_b^α/q_b^α	Rated active/reactive power consumption at state α for the TCL ensemble at bus b
$u_{t,b}$	Voltage magnitude squared at bus b during time interval t
$v_{t,b}$	Voltage magnitude at bus b during time interval t

$\alpha_{t,b}$	Participation factor of the controllable generation resource at bus b during time interval t
$\gamma_{t,b}^{\alpha\beta}$	Cost penalty factor on the transition from state α to state β for the TCL ensemble at bus b during time interval t
$\lambda_{t,b}^{p/q}$	Lagrange multipliers
$\rho_{t,b}^{\alpha/\beta}$	Probability of the TCL ensemble being in state α/β located at bus b during time interval t

Parameters

$\overline{G}_b^p/\underline{G}_b^p$	Maximum/minimum active power output of controllable generation resources at bus b
$\overline{G}_b^q/\underline{G}_b^q$	Maximum/minimum reactive power output of controllable generation resources at bus b
K	Power factor
$\overline{\mathcal{P}}_{t,b}^{\alpha\beta}$	Default transition probability from state β to state α for the TCL ensemble at bus b during time interval t
R_l	Resistance of line l
X_l	Reactance of line l
$\overline{U}_b/\underline{U}_b$	Maximum/minimum limit on the voltage squared at bus b
$\overline{V}_b/\underline{V}_b$	Maximum/minimum limit on the voltage at bus b
V_0	Base voltage squared at root-bus (substation)
δ	Exogenous step-scaling parameter
$\epsilon_{t,b}^{p/q}$	Active/reactive forecast error
$\tilde{\epsilon}_t^{p/q}$	Aggregated active/reactive forecast error
ζ	Tolerance for the termination of the algorithm
η_g/η_v	Violation tolerance on chance constraints
Λ_t	Parameter to monetize active power losses

B. Modeling a TCL Ensemble

We assume that all TCLs which are co-located at the same bus of the distribution system are organized in one TCL ensemble and is operated by one aggregator. Each ensemble is assumed to have a sufficiently large (infinite) number of TCLs. Under this assumption, one can represent each TCL ensemble as a discrete-time and discrete-space MDP and consider that it is capable of maintaining scheduled injections with the distribution system (i.e. there are no fluctuations). The aggregator controls the TCL ensemble by optimizing its transition from one aggregated state to another across the optimization horizon. Similarly to our previous work in [10], we use the MDP framework to build the following model for each TCL ensemble at bus b :

$$\min_{\rho, \mathcal{P}, p, q} \sum_{t \in \mathcal{T}} O_{b,t}^A := \mathbb{E}_\rho \sum_{t \in \mathcal{T}} \sum_{\alpha \in \mathcal{A}} \left(-U_{t+1,b}^\alpha + \sum_{\beta \in \mathcal{A}} \gamma_{t,b}^{\alpha\beta} \log \frac{\mathcal{P}_{t,b}^{\alpha\beta}}{\overline{\mathcal{P}}_{t,b}^{\alpha\beta}} \right) \quad (1)$$

$$\rho_{t+1,b}^\alpha = \sum_{\beta \in \mathcal{A}} \mathcal{P}_{t,b}^{\alpha\beta} \rho_{t,b}^\beta, \quad \forall \alpha \in \mathcal{A}, t \in \mathcal{T} \quad (2)$$

$$p_{t,b} = \sum_{\alpha \in \mathcal{A}} p_b^\alpha \rho_{t,b}^\alpha, \quad \forall t \in \mathcal{T}, b \in \mathcal{N} \quad (3)$$

$$q_{t,b} = \sum_{\alpha \in \mathcal{A}} q_b^\alpha \rho_{t,b}^\alpha, \quad \forall t \in \mathcal{T}, b \in \mathcal{N} \quad (4)$$

$$\sum_{\alpha \in \mathcal{A}} \mathcal{P}_{t,b}^{\alpha\beta} = 1, \quad \forall t \in \mathcal{T}, \beta \in \mathcal{A}, b \in \mathcal{N}^T \quad (5)$$

where $\rho_{t,b}^\alpha \geq 0$ and $\rho_{t,b}^\beta \geq 0$ are decision variables that characterize the probability that TCLs at bus b are operated in states α and β and have active power consumptions modeled by parameters $p_{t,b}^\alpha$ and $p_{t,b}^\beta$, respectively. The set of states for each TCL ensemble is denoted as \mathcal{A} and individual states α and β are $\alpha, \beta \in \mathcal{A}$.¹ These states are obtained by discretizing the range of power consumption for each TCL ensemble given the operating range of each TCL (see Fig. 4). Decision variables $\rho_{t,b}^\alpha$ and $\rho_{t,b}^\beta$ are related via the transition probability $\mathcal{P}_{t,b}^{\alpha\beta}$ that characterizes the probability of the transition of TCLs at bus b from state β at time t to state α at time $t + 1$. On the other hand, parameter vector $\overline{\mathcal{P}}_{t,b}^{\alpha\beta}$ represents the default transition probability, i.e. internal dynamics of the TCL ensemble without actions of the aggregator. In practice, one can dynamically estimate $\overline{\mathcal{P}}_{t,b}^{\alpha\beta}$ from historical observations using reinforcement learning [4].

Eq. (1) represents the objective function of the aggregator that controls the TCL ensemble and aims to maximize the expected utility of the aggregator ($U_{t+1,b}^\alpha$) and to minimize the discomfort cost for the TCL ensemble, which is computed using the exogenous cost penalty ($\gamma_{t,b}^{\alpha\beta}$) and the Kullback-Leibler (KL) distance to penalize the difference between the transition decisions made by the aggregator ($\mathcal{P}_{t,b}^{\alpha\beta}$) and the default transitions of the TCL ensemble ($\overline{\mathcal{P}}_{t,b}^{\alpha\beta}$). The choice of Kullback-Leibler distance for the penalty cost is motivated by its wide use for modeling randomness of discrete and continuous time-series. Other penalty functions can also be used instead. Eq. (2) describes the temporal evolution of the TCL ensemble, where initial conditions over the course of the optimization horizon are given by the decision of the aggregator during the previous optimization horizon. Eq. (3)-(4) computes the expected active and reactive power injections of the TCL ensemble to the distribution system. Eq. (5) imposes the integrality constraint on the transition decisions optimized by the aggregator such that their total probability is equal to one.

The optimization in Eq. (1)-(5) can be solved using dynamic programming that facilitates scalability of our approach and the ability to solve a large number of such optimizations, one for each TCL ensemble, in parallel. This property is particularly helpful when one deals with a large penetration of TCL ensembles anticipated in distribution systems of the future. Thus, in the following, we use a backward-forward algorithm to solve Eq. (1)-(5). This algorithm is an iterative, two-step procedure that is commonly used for inferring probabilities of unknown state probabilities for Markov processes. We customize this procedure to find the optimal TCL transitions ($\mathcal{P}_{t,b}^{\alpha\beta}$) as further described in Appendix A. Note that the optimization in Eq. (1)-(5) can be represented as a Linearly Solvable MDP (LS-MDP) [23], if $\gamma_{t,b}^{\alpha\beta} = 0$, i.e., not state-dependent. Such LS-MDP problems can be solved analytically,

¹Note that the ensemble can remain in the same state at time t and $t + 1$. In this case state β is such that $\beta = \alpha \in \mathcal{A}$.

i.e. without relying on iterative solution techniques [24]–[26], which can be exploited in online dispatch applications.

