A Multiperiod Consensus-Based Transactive Energy System for Unbalanced Distribution Networks

Rui Cheng, *Student Member, IEEE*, Leigh Tesfatsion, *Senior Member, IEEE*, Zhaoyu Wang, *Senior Member, IEEE*

***Abstract*—This study develops a consensus-based transactive energy system design managed by an independent distribution system operator (DSO) for an unbalanced radial distribution network. The network is populated by welfare-maximizing customers with price-sensitive and fixed (non-price-sensitive) demands who make multiple successive power decisions during each real-time operating period OP. The DSO and customers engage in an iterative negotiation process in advance of each OP to determine retail price-to-go sequences for OP that align customer power decisions with network reliability constraints in a manner that respects customer privacy. The convergence properties of a dual decomposition algorithm developed to implement this negotiation process are analytically established. A case study is presented for an unbalanced 123-bus radial distribution network populated by household customers that demonstrates the practical effectiveness of the design.**

***Index Terms*—Transactive energy system, unbalanced radial distribution network, DSO-managed negotiation process, network reliability, DSO-customer decision alignment, customer privacy.**

I. Introduction

THE rapidly growing penetration of distributed energy resources (DERs) poses new challenges for the efficient and reliable management of distribution networks. Researchers and practitioners are exploring a variety of management strategies to meet these challenges.

Transactive energy system (TES) design is an emerging innovative energy management strategy that engages DERs through market interactions. As originally formulated by the GridWise Architecture Council [1], a TES design is a collection of economic and control mechanisms that allows the dynamic balance of power supply and demand across an entire electrical infrastructure using value as the key operational parameter. Typically, valuations for power demands and supplies are expressed by means of purchase and sale prices.

The TES design literature is rapidly expanding. Many different conceptual TES designs are actively under investigation. These designs range from peer-to-peer TES designs based on bilateral customer transactions to TES designs for which customer power requirements are centrally managed by some form of Distribution System Operator (DSO). In addition, researchers have developed software platforms for Integrated Transmission and Distribution (ITD) systems that permit the study of interactions between TES design operations at the distribution level and wholesale power market operations at the transmission level.

Nevertheless, to date, most TES design studies do not carefully take into account network power flow constraints for the empirically relevant case of *unbalanced* distribution networks. Consequently, the studied TES designs cannot ensure the reliable operation of these networks. For example, Nguyen et al. [2] show how voltage violations can arise for an unbalanced distribution network operating under PowerMatcher [3], a well-known TES design that does not explicitly consider the need to satisfy distribution network reliability constraints.

A second TES design issue is that attention is typically focused on the sequential determination of single-period decisions, with no consideration of possible future effects. This single-period focus does not permit decision makers to take into account cross-period correlations among their successive decisions.

A third TES design issue is the desirability of aligning customer goals and constraints with distribution network goals and constraints by means of a decentralized communication process that ensures customer privacy as well as network reliability.[1] To date, the precise form of decentralization that would be needed to achieve this multi-faceted objective for unbalanced distribution networks has not been extensively investigated.

The present study proposes a multiperiod consensus-based TES design that addresses all three of these issues. This TES design can be used by an independent[2] DSO to manage the daily power requirements of customers populating an unbalanced distribution network.

The TES design is *consensus-based* in that retail prices for each operating period OP are determined by means of a negotiation process between the DSO and the customers that preserves customer privacy. The TES design is *multiperiod* in that each operating period OP, of arbitrary duration, is assumed to be partitioned into finitely many successive time-steps; and the negotiation process for OP results in the joint determination of retail prices and customer real and reactive power usage levels for each of these time-steps.

The retail prices for price-sensitive demands that result from the negotiation process between the DSO and the customers have an informative structural form. Each of these prices is the summation of an initial price set by the DSO together

Latest Revision: 4/23/2021. This work has been supported by project award M40 from the Power Systems Engineering Research Center (PSERC). R. Cheng and Z. Wang are with the Department of Electrical & Computer Engineering, Iowa State University, and L. Tesfatsion (corresponding author, tesfatsi@iastate.edu) is with the Department of Economics, Iowa State University, Ames, IA 50011-1054 USA.

[1]The study of institutions mapping private activities into social outcomes by means of decentralized communication processes is referred to as *mechanism design* in the economics literature; see [4], [5].

[2]The qualifier *independent* means that the DSO has no financial or ownership stake in the operations of the distribution network.

with specific price deviations entailed by the need (if any) to ensure that network reliability constraints are met.

The remainder of this study is organized as follows. The relation of this study to previous TES design work is discussed in Section II. The key features of the consensus-based TES design are described in Section III.

An illustration of the consensus-based TES design is formulated analytically in Sections IV through VII for the special case of an unbalanced radial distribution network populated by households. A dual decomposition algorithm implementing the negotiation process for this TES design is developed in Section VIII, and sufficient conditions are established analytically for its convergence to a TES equilibrium ensuring the alignment of DSO and household goals and constraints.

A case study is presented in Section IX to demonstrate the capabilities of the consensus-based TES design. The case study models a DSO-managed unbalanced 123-bus radial distribution network connected to a relatively large transmission network at a single point of connection. The distribution network is populated by household customers with a mixture of thermostatically-controlled load (TCL) and non-TCL, where only TCL is sensitive to price.