C. Chance Constrained Optimal Power Flow

Proliferation of DERs imposes uncertainty on the nodal power injections (e.g. due to the solar irradiance). Unlike TCL ensembles, this uncertainty can be accurately parameterized using standard probability distributions and thus endogenously modeled in decision support tools used by the DSO in a computationally tractable manner. We therefore formulate a CC-OPF that takes the DSO perspective and seeks the least-cost strategy to operate the distribution system given its technical limits and PV uncertainty².

1) *Deterministic OPF*: The CC-OPF is built based on the following deterministic OPF model that considers AC power flows using the *LinDistFlow* model [11] and disregards the PV uncertainty:

$$\min_{\substack{p_t^g, q_t^g, p_t^c, \\ q_t^c, u_t, f_t^p, f_t^q}} \sum_{t \in \mathcal{T}} \sum_{l \in \mathcal{E}} R_l \frac{f_{t,l}^p{}^2 + f_{t,l}^q{}^2}{V_0^2} \quad (6)$$

$$f_{t,l(b)}^p + p_{t,b}^g + p_{t,b}^{\mathcal{PV}} = p_{t,b}^c + p_{t,b} + \sum_{l|o(l)=b} f_{t,l}^p, \forall t \in \mathcal{T}, b \in \mathcal{N} \quad (7)$$

$$f_{t,l(b)}^q + q_{t,b}^g + q_{t,b}^{\mathcal{PV}} = q_{t,b}^c + q_{t,b} + \sum_{l|o(l)=b} f_{t,l}^q, \forall t \in \mathcal{T}, b \in \mathcal{N} \quad (8)$$

$$u_{t,r(l)} = u_{t,o(l)} - 2(R_l f_{t,l}^p + X_l f_{t,l}^q), \forall t \in \mathcal{T}, l \in \mathcal{E} \quad (9)$$

$$\underline{G}_b^p \leq p_{t,b}^g \leq \overline{G}_b^p, \forall t \in \mathcal{T}, b \in \mathcal{N}^{\mathcal{G}} \quad (10)$$

$$\underline{G}_b^q \leq q_{t,b}^g \leq \overline{G}_b^q, \forall t \in \mathcal{T}, b \in \mathcal{N}^{\mathcal{G}} \quad (11)$$

$$\underline{U}_b \leq u_{t,b} \leq \overline{U}_b, \forall t \in \mathcal{T}, b \in \mathcal{N} \quad (12)$$

Eq. (6) minimizes the active power losses in the distribution system. Note that the proposed formulation and algorithm can accommodate other choices of the objective function (e.g. cost-minimization). Eq. (7)–(9) are nodal active and reactive power balances as in the *LinDistFlow* model [11]. Although Eq. (7)–(9) neglect the effect of power losses, the objective function in Eq. (6) can still be formulated in a loss-minimization manner based on the active and reactive power flows $f_{t,l}^p$ and $f_{t,l}^q$ provided by the *LinDistFlow* model. In Eq. (7)–(8), $p_{t,b}$ and $q_{t,b}$ are parameterized and obtained from the MDP optimization in Eq. (1)–(5). Eq. (10)–(11) enforce the minimum and maximum limits on the active and reactive power output of controllable generation resources. Eq. (12) limits voltage magnitudes squared within their minimum and maximum values.

2) *PV Uncertainty*: The uncertain PV output at every bus b and time interval t is defined as $p_{t,b}^{\mathcal{PV}} = p_{t,b}^p - \epsilon_{t,b}^p$, where $p_{t,b}^p$ is the forecast value and $\epsilon_{t,b}^p$ is a forecast error. We assume that this forecast error follows a zero-mean, normal distribution with variance $\sigma_{t,b}^2$, i.e. $\epsilon_{t,b}^p \sim N(0, \sigma_{t,b}^2)$. The forecast error in that form is commonly provided by forecast vendors (e.g.

² Additional uncertainty may also arise from the TCL ensembles due to a large but finite number of TCL users. However, as per the law of large numbers, these fluctuations scale as $\sim 1/\sqrt{N}$, where N is a number of users, and are thus significantly smaller than $O(1)$ fluctuations of the PV resources.

[27]). Furthermore, the inaccuracy of assuming normally distributed forecast errors can be mitigated in the CC-OPF using data-robust approaches as in [13], [28]. Since active power forecast errors also cause fluctuations of the reactive power, we assume that the latter errors are proportional, i.e. $\epsilon_{t,b}^q = \epsilon_{t,b}^p K$, where K is a parameter computed for a given power factor. To compensate for the forecast error and thus to maintain the generation-load balance, controllable generators operated by the DSO are assumed to adjust their output based on a proportional control law, [13], [16], [28]. This control assumes that each controllable generator compensates a fraction of the aggregated forecast error by changing its real-time active ($p_{t,b}^g$) and reactive ($q_{t,b}^g$) power outputs around its generation setpoints $p_{t,b}^g$ and $q_{t,b}^g$ optimized for a given forecast based on optimized participation factors $\alpha_{t,b}$. This control is formalized as:

$$p_{t,b}^g = p_{t,b}^g + \alpha_{t,b} \tilde{\epsilon}_t^p, \quad \forall t, b \in \mathcal{N}^{\mathcal{G}} \quad (13)$$

$$q_{t,b}^g = q_{t,b}^g + \alpha_{t,b} \tilde{\epsilon}_t^q, \quad \forall t, b \in \mathcal{N}^{\mathcal{G}} \quad (14)$$

where $\tilde{\epsilon}_t^p = \sum_{b \in \mathcal{N}^{\mathcal{PV}}} \epsilon_{t,b}^p$ and $\tilde{\epsilon}_t^q = \sum_{b \in \mathcal{N}^{\mathcal{PV}}} \epsilon_{t,b}^q$ are the aggregated forecast errors for every time interval t . Participation factors $\alpha_{t,b}$ are optimized to accommodate different technical and cost characteristics of controllable generators with the condition $\sum_{b \in \mathcal{N}^{\mathcal{G}}} \alpha_{t,b} = 1, \forall t \in \mathcal{T}$, i.e. the total change in the output of controllable generators is equal to the aggregated forecast error.

Following the realization of $\epsilon_{t,b}^p$ and $\epsilon_{t,b}^q$, the real-time active ($f_{t,l}^p$) and reactive ($f_{t,l}^q$) power flows are modeled as:

$$f_{t,l}^p = f_{t,l}^p + a_{l*}(\epsilon_t^p - \alpha_t \tilde{\epsilon}_t^p), \quad \forall t, \forall l \quad (15)$$

$$f_{t,l}^q = f_{t,l}^q + a_{l*}(\epsilon_t^q - \alpha_t \tilde{\epsilon}_t^q), \quad \forall t, \forall l \quad (16)$$

where a_{l*} denotes the l^{th} row of a matrix $A : |\mathcal{E}| \times |\mathcal{N}|$ with elements $a_{(lb)}$ which we define such that:

$$a_{(lb)} = \begin{cases} 1, & \text{if line } l \text{ is part of the path} \\ & \text{from root to bus } b \quad l \in \mathcal{E}, b \in \mathcal{N} \\ 0, & \text{otherwise} \end{cases}$$

and α_t is a $|\mathcal{N}| \times 1$ vector with elements $\alpha_{t(b)}$ such that:

$$\alpha_{t(b)} = \begin{cases} \alpha_{t,b}, & \text{if } b \in \mathcal{N}^{\mathcal{G}} \\ 0, & \text{otherwise.} \end{cases}$$

Accordingly, one can use Eq. (15)–(16) to derive the real-time voltage magnitudes squared ($u_{t,b}$):

$$u_{t,r(l)} = u_{t,r(l)} - 2a_{*r(l)}^{\top} [RA(\epsilon_t^p - \alpha_t \tilde{\epsilon}_t^p) + XA(\epsilon_t^q - \alpha_t \tilde{\epsilon}_t^q)], \quad \forall t \in \mathcal{T}, b \in \mathcal{N} \quad (17)$$

where R, X are $|\mathcal{E}| \times |\mathcal{E}|$ matrices with diagonal entries consisting of the line resistances and reactances respectively: $R_{(ii)} = R_i, R_{(ij,i \neq j)} = 0, X$ in analogy.

3) *Formulation*: Using the results in Eqs. (13)–(17), the deterministic OPF in Eq. (6)–(12) can be converted into the following CC-OPF formulation that accounts for real-time

quantities $\mathbf{p}_{t,b}^g$, $\mathbf{q}_{t,b}^g$, $\mathbf{f}_{t,l}^p$, $\mathbf{f}_{t,l}^q$ and $\mathbf{u}_{t,b}$ following realizations of $\epsilon_{t,b}^p$ and $\epsilon_{t,b}^q$:

$$\min_{\substack{p^g, q^g, \\ p^c, q^c, u, f^p, \\ f^q, \alpha}} \sum_{t \in \mathcal{T}} O_t^D = \mathbb{E}_{\epsilon^p, \epsilon^q} \left[\sum_{t \in \mathcal{T}} \sum_{l \in \mathcal{E}} R_l \frac{\mathbf{f}_{t,l}^{p^2} + \mathbf{f}_{t,l}^{q^2}}{V_0^2} \right] \quad (18)$$

$$\text{Eq. (7)-(9)}, \quad (19)$$

$$\text{Eq. (13)-(14), (17)} \quad (20)$$

$$\mathbb{P}(\mathbf{p}_{t,b}^g \leq \overline{G}_b^p) \geq (1 - \eta_g), \quad \forall t \in \mathcal{T}, b \in \mathcal{N}^G \quad (21)$$

$$\mathbb{P}(\underline{G}_b^p \leq \mathbf{p}_{t,b}^g) \geq (1 - \eta_g), \quad \forall t \in \mathcal{T}, b \in \mathcal{N}^G \quad (22)$$

$$\mathbb{P}(\mathbf{q}_{t,b}^g \leq \overline{G}_b^q) \geq (1 - \eta_g), \quad \forall t \in \mathcal{T}, b \in \mathcal{N}^G \quad (23)$$

$$\mathbb{P}(\underline{G}_b^q \leq \mathbf{q}_{t,b}^g) \geq (1 - \eta_g), \quad \forall t \in \mathcal{T}, b \in \mathcal{N}^G \quad (24)$$

$$\mathbb{P}(\mathbf{u}_{t,b} \leq \overline{U}_b) \geq (1 - \eta_v), \quad \forall t \in \mathcal{T}, b \in \mathcal{N} \quad (25)$$

$$\mathbb{P}(\underline{U}_b \leq \mathbf{u}_{t,b}) \geq (1 - \eta_v), \quad \forall t \in \mathcal{T}, b \in \mathcal{N}, \quad (26)$$

where η_g and η_v are exogenous parameters that define tolerance to constraint violations. Eq. (21)-(26) are chance constraints on the power outputs of conventional generators and voltage magnitudes squared that replace deterministic constraints in Eq. (10)-(12). Under the assumption that $\eta_g, \eta_v < 0.5$, one can recast Eqs. (21)-(26) into SOC constraints that can in turn be solved efficiently using off-the-shelf solvers, [16], [28]. The reformulation process for Eq. (21)-(26) is shown in Appendix B. Accordingly, the expected value in the objective function given by Eq. (18) is derived below as:

$$\sum_{t \in \mathcal{T}} O_t^D = \sum_{t \in \mathcal{T}} \sum_{l \in \mathcal{E}} \frac{R_l}{V_0^2} (\mathbb{E}[\mathbf{f}_{t,l}^{p^2}] + \mathbb{E}[\mathbf{f}_{t,l}^{q^2}]), \quad (27)$$

where $\mathbb{E}[\mathbf{f}_{t,l}^{p^2}]$ and $\mathbb{E}[\mathbf{f}_{t,l}^{q^2}]$ are as follows:

$$\begin{aligned} \mathbb{E}[\mathbf{f}_{t,l}^{p^2}] &= \text{Var}(\mathbf{f}_{t,l}^p) + (f_{t,l}^p)^2 \\ &= \sum_{j \in \mathcal{N}} [a_{lj} (\text{Var}(\epsilon_{t,j}^p) - \alpha_{t,j}^2 \text{Var}(\tilde{\epsilon}_t^p))] + (f_{t,l}^p)^2 \end{aligned} \quad (28)$$

$$\begin{aligned} \mathbb{E}[\mathbf{f}_{t,l}^{q^2}] &= \text{Var}(\mathbf{f}_{t,l}^q) + (f_{t,l}^q)^2 \\ &= \sum_{j \in \mathcal{N}} [a_{lj} (\text{Var}(\epsilon_{t,j}^q) - \alpha_{t,j}^2 \text{Var}(\tilde{\epsilon}_t^q))] + (f_{t,l}^q)^2 \end{aligned} \quad (29)$$

Given Eq. (28)-(29), the right hand-side of the original objective function in Eq. (27) can be re-written as:

$$\begin{aligned} \sum_{t \in \mathcal{T}} O_t^D &= \sum_{t \in \mathcal{T}} \sum_{l \in \mathcal{E}} \frac{R_l}{V_0^2} \left[\sum_{j \in \mathcal{N}} [a_{lj} (\text{Var}(\epsilon_{t,j}^p) - \alpha_{t,j}^2 \text{Var}(\tilde{\epsilon}_t^p))] + f_{t,l}^{p^2} \right. \\ &\quad \left. + \sum_{j \in \mathcal{N}} [a_{lj} (\text{Var}(\epsilon_{t,j}^q) - \alpha_{t,j}^2 \text{Var}(\tilde{\epsilon}_t^q))] + f_{t,l}^{q^2} \right]. \end{aligned} \quad (30)$$

Thus, Eq. (30) is a quadratic, deterministic equivalent of the original objective function permissible for off-the-shelf solvers.