Concluding comments are given in Section X. A nomenclature table, plus important technical details regarding unbalanced distribution network modeling, dual decomposition, and proposition proofs, are provided in appendices.

II. Relationship of Current Work to Existing Energy Management Work

As extensively surveyed in ([6], [7]), current energy management strategies can roughly be classified into four categories: top-down switching, centralized optimization; price reaction; and TES design.

A *top-down switching method* for a group of electrical devices, implemented by a utility or other entity, is a method for simultaneously controlling the energy usage of these devices by means of signals that are commonly and simultaneously communicated to each device. Top-down switching methods are easy to implement. However, they cannot fully exploit the response potential of individual electrical devices based on differences in physical attributes and owner preferences.

In contrast, *centralized optimization* ([8], [9]) for an electrical system is the centrally-managed formulation and solution of system-wide optimizations during successive operating periods. The main advantage of centralized optimization is that the manager has more certain control over system outcomes [10]. However, a major disadvantage is that effective centralized optimization can require the violation of customer privacy. Moreover, centralized optimization can entail large amounts of computation time, hindering scalability. For example, the computations required for centralized optimization become extremely challenging when distribution network power flow is explicitly considered; see [11].

A *price-reaction method* is an energy management method based on one-way communication that uses price signals communicated to customers to modify their energy usage patterns ([12], [13]). Price-reaction methods are simple to deploy. However, price-reaction methods can result in reliability problems due to the the difficulty of accurately predicting customer responses to price signals. In addition, price-reaction methods can result in divergent price and quantity outcomes over time [14], a well-known issue referred to in the economics literature as "cobweb dynamics."

A *TES design* is an energy management strategy that uses market mechanisms to ensure the continual balancing of power demands and supplies across an entire electrical infrastructure based on customer purchase and sale valuations ([15], [16], [17]). For example, *peer-to-peer TES designs* [18] posit direct bilateral transactions among design participants with no central management. In contrast, *centrally-managed bid-based TES designs* posit the existence of a central manager that repeatedly communicates prices to power customers for successive operating periods based on bid functions received from these customers that express their updated power demands, power supplies, and/or ancillary service supplies; see, e.g., [19].

The potential advantages of TES design relative to the previous three energy management approaches include the ability to align distribution network goals and constraints with local customer goals and constraints in a tractable scalable manner while respecting customer privacy. These potential advantages have resulted in rapidly expanding TES research efforts and demonstration projects ([20]-[26]).

The study by Hu et al. [27] is closest to our current study. The authors develop a DSO-managed multiperiod TES design based on a negotiation process between the DSO and a collection of aggregators managing the charging schedule for electric vehicle (EV) owners. The DSO and aggregators exchange primal and dual variable information in order to determine a retail energy price sequence for EV charging from a single-phase distribution network, where the resulting EV charging schedule is able to contribute to the support of power balance for a day-ahead market operating over a transmission network connected to the distribution network. The DSO undertakes an optimization to ensure the charging schedule satisfies voltage and peak demand constraints for the distribution network.

Our study differs from Hu et al. [27] in four important regards. First, our proposed TES design is suitable for managing the operations of an *unbalanced* distribution network. Second, our proposed TES design ensures the satisfaction of distribution network constraints without requiring the DSO to solve an optimization problem.

Third, our TES design aligns DSO goals and constraints with customer goals and constraints, where customer goals are explicitly expressed in terms of customer net benefits, and customer constraints are explicitly expressed in terms of customer physical and financial considerations. In contrast, Hu et al. do not consider whether the EV charging schedule determined by their proposed negotiation process is in fact the best possible schedule for EV owners, measured in terms of the goals and constraints of these EV owners.

Fourth, the negotiation process postulated by our TES design for each operating period is based on a more intuitive, readily interpreted exchange of information than in Hu et al. [27]. At the beginning of this negotiation process the DSO is assumed to receive two types of structural information from

each customer: namely, power factor rating information for the customer's price-sensitive demand; and a thermostat slider-knob control setting between 0 and 1 indicating the customer's preferred emphasis on power-usage benefit relative to power-usage cost. Given this information, the DSO iteratively communicates retail price-to-go sequences to customers who in turn indicate their power usage responses. This iterative process comes to a halt once the customers' power usage responses satisfy all distribution network constraints.

III. The General Consensus-Based TES Design

A. Design Overview

Retail customer participation in existing U.S. RTO/ISO-managed wholesale power markets is typically handled by some form of intermediary, such as a Qualified Scheduling Entity (QSE) or a Load-Serving Entity (LSE). However, FERC Order 2222 [28] promotes participation by a broader range of distributed resource aggregators.

This study develops a consensus-based TES design managed by a DSO within an integrated transmission and distribution system. This DSO operates as a linkage agent at a transmission grid bus b^* that connects an unbalanced radial distribution network to a relatively large transmission network. The DSO manages the power usage needs for a collection of customers located across the distribution network who have a mix of fixed and price-sensitive demands. The DSO is an independent entity with a fiduciary responsibility for ensuring the welfare of the customers, conditional on the maintenance of distribution network reliability.

Each operating day is partitioned into a finite number of operating periods OP. Prior to each OP, the DSO engages its customers in a negotiation process that results in retail prices for OP. The objective of this negotiation process is to permit customers to select power usage levels for OP that maximize their net benefit, subject to retail prices and local constraints, in a manner that ensures both the reliability of distribution network operations and customer privacy.

The power usage needs of the customers must be met by power procured from the transmission network.[3] The DSO manages this power procurement on behalf of the customers. All DSO net procurement costs are allocated back to customers on the basis of their relative power usage.

B. Design Timing Relative to Real-Time Market Processes

In advance of each operating period OP, the ISO/RTO conducts a real-time market (RTM) consisting of a security-constrained economic dispatch (SCED) optimization. Hereafter this RTM will be denoted by RTM(OP).

Figure 1 depicts the timing of the consensus-based TES design outlined in Section III-A in relation to RTM(OP). The look-ahead horizon for RTM(OP), denoted by LAH(OP), is the time interval between the close of RTM(OP) and the start of OP. Each operating period OP is assumed to be partitioned into NK sub-periods t. The negotiation process N(OP) between the DSO and its customers to determine retail prices for each sub-period t of OP takes place during LAH(OP).

[3]As noted in Sec. I, distributed generation is not considered in this study.

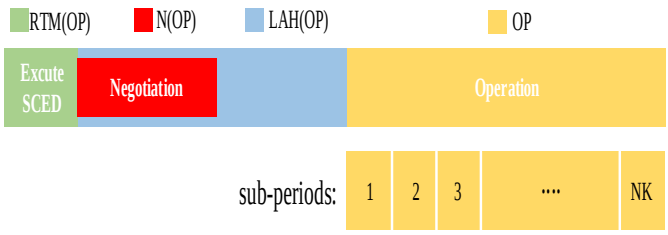

Fig. 1. Timing of the consensus-based TES design in relation to a standard real-time market RTM(OP) for an operating period OP

C. Design Negotiation Process: General Outline

The negotiation process N(OP) for each operating period OP consists of three basic components, as follows:

C1. Initialization: At the start of N(OP), the DSO receives from each customer ψ a power-factor function f_ψ for price-sensitive demand permitting reactive power usage to be determined from real power usage. In addition, the DSO receives from each customer ψ a thermostat slider-knob control setting γ_ψ between 0 ("Benefit") and 1 ("Cost") whose closeness to 0 indicates the degree to which ψ prefers to emphasize the benefit of power usage relative to its cost.[4] The DSO then sets retail prices for all fixed loads during each sub-period t of OP, as well as *initial* retail prices for all price-sensitive demands during each sub-period t of OP, and communicates these prices to its customers.

C2. Adjustment: Upon receipt of prices from the DSO, each customer ψ communicates back to the DSO its optimal price-sensitive real power usage level for each sub-period t of OP. The DSO then determines whether estimated customer real and reactive power usage levels for OP would result in any violation of network reliability constraints. If so, and if the DSO's stopping rule has not been activated, the DSO determines adjusted prices for price-sensitive demands during each sub-period t of OP and communicates these adjusted prices back to its customers to initiate another negotiation round. Otherwise, the DSO terminates the negotiation process.

C3. Stopping Rule: If the negotiation process has not terminated by a designated time prior to the end of LAH(OP), the DSO invokes a standardized procedure to set final prices for customer price-sensitive demands that ensure the reliability of distribution network operations.

D. Design Negotiation Process: Implementation Details

In greater detail, the implementation of our TES design negotiation process N(OP) for any given operating period OP proceeds as follows:

- At the start of RTM(OP), the ISO/RTO submits a forecast to RTM(OP) for load during OP, including distribution network load at the linkage bus b^*. The ISO/RTO then conducts a SCED optimization for RTM(OP), which determines the locational marginal price LMP(b^*, OP) (cents/kWh) at bus b^* for OP.

[4]See Appendix D for a more detailed constructive definition of γ_ψ.

- At the close of RTM(OP), i.e., the start of LAH(OP), the ISO/RTO communicates LMP(b^*, OP) to the DSO. The DSO also receives from each customer ψ a power-factor function f_ψ for price-sensitive demand and a slider-knob control setting γ_ψ.
- At the start of N(OP) during LAH(OP), the DSO communicates to each of its customers that LMP(b^*, OP) will be the price it charges for all fixed load during each sub-period t of OP.
- The DSO then conducts a multi-round negotiation with each customer ψ to determine a retail price $\pi_\psi(t)$ for the price-sensitive demand of customer ψ during each sub-period t of OP, $t = 1, \ldots, NK$.
- At the start of this multi-round negotiation process, the DSO communicates to each customer ψ an *initial retail price-to-go sequence* $\boldsymbol{\pi}^o_\psi(\mathcal{K})$ for price-sensitive demand during OP taking the row-vector form $\boldsymbol{\pi}^o_\psi(\mathcal{K}) = [\pi^o_\psi(1), ..., \pi^o_\psi(NK)]$, where: (i) $\mathcal{K} = (1, \ldots, NK)$ is a partition of OP into sub-periods t; and (ii) $\pi^o_\psi(t) =$ LMP(b^*, OP) for each sub-period t.
- Each customer ψ then communicates back to the DSO its optimal price-sensitive real power usage sequence for OP, conditional on its own local constraints plus the initial retail price-to-go sequence $\boldsymbol{\pi}^o_\psi(\mathcal{K})$ received from the DSO. This permits the DSO to estimate total customer real and reactive power usage sequences for OP.
- The DSO continues to conduct successive negotiation rounds until either its estimated power usage sequences for OP satisfy all network reliability constraints or its stopping rule is activated.
- At the termination of the multi-round negotiation process, each customer ψ implements its optimal real and reactive power usage sequences for OP, conditional on its final received retail price-to-go sequence and its own local constraints.