Remark 1: While the Gaussian assumption to represent the PV uncertainty is sufficient for the needs of this paper, it is not generally restrictive. As shown in [13], [29], [30], non-Gaussian distributions or their approximation via a Gaussian mixture can be used to reformulate the chance constraints without increasing computational complexity of the model.

Such reformulations tend to yield a more robust, but expensive solution.

Remark 2: Note that the CC-OPF in Eq.(18)-(26) does not impose power flow limits on $\mathbf{f}_{t,l}^p$ and $\mathbf{f}_{t,l}^q$, because distribution systems are typically voltage-constrained and power flow limits can be disregarded. In [31], we describe an approach to enforce chance-constrained apparent power limits.

Remark 3: The PV uncertainty is modeled with individual chance constraints as given by Eq. (21)-(26), which are computationally tractable. Alternatively, one can consider the joint chance constraint over Eq. (21)-(26), but such a formulation would be computationally unbearable [32]. Since replacing the joint chance constraint with individual chance constraints may lead to a conservative solution, especially for large networks, one needs to tune the value of parameters η_g and η_v .

D. Integrated Optimization Problem

Based on the models in Eq. (1)-(5) and in Eq. (18)-(26), the integrated optimization problem that includes MDP and CC-OPF is formulated as follows:

$$\min_{\substack{\rho, \mathcal{P}, u, p, q \\ f^p, f^q, \\ p^c, q^c, p^g, q^g}} \sum_{t \in \mathcal{T}} \left[\sum_{b \in \mathcal{N}^T} O_{b,t}^A + \Lambda_t O_t^D \right] \quad (31)$$

$$\text{Eq. (2) - (5)} \quad (32)$$

$$\text{Eq. (19) - (26)}, \quad (33)$$

where parameter Λ_t is a tariff that monetizes the active power losses to make them comparable to the MDP objective function. The optimization in Eq. (31)-(33) cannot be solved efficiently using existing dynamic programming methods and off-the-shelf solvers. This motivates the solution technique described in Section III.

III. SOLUTION TECHNIQUE

To solve the integrated problem in (31)-(33), we propose a Spatio-Temporal Dual Decomposition (ST-D2) algorithm that exploits two ideas. First, we seek the consensus between the MDP and CC-OPF and thus use a dual decomposition of the original problem. Second, we decouple some spatially- and temporally-independent decision variables. The spatial separation is applied because each TCL ensemble is located at a unique bus and therefore can be optimized separately. In this case, the MDP optimization for each TCL ensemble is performed over the entire optimization horizon to capture inter-temporal constraints on each TCL ensemble. On the other hand, the CC-OPF decisions are temporally separable since controllable generators located in the distribution system typically do not have such inter-temporal constraints as ramping rate and minimum up/down time limits, which are customary for transmission systems. Therefore, the CC-OPF can be solved separately for each time interval t , see [12] for a time-decoupled OPF example. In the presence of temporally coupled resources, e.g. energy storage systems, which introduce the inter-temporal constraints, the CC-OPF can be solved over the entire optimization horizon. This will increase computing times, but is still computationally tractable as shown in [13].

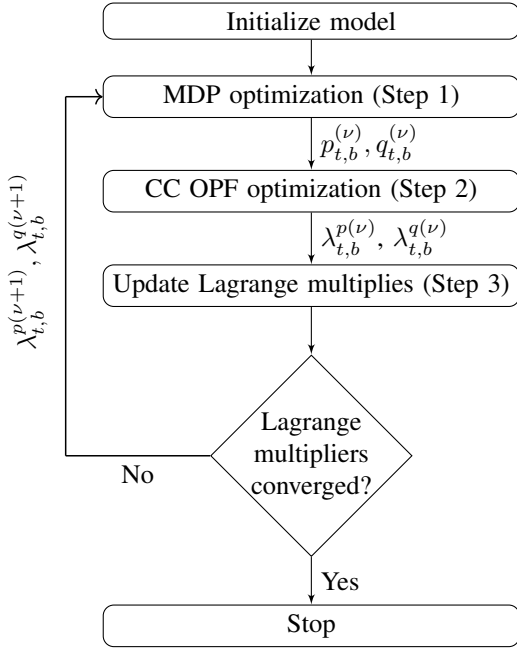


Figure 2. Spatio-Temporal Dual Decomposition (ST-D2) Algorithm.

The proposed ST-D2 algorithm iterates as illustrated in Fig. 2 and each step is further itemized below, where ν is an iteration counter:

- 1) Solve the MDP for each TCL ensemble:

$$\forall b \in \mathcal{N}^T : \min_{\rho, \mathcal{P}} \sum_{t \in \mathcal{T}} O_{b,t}^{A(\nu)}$$

Eq. (2) – (5)

$$U_{t+1,b}^{\alpha(\nu)} = U_{t,b}^{\alpha(\nu)} + \lambda_{t,b}^{p(\nu)} p_b^\alpha + \lambda_{t,b}^{q(\nu)} q_b^\alpha, \quad \forall \alpha \in \mathcal{A}, t \in \mathcal{T}$$

where $\lambda_{t,b}^{p(\nu)}$ and $\lambda_{t,b}^{q(\nu)}$ are the Lagrange multipliers of Eq. (3) and (4), respectively, obtained at the previous iteration of the ST-D2 algorithm. Hence, $\lambda_{t,b}^{p(\nu=1)} = \lambda_{t,b}^{q(\nu=1)} = 0$ during the first iteration.

- 2) Solve the CC-OPF problem, where each TCL ensemble is parameterized using the values of Lagrange multipliers $\lambda_{t,b}^{p(\nu)}$:

$$\forall t \in \mathcal{T} : \min_{p^g, q^g, p^c, q^c, u, f^p, f^q} \sum_{t \in \mathcal{T}} \Lambda_t O_t^{D(\nu)} - \sum_{b \in \mathcal{N}} (\lambda_{t,b}^{p(\nu)} p_{t,b}^{(\nu)} + \lambda_{t,b}^{q(\nu)} q_{t,b}^{(\nu)})$$

Eq. (19) – (26),

where the CC-OPF problems for all time intervals are solved in parallel.

- 3) Update the Lagrange multipliers:

$$\lambda_{t,b}^{p(\nu+1)} \leftarrow \lambda_{t,b}^{p(\nu)} + \delta \left(\sum_{\alpha \in \mathcal{A}} p_b^\alpha \rho_{t,b}^{\alpha(\nu)} - p_{t,b}^{(\nu)} \right)$$

$$\lambda_{t,b}^{q(\nu+1)} \leftarrow \lambda_{t,b}^{q(\nu)} + \delta \left(\sum_{\alpha \in \mathcal{A}} q_b^\alpha \rho_{t,b}^{\alpha(\nu)} - q_{t,b}^{(\nu)} \right)$$

where δ is an exogenous parameter that can be tuned to improve computational performance [20].