Important Remark on N(OP) Retail Price Initialization:

In the seven existing U.S. RTO/ISO-managed wholesale power markets, RTM LMPs are measured in \$/MWh. Moreover, the prices charged to QSE/LSE intermediaries for the wholesale power usage of their managed customers are typically based on temporal and/or spatial *averages* of RTM LMP realizations in order to mitigate customer exposure to price volatility.

The presumption in this study that the DSO sets initial retail prices for N(OP) equal to RTM LMPs measured in cents/kWh is made purely for ease of exposition. These initial retail prices could instead be set as averages of RTM LMPs measured in \$/MWh, in conformity with current practices. This change would not have any substantive effect on reported results.

IV. TES Design Illustration: Overview

For concreteness, this section illustrates our proposed DSO-managed consensus-based TES design for the special case of an unbalanced radial distribution network populated by households. More precisely, the distribution network consists of a collection of buses connected by multi-phase line segments; and each household is connected to a particular bus by an external 1-phase line; see Fig. 2.

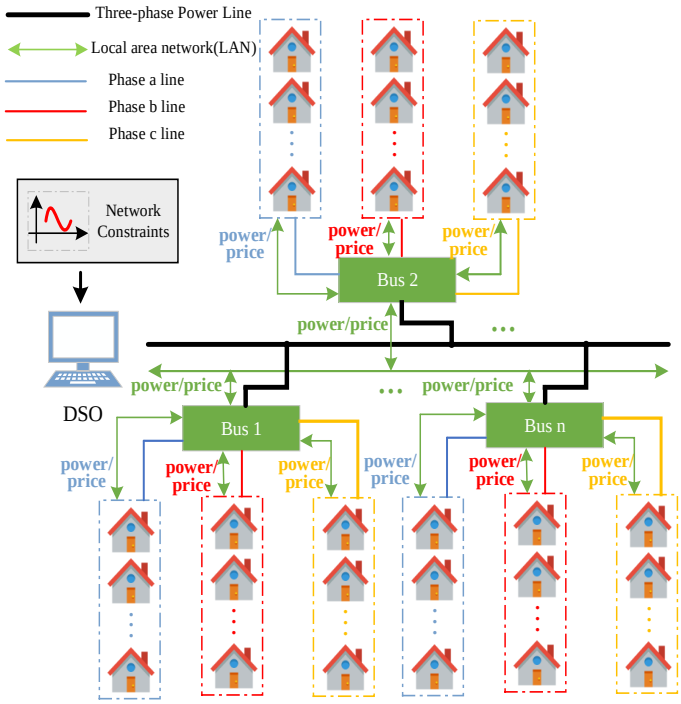

Fig. 2. General features of the household TES design illustration

The goal of each household ψ is to maximize its own net benefit, subject to retail prices and local constraints. Every household owns a mixture of TCL and non-TCL devices. Each non-TCL device is a fixed-load device, meaning its load is not sensitive to charged prices. In contrast, each TCL device is a smartly controlled device with price-sensitive power usage.

Two important restrictions are imposed on household loads for the TES design illustration. First, in the absence of all household TCL, household non-TCL satisfies all network reliability constraints. Second, the price-sensitive TCL of each household reduces to zero at a finite sufficiently-high TCL price; and these sufficiently-high TCL prices are known to the DSO based on historical experience.

Household load for each operating period OP must be serviced by power obtained from a transmission network connected to the distribution network at a linkage bus b^*. ISO-forecasted power deliveries to households during OP are prescheduled in RTM(OP), an ISO-managed RTM operating over the transmission network. The rules governing the operations of RTM(OP) are based on the rules governing RTM operations in the U.S. ERCOT energy region ([29], [30]).

Household power usage for each operating period OP is managed by means of the DSO-managed consensus-based TES design outlined in Section III. The DSO tasked with managing this TES design conducts a negotiation process N(OP) with households during the look-ahead horizon LAH(OP) for RTM(OP). As depicted in Fig.3, this negotiation process N(OP) consists of the following three components:

C1. Initialization: The DSO receives from each household ψ a power-factor function f_ψ for price-sensitive demand and a

slider-knob control setting γ_ψ. The DSO sets the non-TCL retail price for all households during OP equal to LMP(b^*, OP), the LMP determined in RTM(OP) for the linkage bus b^*. The DSO also initially sets the TCL retail price for all households during OP equal to LMP(b^*, OP). The DSO then ensures these non-TCL and TCL retail prices are signaled to the households. Finally, the DSO sets a maximum permitted number $I_{\max}$ of iterations for N(OP) that ensures N(OP) will come to a close prior to the end of LAH(OP).

C2. Adjustment: Each household adjusts its real TCL power usage for OP to maximize its net benefit, conditional on local constraints and its latest received TCL price-to-go sequence for OP. The resulting household real TCL power usage schedules are communicated to the DSO, which permits the DSO to estimate total real and reactive power schedules for OP. If these schedules result in violations of network reliability constraints, the DSO updates the household TCL price-to-go sequences and communicates these updated sequences back to the households.

C3. Stopping Rule: The negotiation process continues until either there are no network reliability constraint violations or the number of iterations reaches $I_{\max}$, whichever comes first. If violations remain after $I_{\max}$ is reached, the DSO sets final household TCL prices for OP to levels that are sufficiently high to reduce household TCL to zero.

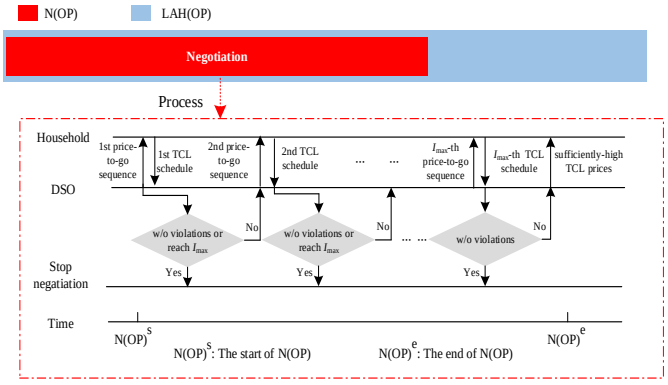

Fig. 3. Negotiation Process N(OP) for the household TES design illustration

V. TES Design Illustration: Network Model

A. Radial network model for a distribution system

Consider a radial network with N+1 buses. Let $\{0\} \bigcup \mathcal{N}$ denote the index set for these buses, where $\mathcal{N} = \{1, 2, ..., N\}$. As shown in Fig. 4, bus 0 is the head bus of the radial network. In addition, bus 0 is the *linkage bus* b^*, i.e., the bus at which the distribution network connects to the transmission network.

By definition of a radial network, each bus located on a radial network can have at most one bus that is an immediate predecessor, measured relative to the feeder head. For each bus $j \in \mathcal{N}$, let $BP(j) \in \{0\} \bigcup \mathcal{N}$ denote the bus that immediately precedes bus j along the radial network headed by bus 0. Thus, as seen in Fig. 4, the distance between bus $BP(j)$ and bus 0 is strictly smaller than the distance between bus j and bus 0.

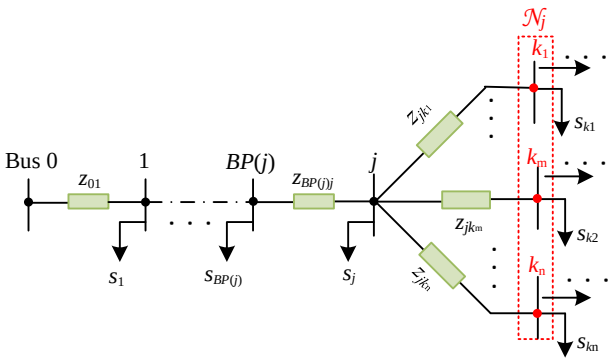

Fig. 4. The radial network for the household TES design illustration from the vantage point of a bus $j \in \mathcal{N}$

Also, let $\mathcal{N}_j$ denote the set of all buses located strictly after bus j along the radial network. For example, in Fig. 4, the depicted buses $k_1, k_2, ..., k_n$ comprise all buses that strictly follow bus j. Thus, $\mathcal{N}_j = \{k_1, k_2, ..., k_n\}$ for this case.

Finally, the edge set for any radial network with $N+1$ buses consists of N distinct *line segments* connecting pairs of adjacent buses along the radial network. More precisely, this edge set is given by $\mathcal{L} = \{\ell_j = (i,j) | i = BP(j), j \in \mathcal{N}\}$, i.e., the collection of all line segments connecting bus $BP(j)$ to bus j for each $j \in \mathcal{N}$. For an unbalanced radial network, each line segment can consist of single-phase, two-phase, or three-phase circuits. Hereafter, such line segments are simply referred to as 1-phase, 2-phase, and 3-phase line segments, respectively.

B. Single-phase radial network

Consider a radial network consisting of 1-phase line segments with a common phase. For each bus $i \in \{0\} \bigcup \mathcal{N}$, let $V_i(t)$ denote its voltage magnitude, let $v_i(t) = |V_i(t)|^2$ denote its squared voltage magnitude, and let $p_i(t)$ and $q_i(t)$ denote its active and reactive bus load, all measured per unit (p.u.). Also, for each line segment $(i,j) \in \mathcal{L}$, let $z_{ij} = r_{ij} + jx_{ij}$ denote its line impedance, and let $P_{ij}(t)$ and $Q_{ij}(t)$ denote the real and reactive power flow from bus i to j, respectively, all measured per unit (p.u.).

The following *Linearized Distribution Flow (LinDistFlow)* equations,[5] adapted from [31] and [32], model the single-phase radial network power flow. For all $(i,j) \in \mathcal{L}$,

$$P_{ij}(t) = \sum_{k \in \mathcal{N}_j} P_{jk}(t) + p_j(t) \tag{1a}$$

$$Q_{ij}(t) = \sum_{k \in \mathcal{N}_j} Q_{jk}(t) + q_j(t) \tag{1b}$$

$$v_i(t) = v_j(t) + 2 \cdot \left(P_{ij}(t) r_{ij} + Q_{ij}(t) x_{ij} \right) \tag{1c}$$

C. Extension to an unbalanced radial network

Section V-B focuses on the case of a single-phase radial network for a distribution system. However, distribution networks

[5]The derivation of the LinDisFlow power flow equations assumes that the power loss on each radial network line segment is negligible relative to the power flow on this line segment.

are typically unbalanced as well as radial, with multi-phase line segments. The common form of a 3-phase line segment for a radial distribution network is shown in Fig.5.