These iterations continue until $\lambda_{t,b}^{p(\nu)}$ and $\lambda_{t,b}^{q(\nu)}$ converge with a given termination tolerance (ζ).

Remark 4: The proposed ST-D2 algorithm is based on the dual decomposition and, therefore, it inherits algorithmic properties of the dual decomposition, including convergence properties and the ability to deal with non-convex decisions, e.g. [33]. Furthermore, it can be extended to the Alternating Direction Method of Multipliers (ADMM), which has gained attention in distributed power grid applications [34], [35], by adding the Augmented Lagrangian terms to the objective function of the integrated problem. As pointed out in [36], the ADMM algorithm blends the benefits of the dual decomposition and Augmented Lagrangian methods.

IV. CASE STUDY

The case study uses the IEEE 33-bus distribution system [21], as shown in Fig. 3, where the root bus of the distribution system is connected to the transmission network. The root bus is considered as an infinite source of power supply. One controllable distributed generator with the maximum capacity of 30 kW is placed at bus # 14. We consider that the power supply cost from the root bus and from the distributed generator is equal and set $\Lambda_t = \$10/\text{kWh}$. Seven PV resources, with the rated capacity of 1.5 kW, are installed at buses # 2, 3, 6, 18, 21, 25 and 32 and produce at zero cost. The forecast error of each PV resource is zero-mean and its standard deviation is set to 30% of the forecast output unless stated otherwise. The loads at buses # 17, 20, 23 and 26 are replaced with TCL ensembles of the equivalent capacity. Each ensemble is discretized in 8 states, as shown in Fig. 4, with the default transition probabilities ($\bar{\mathcal{P}}_{t,b}$) between the states as shown in Table I. Each TCL ensemble can be dispatched in the range of 10% - 200% of its average load. To assess the impact of TCL users' comfort, the cost penalty ($\gamma_{t,b}^{\alpha\beta}$) is considered for two cases. The first case, referred to in the following as the *uniform*

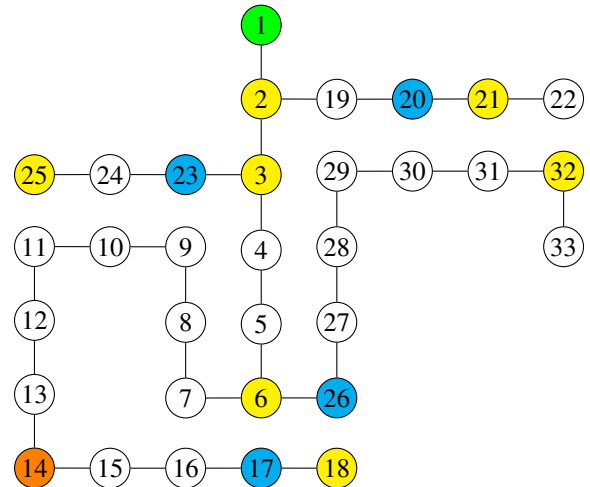


Figure 3. A schematic representation of the IEEE 33-bus distribution test system [21], where the root bus is denoted in green (# 1), bus with distributed generator in orange (# 14), buses with PV resources in yellow (# 2, 3, 6, 18, 21, 25, 32), and buses with TCL ensembles in blue (# 17, 20, 23, 26).

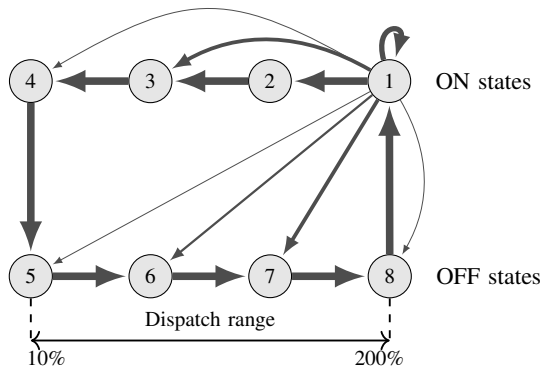


Figure 4. A MDP representation of the TCL ensemble with 8 states displaying all possible transitions from state 1. The active and reactive power consumptions at each state are obtained as uniform intervals within 10% - 200% of the average load of the TCL ensemble.

Table I
DEFAULT TRANSITION PROBABILITIES OF EACH TCL ENSEMBLE BETWEEN EIGHT STATES IN FIGURE 4.

State	1	2	3	4	5	6	7	8
1	0.2	0.5	0.1	0.03	0.02	0.03	0.1	0.02
2	0.02	0.2	0.5	0.1	0.03	0.02	0.03	0.1
3	0.1	0.02	0.2	0.5	0.1	0.03	0.02	0.03
4	0.03	0.1	0.02	0.2	0.5	0.1	0.03	0.02
5	0.02	0.03	0.1	0.02	0.2	0.5	0.1	0.03
6	0.03	0.02	0.03	0.1	0.02	0.2	0.5	0.1
7	0.1	0.03	0.02	0.03	0.1	0.02	0.2	0.5
8	0.5	0.1	0.03	0.02	0.03	0.1	0.02	0.2

cost penalty case, assumes the same penalty for each possible transition shown in Fig. 4, i.e. $\gamma_{t,b}^{\alpha\beta} = 1\$$. The second case, referred to in the following as the *non-uniform cost penalty* case, differentiates between the transitions along the cycle (e.g., $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 7 \rightarrow 8 \rightarrow 1$) and other transitions (e.g., $1 \rightarrow 3, 1 \rightarrow 4, 1 \rightarrow 5, 1 \rightarrow 6, 1 \rightarrow 7, 1 \rightarrow 8$), where the former transitions are penalized using $\gamma_{t,b}^{\alpha\beta} = 1\$$ and the latter transitions are penalized with $\gamma_{t,b}^{\alpha\beta} = 10\$$. This differentiation allows to put a higher cost penalty on transitions with a larger power change that are more disruptive for TCL users' comfort. The optimization horizon consists of 24 hourly time intervals. For the sake of simplicity, we assume that $\eta_v = \eta_g$.

All simulations are performed in Julia JuMP [37] using the Ipopt solver on an Intel Core i5 1.6 GHz processor with 4 GB of RAM. The value of ζ is set to 0.0001. The code and input data used in this paper are available in [38].

A. Computational Performance

In the following numerical experiments, the proposed ST-D2 algorithm converges in 4-7 iterations. There is no explicit correlation observed between the complexity of the problem (number of TCL ensembles considered) and the number of iterations and computing times required for convergence. For example, the most complex instance with four TCL ensembles is solved in 189.90 seconds. Table II itemizes this computing time for the MDP optimization (Step 1) and CC-OPF optimization (Step 2). Within these four iterations, only 19.92 seconds ($\approx 10.4\%$ of the total computing time) is spent on the MDP optimization in Step 1, while the rest of the time is

Table II
COMPUTATIONAL PERFORMANCE OF THE PROPOSED ST-D2 ALGORITHM WITH FOUR TCL ENSEMBLES OVER 24 TIME INTERVALS.