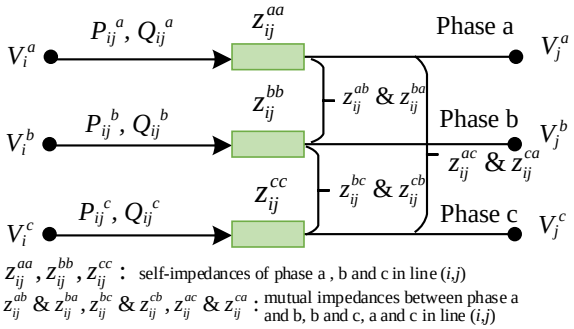

Fig. 5. Depiction of a 3-phase line segment in a radial network.

Thus, making use of previous work ([11], [33],[34]), this section extends the LinDistFlow power flow equations (1a)-(1c) to handle the case of an unbalanced radial distribution network. More precisely, as will be seen below, this extended model assumes: (i) unbalanced phases $\{a, b, c\}$ for which the successive phase differences are given by $\frac{2}{3}\pi$; and (ii) voltage magnitudes across phases are approximately the same. This results in approximately balanced 3-phase voltages. Without loss of generality,[6] each line segment in the network edge set $\mathcal{L}$ is assumed to be a 3-phase line segment.

The following column vectors and matrices are used to depict squared 3-phase voltage magnitudes, bus active and reactive loads, real and reactive power flows over line segments, and line segment impedance values for this unbalanced radial network, all measured per unit (p.u.): For each $(i, j) \in \mathcal{L}$,

$$\begin{aligned}
\boldsymbol{v}_i(t) &= [v_i^a(t), v_i^b(t), v_i^c(t)]^T \\
\boldsymbol{p}_i(t) &= [p_i^a(t), p_i^b(t), p_i^c(t)]^T \\
\boldsymbol{q}_i(t) &= [q_i^a(t), q_i^b(t), q_i^c(t)]^T \\
\boldsymbol{P}_{ij}(t) &= [P_{ij}^a(t), P_{ij}^b(t), P_{ij}^c(t)]^T \\
\boldsymbol{Q}_{ij}(t) &= [Q_{ij}^a(t), Q_{ij}^b(t), Q_{ij}^c(t)]^T \\
\boldsymbol{Z}_{ij} = \boldsymbol{R}_{ij} + j\boldsymbol{X}_{ij} &= \begin{bmatrix} z_{ij}^{aa} & z_{ij}^{ab} & z_{ij}^{ac} \\ z_{ij}^{ba} & z_{ij}^{bb} & z_{ij}^{bc} \\ z_{ij}^{ca} & z_{ij}^{cb} & z_{ij}^{cc} \end{bmatrix}
\end{aligned}$$

where $\boldsymbol{Z}_{ij} \in \mathbb{C}^{3\times 3}$ is a symmetric matrix.

Using this notation, the extended LinDistFlow model equations can be expressed as follows. For all $(i, j) \in \mathcal{L}$,

$$\boldsymbol{P}_{ij}(t) = \sum_{k\in\mathcal{N}_j} \boldsymbol{P}_{jk}(t) + \boldsymbol{p}_j(t) \tag{2a}$$

$$\boldsymbol{Q}_{ij}(t) = \sum_{k\in\mathcal{N}_j} \boldsymbol{Q}_{jk}(t) + \boldsymbol{q}_j(t) \tag{2b}$$

$$\boldsymbol{v}_i(t) = \boldsymbol{v}_j(t) + 2(\bar{\boldsymbol{R}}_{ij}\boldsymbol{P}_{ij}(t) + \bar{\boldsymbol{X}}_{ij}\boldsymbol{Q}_{ij}(t)) \tag{2c}$$

[6]As discussed and illustrated in Appendix B, any k-phase line segment in the network edge set $\mathcal{L}$ with $k < 3$ can be represented as a 3-phase line segment by introducing an appropriate number of additional "virtual" circuits for this line segment with "virtual" phases whose self-impedance and mutual impedance are set to zero. The introduction of these virtual elements does not affect the resulting power flow solutions.

where

$$\begin{aligned}
\boldsymbol{a} &= [1, e^{-\mathrm{j}2\pi/3}, e^{\mathrm{j}2\pi/3}]^T \\
\bar{\boldsymbol{R}}_{ij} &= Re(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{R}_{ij} + Im(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{X}_{ij} \\
\bar{\boldsymbol{X}}_{ij} &= Re(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{X}_{ij} - Im(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{R}_{ij}
\end{aligned}$$

$\odot$ denotes element-wise multiplication

$\boldsymbol{a}^H$ denotes the conjugate transpose of $\boldsymbol{a}$

D. Representation of Unbalanced Radial Networks

Consider, first, the standard matrix representation $\bar{\boldsymbol{M}} = [\boldsymbol{m}_0, \boldsymbol{M}^T]^T \in \mathbb{R}^{(N+1)\times N}$ for the incidence matrix of a single-phase radial network. The rows of this matrix correspond to the buses i in the bus set $\{0\} \bigcup \mathcal{N}$, ordered from lowest to highest i value. The columns of this matrix correspond to the line segments ℓ_j in the edge-set $\mathcal{L}$, ordered from lowest to highest j value. The entries of the matrix indicate, for each bus and line segment, whether or not the bus is a "from" node or a "to" node for this line segment.

More precisely, the incidence matrix $\bar{\boldsymbol{M}}$ with an entry 1 for each "from" node and -1 for each "to" node takes the following form:

$$\bar{\boldsymbol{M}} = \begin{bmatrix} \mathbb{J}(0,\ell_1) & \mathbb{J}(0,\ell_2) & \dots & \mathbb{J}(0,\ell_N) \\ \mathbb{J}(1,\ell_1) & \mathbb{J}(1,\ell_2) & \dots & \mathbb{J}(1,\ell_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{J}(N,\ell_1) & \mathbb{J}(N,\ell_2) & \dots & \mathbb{J}(N,\ell_N) \end{bmatrix} \tag{4}$$

where $\mathbb{J}(\cdot)$ is an indicator function defined as

$$\mathbb{J}(i,\ell_j) = \begin{cases} 1 & \text{if } i = BP(j) \\ -1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

The first row $\boldsymbol{m}_0^T$ of the matrix $\bar{\boldsymbol{M}}$ represents the connection structure between bus 0 and the line segments in $\mathcal{L}$; the remaining submatrix, denoted by $\boldsymbol{M}$, represents the connection structure between the remaining buses and the line segments in $\mathcal{L}$. Since $\boldsymbol{M}$ is a square matrix with full rank [35], it is an invertible matrix. A numerical example illustrating the construction of $\bar{\boldsymbol{M}}$ for a single-phase radial network is given in Appendix C.

Next consider, instead, an unbalanced radial network with a bus set $\{0\} \bigcup \mathcal{N}$ and edge set $\mathcal{L}$ for which each line segment $\ell_j \in \mathcal{L}$ is a 3-phase line.[7] An extended incidence matrix $\bar{\boldsymbol{A}} = [\boldsymbol{A}_0, \boldsymbol{A}^T]^T \in \mathbb{R}^{3(N+1)\times 3N}$ for this unbalanced radial network is constructed as follows:

$$\bar{\boldsymbol{A}} = \bar{\boldsymbol{M}} \otimes \boldsymbol{I}_3 = \begin{bmatrix} \mathbb{J}(0,\ell_1)\boldsymbol{I}_3 & \mathbb{J}(0,\ell_2)\boldsymbol{I}_3 & \dots & \mathbb{J}(0,\ell_N)\boldsymbol{I}_3 \\ \mathbb{J}(1,\ell_1)\boldsymbol{I}_3 & \mathbb{J}(1,\ell_2)\boldsymbol{I}_3 & \dots & \mathbb{J}(1,\ell_N)\boldsymbol{I}_3 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{J}(N,\ell_1)\boldsymbol{I}_3 & \mathbb{J}(N,\ell_2)\boldsymbol{I}_3 & \dots & \mathbb{J}(N,\ell_N)\boldsymbol{I}_3 \end{bmatrix} \tag{5}$$

[7]As previously noted in footnote 6, for the purposes of the current study any unbalanced radial network whose edges consist of a mix of 1-phase, 2-phase, and 3-phase line segments can equivalently be represented as an unbalanced radial network consisting entirely of 3-phase line segments.

where

$$\begin{aligned} \boldsymbol{I}_3 &= 3 \times 3 \text{ identity matrix;} \\ \otimes &= \text{Kronecker product operation.} \end{aligned}$$

In (5), the term $\mathbb{J}(i, \ell_j)\boldsymbol{I}_3 \in \mathbb{R}^{3\times 3}$ represents the 3-phase connection structure between bus i and line segment ℓ_j. The submatrix $\boldsymbol{A}_0^T \in \mathbb{R}^{3\times 3N}$ represents the 3-phase connection structure between the feeder head bus 0 and each of the line segments in $\mathcal{L}$. Finally, the submatrix $\boldsymbol{A} \in \mathbb{R}^{3N\times 3N}$ represents the 3-phase connection structure between each remaining bus i and the line segments in $\mathcal{L}$. A numerical example illustrating the construction of $\bar{\boldsymbol{A}}$ for an unbalanced radial network is given in Appendix C.

The squared voltage magnitudes, real and reactive loads, and real and reactive power flows over line segments for the unbalanced radial distribution network are compactly denoted by the following column vectors:[8]

$$\begin{aligned} \boldsymbol{v}(t) &= [\boldsymbol{v}_1(t)^T, \boldsymbol{v}_2(t)^T, ..., \boldsymbol{v}_N(t)^T]^T \\ \boldsymbol{p}(t) &= [\boldsymbol{p}_1(t)^T, \boldsymbol{p}_2(t)^T, ..., \boldsymbol{p}_N(t)^T]^T \\ \boldsymbol{q}(t) &= [\boldsymbol{q}_1(t)^T, \boldsymbol{q}_2(t)^T, ..., \boldsymbol{q}_N(t)^T]^T \\ \boldsymbol{P}(t) &= [\boldsymbol{P}_{BP(1)1}(t)^T, ..., \boldsymbol{P}_{BP(N)N}(t)^T]^T \\ \boldsymbol{Q}(t) &= [\boldsymbol{Q}_{BP(1)1}(t)^T, ..., \boldsymbol{Q}_{BP(N)N}(t)^T]^T \end{aligned}$$