Iteration No.	Computational Time (s)	
	MDP (Step 1)	CC-OPF (Step 2)
1	6.20	36.93
2	4.12	42.48
3	5.68	43.15
4	3.92	47.42
Total time	189.90	

Table III
CONVERGENCE OF THE ST-D2 ALGORITHM FOR TIME INTERVALS AT 8, 16 AND 24 HOURS FOR THE TCL ENSEMBLE AT BUS # 17.

Iteration No.	Lagrange Multipliers					
	t=8		t=16		t=24	
	λ^p	λ^q	λ^p	λ^q	λ^p	λ^q
1	15.4205	0.2575	15.4205	0.2575	15.4205	0.2575
2	12.5518	-0.6987	12.6090	-0.6796	12.8543	-0.5978
3	12.6305	-0.6724	12.6837	-0.6547	12.9228	-0.5750
4	12.6284	-0.6731	12.6818	-0.6553	12.9210	-0.5756
5	12.6284	-0.6731	12.6818	-0.6553	12.9210	-0.5756

Table IV
COMPUTATIONAL PERFORMANCE OF THE PROPOSED ST-D2 ALGORITHM FOR A DIFFERENT NUMBER OF STATES WITHIN A GIVEN TCL ENSEMBLE OVER 24 TIME INTERVALS.

No. of States	Computational Time (s)		
	MDP (Step 1)	CC-OPF (Step 2)	Total
4	5.56	137.08	142.64
8	19.92	169.98	189.90
12	35.81	178.68	214.49
24	122.29	163.41	285.7

incurred by the CC-OPF optimization in Step 2. These results demonstrate that the MDP optimization has one-tenth bearing on the overall complexity of the ST-D2 algorithm relative to the computational burden of the CC-OPF optimization. The convergence of this instance is detailed in Table III, where values of Lagrange multipliers $\lambda_{t,b}^p$ and $\lambda_{t,b}^q$ for the TCL ensemble at bus # 17 are itemized for each iteration for time intervals at 8, 16, and 24 hours. Comparing the results for the 4th and 5th iterations in Table III reveals that desired tolerance ζ is achieved.

Table IV and Table V demonstrate the computational performance of the ST-D2 algorithm for a different number of MDP states used to represent a given TCL ensemble and for a different number of TCL ensembles hosted in the distribution system, respectively. Naturally, increasing the number of states in each ensemble and the number of TCL ensembles in the system leads to greater computing times. Table VI summarizes the computing times needed to solve the proposed model using the proposed ST-D2 algorithm on the 37-, 123- and 8500-bus IEEE systems, [22], with a different number of TCL ensembles. As expected, the computational time increases for a greater number of buses and TCL ensembles. However, in all cases considered in Table VI, the optimal solution is returned by the ST-D2 algorithm within an acceptable time for operational tasks.

B. System Perspective

Solving the integrated optimization problem as in Eq. (31)-(33) leads to the following two main results. First, it reduces the active power losses in the distribution system as explicitly

Table V
COMPUTATIONAL PERFORMANCE OF THE PROPOSED ST-D2 ALGORITHM WITH A DIFFERENT NUMBER OF TCL ENSEMBLES OVER 24 TIME INTERVALS.

No. of TCL Ensembles	Computational Time (s)		
	MDP (Step 1)	CC-OPF (Step 2)	Total
1	19.96	166.44	186.40
2	19.22	167.76	186.98
3	19.08	167.22	186.30
4	19.92	169.98	189.90

Table VI
COMPUTATIONAL PERFORMANCE OF THE PROPOSED ST-D2 ALGORITHM FOR LARGER IEEE SYSTEMS.

	Total Computational Time (s)		
	1 TCL ensemble	37 TCL ensemble	123 TCL ensemble
IEEE 37-bus	164.3*	178.9	N/A
IEEE 123-bus	181.3*	201.0*	267.4
IEEE 8500-bus	611.2*	624.7*	902.3*

* All TCLs are located in the electrically most remote nodes from the root node of the distribution system.

formulated in the objective function. Second, it improves compliance with voltage limits.

Fig. 5 illustrates the effect of dispatching TCL ensembles within the integrated optimization on the active power losses. If there is no MDP optimization, the losses remain constant during the optimization horizon. The uniform cost penalty that dispatches the TCLs more aggressively than the non-uniform cost penalty, as further discussed in Section C, is more effective in reducing the active power losses relative to the case without the TCL resources. As shown in Fig. 6, the effectiveness of the TCL dispatch for reducing the losses is particularly important for large standard deviations of the forecast error that cannot otherwise be dealt with efficiently using traditional controls considered in the CC-OPF formulation. Similarly, it helps improve the voltage profile as the uncertainty of the forecast error increases as shown in Fig. 7.

To further evaluate the effect of the TCL dispatch on the compliance with voltage limits enforced in Eq. (25)-(26), we generate 500 random samples representing the PV outputs and assess the feasibility of the solution obtained by the integrated optimization problem for different values of η_v . This assessment is performed by re-dispatching the obtained solution for each random sample. Fig. 8 presents the statistics

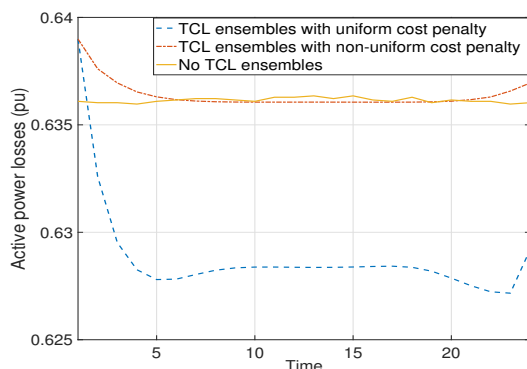


Figure 5. The active power losses for the uniform and non-uniform cost penalty cases, as well as without the dispatch of TCL resources.

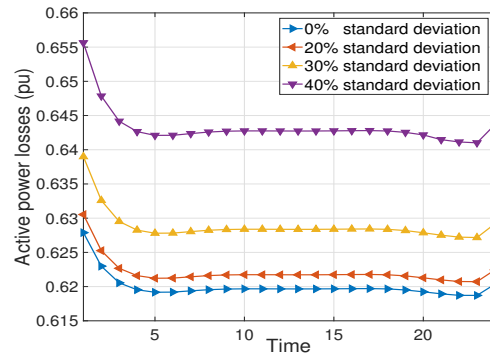


Figure 6. The active power losses for different values of the standard deviation ($\sigma_{t,b}$) on the forecast error with the non-uniform cost penalty.

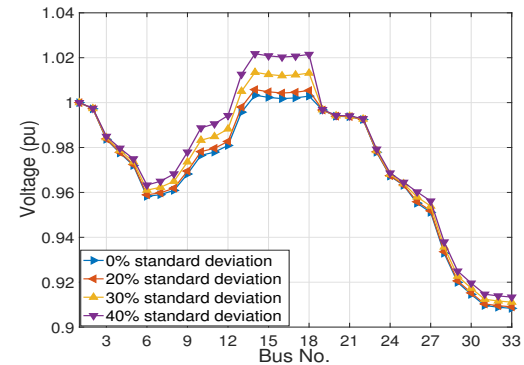


Figure 7. Voltage profile for different values of the standard deviation ($\sigma_{t,b}$) on the forecast error at a given time interval ($t = 24$ hours). The spikes around buses # 14 is due to the injection of the controllable distributed generator.

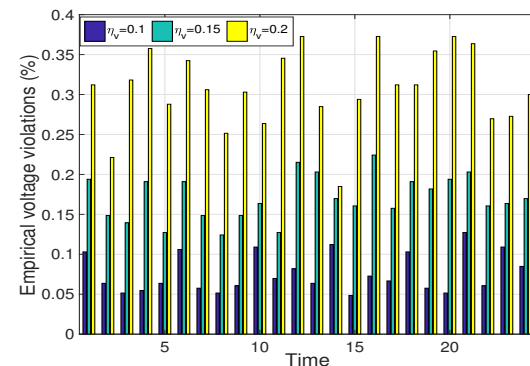


Figure 8. Statistics of the voltage limit violations for different values of η_v over 24 time intervals in 500 randomly generated samples.

on the total number of voltage constraint violations during the optimization horizon for different values of η_v . In all instances observed in Fig. 8 the empirical probability of violation is below the values postulated on η_v in (25)-(26). An improvement in the compliance with voltage limits for tighter values of η_v comes at an incremental increase in the objective function of the DSO ($< 0.1\%$).

C. Perspective of TCL Ensembles

From the perspective of the TCL ensembles, their ability to exercise dispatch flexibility mainly depends on the value of parameter $\gamma_{t,b}^{\alpha\beta}$. Fig. 9 compares the TCL decisions for

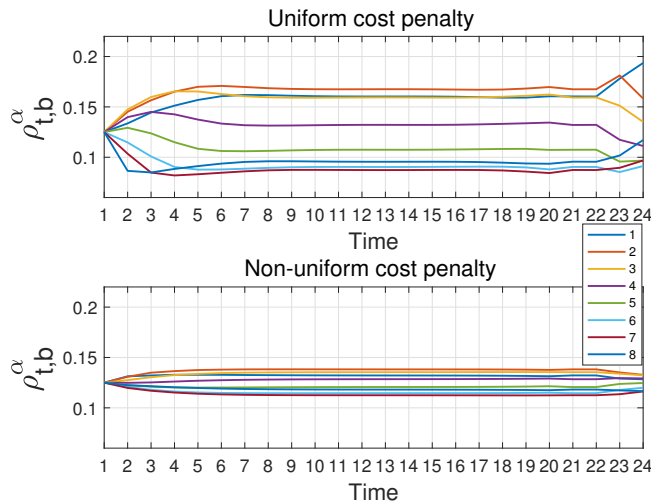


Figure 9. Optimal steady-state probabilities $\rho_{t,b}^{\alpha}$ for different states α at the aggregator on the TCL ensemble at bus # 26 with the uniform and non-uniform cost penalty for the transitions between different states. The legend denotes different states of the TCL ensemble as shown in Fig. 4.

the uniform and non-uniform cost penalty cases. As the penalty $\gamma_{t,b}^{\alpha\beta}$ considered in the non-uniform cost penalty setting weights out-of-cycle transition probabilities higher than next-step probabilities, the non-uniform cost penalty case will return more homogeneous transition decisions $\mathcal{P}_{t,b}^{\alpha\beta}$ of the aggregator during the optimization horizon. As a result, the steady-state probability $\rho_{t,b}^{\alpha}$ for each state will be more homogeneous. The probability masses associated with one-step ahead transitions (largest as shown in Table I) can be exercised at a lower cost compared to the other transitions and hence lead to more homogeneity. In practice, this homogeneity implies that the non-uniform cost penalty case does not lead to drastic changes in the power consumption of the TCL ensemble, as per Eq. (3)-(4), and is thus more suitable for accommodating comfort constraints of TCL users. On the other hand, the uniform cost penalty does not discriminate abrupt changes in the power consumption of the TCL ensemble and leads to more dispersed transitions. The difference between the uniform and non-uniform cost penalty cases presented in Fig. 9 indicates that there is a subtle trade-off between the comfort preferences of the TCL users driven by their default dynamics ($\bar{\mathcal{P}}_{t,b}^{\alpha\beta}$) and their ability to exercise dispatch flexibility. Both parameters $\bar{\mathcal{P}}_{t,b}^{\alpha\beta}$ and $\gamma_{t,b}^{\alpha\beta}$ can be refined within the proposed MDP optimization by using reinforcement learning [4]. We leave it for our future work.

V. CONCLUSION

This paper presents a modeling framework and algorithm to integrate TCL ensembles in PV-dominant distribution systems and co-optimize their dispatch flexibility with the rest of the distribution system resources. The case study demonstrates that the proposed model is capable of leveraging the dispatch flexibility of TCL ensembles to reduce active power losses and maintain nodal voltage magnitudes within an acceptable range. Comparison between the uniform and non-uniform cost

penalty cases reveals that accounting for comfort preferences of TCL users can significantly influence the effect of TCL ensembles on the distribution system. The use of chance constraints on voltage limits also provides a flexible mechanism to address the conservatism of the solution and is effective in reducing violations of voltage limits.

REFERENCES

- [1] R. Hledik and J. Lazar, "Distribution system pricing with distributed energy resources," Tech. Rep., 2016.
- [2] NYISO, "A review of distributed energy resources," 2015. [Online]. Available: <https://goo.gl/tSHzKQ>
- [3] D. S. Callaway and I. A. Hiskens, "Achieving controllability of electric loads," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 184–199, Jan 2011.
- [4] F. Ruelens and et al, "Residential demand response of thermostatically controlled loads using batch reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2149–2159, Sept 2017.
- [5] D. Angeli and P. Kountouriotis, "A stochastic approach to dynamic-demand refrigerator control," *IEEE Tran. Contr. Syst. Tech.*, vol. 20, 2012.
- [6] J. L. Mathieu and D. S. Callaway, "State estimation and control of heterogeneous thermostatically controlled loads for load following," in *2012 45th Hawaii Int. Conf. Syst. Sc.*, Jan 2012, pp. 2002–2011.
- [7] M. Chertkov, V. Y. Cheryak, and D. Deka, "Ensemble control of cycling energy loads: Markov decision approach," vol. abs/1701.04941, 2017.
- [8] A. Bušić and S. Meyn, "Distributed randomized control for demand dispatch," in *2016 IEEE 55th Conf. Dec. Contr.*, Dec 2016, pp. 6964–71.
- [9] D. S. Callaway, "Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy," *En. Conv. Mngmt.*, vol. 50, no. 5, pp. 1389 – 1400, 2009.
- [10] M. Chertkov, D. Deka, and Y. Dvorkin, "Optimal ensemble control of loads in distribution grids with network constraints," in *Pwr. Syst. Comp. Conf. (accepted), Dublin, Ireland*. [Online]. Available: <https://dblp.org/rec/bib/journals/corr/abs-1710-09924>
- [11] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Pwr. Del.*, vol. 4, pp. 735–743, 1989.
- [12] E. Dall'Anese, K. Baker, and T. Summers, "Chance-constrained ac optimal power flow for distribution systems with renewables," *IEEE Trans. Pwr. Syst.*, vol. 32, no. 5, pp. 3427–3438, Sept 2017.
- [13] R. Mieth and Y. Dvorkin, "Data-driven distributionally robust optimal power flow for distribution systems," in *IEEE Contr. Syst. Let. (under review)*, 2018. [Online]. Available: <https://arxiv.org/abs/1803.04912>
- [14] Y. Cao and et al, "Chance-constrained optimization-based unbalanced optimal power flow for radial distribution networks," *IEEE Trans. Pwr. Del.*, vol. 28, no. 3, pp. 1855–1864, July 2013.
- [15] Hassan, Ali et al, "Chance-constrained ADMM approach for decentralized control of distributed energy resources," in *Pwr. Syst. Comp. Conf. (accepted), Dublin, Ireland*, 2018. [Online]. Available: <https://arxiv.org/abs/1710.09924>
- [16] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained optimal power flow: Risk-aware network control under uncertainty," *SIAM Review*, vol. 56, no. 3, pp. 461–495, 2014.
- [17] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, "Analytical reformulation of security constrained optimal power flow with probabilistic constraints," in *2013 IEEE Grenoble Conference*, June 2013, pp. 1–6.
- [18] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Math. Prog.*, vol. 158, no. 1, pp. 291–327, Jul 2016.
- [19] B. Li, M. Vrakopoulou, and J. L. Mathieu, "Chance constrained reserve scheduling using uncertain controllable loads part ii: Analytical reformulation," *IEEE Transactions on Smart Grid*, 2017.
- [20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [21] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, Apr 1989.
- [22] K. P. Schneider, B. Mather, B. C. Pal, C.-W. Ten, G. Shirek, H. Zhu, J. Fuller, J. L. R. Pereira, L. Ochoa, L. Araujo et al., "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Trans. Power Sys.*, 2017.
- [23] E. Todorov, "Linearly-solvable markov decision problems," in *Advances in Neural Information Processing Systems 19*, B. Schölkopf, J. C. Platt, and T. Hoffman, Eds. MIT Press, 2007, pp. 1369–1376.
- [24] K. Dvijotham and E. Todorov, "A unifying framework for linearly solvable control," vol. abs/1202.3715, 2012.
- [25] —, *Linearly Solvable Optimal Control*. Wiley-IEEE, 2013, p. 648.

- [26] S. P. Meyn et al, “Ancillary service to the grid using intelligent deferrable loads,” *IEEE Trans. Aut. Cntr.*, vol. 60, pp. 2847–2862, Nov 2015.
- [27] P. Bacher, H. Madsen, and H. A. Nielsen, “Online short-term solar power forecasting,” *Solar Energy*, vol. 83, no. 10, pp. 1772 – 1783, 2009.
- [28] M. Lubin, Y. Dvorkin, and S. Backhaus, “A robust approach to chance constrained optimal power flow with renewable generation,” *IEEE Trans. Pwr. Syst.*, vol. 31, no. 5, pp. 3840–3849, Sept 2016.
- [29] L. Roald, F. Oldewurtel, B. Van Parys, and G. Andersson, “Security Constrained Optimal Power Flow with Distributionally Robust Chance Constraints,” *ArXiv e-prints*, Aug. 2015.
- [30] Y. Dvorkin, M. Lubin, S. Backhaus, and M. Chertkov, “Uncertainty sets for wind power generation,” *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3326–3327, July 2016.
- [31] M. Lubin, Y. Dvorkin, and R. Line, “Chance constraints for improving the security of ac optimal power flow,” in *IEEE Transactions on Power Systems (under review)*, 2019. [Online]. Available: <https://arxiv.org/abs/1803.04912>
- [32] S. Zymmler, D. Kuhn, and B. Rustem, “Distributionally robust joint chance constraints with second-order moment information,” *Mathematical Programming*, vol. 137, no. 1, pp. 167–198, Feb 2013. [Online]. Available: <https://doi.org/10.1007/s10107-011-0494-7>
- [33] Q. T. Dinh, I. Necoara, and M. Diehl, “A dual decomposition algorithm for separable nonconvex optimization using the penalty function framework,” in *52nd IEEE Conference on Decision and Control*, Dec 2013, pp. 2372–2377.
- [34] P. Šulc, S. Backhaus, and M. Chertkov, “Optimal distributed control of reactive power via the alternating direction method of multipliers,” *IEEE Transactions on Energy Conversion*, vol. 29, no. 4, pp. 968–977, Dec 2014.
- [35] W. Shi, X. Xie, C. C. Chu, and R. Gadh, “Distributed optimal energy management in microgrids,” *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1137–1146, May 2015.
- [36] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011. [Online]. Available: <http://dx.doi.org/10.1561/22000000016>
- [37] I. Dunning, J. Huchette, and M. Lubin, “Jump: A modeling language for mathematical optimization,” *SIAM Rev.*, vol. 59, pp. 295–320, 2017.
- [38] A. Hassan et al, “Code supplement for optimal load ensemble control in chance-constrained optimal power flow,” 2018. [Online]. Available: https://github.com/ahassansan/CC_MDP

can be represented in the following form [20]:

$$\mu^T x + \Phi^{-1}(1 - \epsilon)\sqrt{x^T \Sigma x} \leq b \quad (35)$$

where $\epsilon \in [0, 1/2]$ is a given tolerance to violations and Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution. Eq. (35) is then convex and equivalent to the following SOC constraint [20]:

$$t \geq \left\| \Sigma^{\frac{1}{2}} x \right\|_2 \quad (36)$$

$$\mu^T x + \Phi^{-1}(1 - \epsilon)t \leq b. \quad (37)$$

APPENDIX A

BACKWARD-FORWARD ALGORITHM

We overview the backward-forward algorithm below:

- **Backward in time step.** Starting at $t = |\mathcal{T}|$, solve (1)-(5) for \mathcal{P} recursively backward in time, i.e. $t \rightarrow 1$. This process returns optimal $\mathcal{P}_{t,b}^{\alpha\beta}$ for transitions to all states α from state β at time t given associated cost functions. Each problem can be solved either by a Lagrange relaxation or by minimizing a convex function.
- **Forward in time step.** Reconstruct ρ using the relationship in (2) forward in time, i.e. $t \rightarrow |\mathcal{T}|$, with the initial condition on $\rho_{t=0,b}^{\alpha} = \rho_{in;b}^{\alpha}, \forall \alpha$, where $\rho_{in;b}^{\alpha}$ is given.

Interested readers are referred to Appendix 1.9 in [7] and to [10] for more details.

APPENDIX B

SOC REFORMULATION OF THE CHANCE CONSTRAINTS

Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ be the vector of random variable with the means and variances given by the vector μ and covariance-matrix Σ , respectively, and let b and x be the vectors of parameters and decision variables. The chance constraint of the form:

$$\mathbb{P}(\xi^T x \leq b) \geq 1 - \epsilon \quad (34)$$