Resistances and reactances for the line segments in the edge-set $\mathcal{L}$ are compactly denoted by $3N \times 3N$ block diagonal matrices $\boldsymbol{D}_r$ and $\boldsymbol{D}_x$ such that the main-diagonal blocks are 3×3 square matrices and all off-diagonal blocks are zero matrices:

$$\begin{aligned} \boldsymbol{D}_r &= \text{diag}(\bar{\boldsymbol{R}}_{BP(1)1}, ..., \bar{\boldsymbol{R}}_{BP(N)N}) \\ \boldsymbol{D}_x &= \text{diag}(\bar{\boldsymbol{X}}_{BP(1)1}, ..., \bar{\boldsymbol{X}}_{BP(N)N}) \end{aligned}$$

The j-th main-diagonal blocks of $\boldsymbol{D}_r$ and $\boldsymbol{D}_r$ are $\bar{\boldsymbol{R}}_{BP(j)j}$ and $\bar{\boldsymbol{X}}_{BP(j)j}$ corresponding to a particular line segment $\ell_j \in \mathcal{L}$. The squared voltage magnitudes for bus 0 in the unbalanced radial distribution network are represented by $\boldsymbol{v}_0 = [v_0^a, v_0^b, v_0^c]^T$.

Given these notational conventions, the LinDistFlow equations (2a)-(2c) for an unbalanced radial distribution network can be compactly expressed as follows:

$$\boldsymbol{A}\boldsymbol{P}(t) = -\boldsymbol{p}(t) \tag{6a}$$

$$\boldsymbol{A}\boldsymbol{Q}(t) = -\boldsymbol{q}(t) \tag{6b}$$

$$\begin{bmatrix} \boldsymbol{A}_0 & \boldsymbol{A}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_0(t) \\ \boldsymbol{v}(t) \end{bmatrix} = 2\Big(\boldsymbol{D}_r\boldsymbol{P}(t) + \boldsymbol{D}_x\boldsymbol{Q}(t)\Big) \tag{6c}$$

Since the matrix $\boldsymbol{M}^T$ is invertible, the matrix $\boldsymbol{A}^T$ is also invertible. Substituting (6a) and (6b) into (6c), it is seen that the LinDistFLow formulation (6) can equivalently be expressed as

$$\boldsymbol{v}(t) = -[\boldsymbol{A}^T]^{-1}\boldsymbol{A}_0\boldsymbol{v}_0(t) - 2\boldsymbol{R}_D\boldsymbol{p}(t) - 2\boldsymbol{X}_D\boldsymbol{q}(t) \tag{7a}$$

$$\boldsymbol{R}_D = [\boldsymbol{A}^T]^{-1}\boldsymbol{D}_r\boldsymbol{A}^{-1} \tag{7b}$$

$$\boldsymbol{X}_D = [\boldsymbol{A}^T]^{-1}\boldsymbol{D}_x\boldsymbol{A}^{-1} \tag{7c}$$

[8]The squared voltage magnitudes and the real and reactive loads at buses i are sorted in accordance with the ordering of these buses from small to large values of i. The real and reactive power flows over line segments ℓ_j are sorted in accordance with the ordering of these line segments from small to large values of j.

VI. TES Design Illustration: Household Model

For ease of notation, let $\psi = (u, \phi, i)$ denote a household with structural and preference attributes u located on an external 1-phase line with phase $\phi \in \Phi = \{a, b, c\}$, where this 1-phase line is connected to the distribution network at bus $i \in \mathcal{N}$; see Fig. 2. Recall from Section III-B that the operating period OP is assumed to be partitioned into a sequence $\mathcal{K}$ = $(1, \ldots, NK)$ of NK successive sub-periods t.

To determine the optimal TCL power usage sequence for OP at the start of the look-ahead horizon LAH(OP), each household ψ needs to update its forecast for inside air temperature $T_\psi(0)$ (^oF) at the start of OP as well as its forecast for ambient outside air temperature $T_o(t)$ (^oF) at the start of each sub-period $t \in \mathcal{K}$. In addition, each household ψ also needs to forecast its real and reactive non-TCL power usage levels $p_\psi^{\text{non}}(t)$ and $q_\psi^{\text{non}}(t)$ (p.u.) for each sub-period $t \in \mathcal{K}$. The sequences of real and reactive non-TCL power usage levels for a household ψ during OP are denoted by the following $NK \times 1$ column vectors:

$$\begin{aligned} \mathcal{P}_\psi^{\text{non}}(\mathcal{K}) &= [p_\psi^{\text{non}}(1), ..., p_\psi^{\text{non}}(NK)]^T \\ \mathcal{Q}_\psi^{\text{non}}(\mathcal{K}) &= [q_\psi^{\text{non}}(1), ..., q_\psi^{\text{non}}(NK)]^T \end{aligned}$$

The goal of each household ψ at the beginning of operating period OP is to maximize its net benefit attained during OP. This net benefit is assumed to take the general form:

$$\text{NetBen}_\psi = \text{Comfort}_\psi - \mu_\psi[\text{Electricity Cost}]_\psi \tag{8}$$

In (8), Comfort_ψ (utils) denotes the benefit (thermal comfort) attained by household ψ from TCL power usage during OP, and Electricity Cost (cents) denotes the cost incurred by household ψ for TCL power usage during OP.[9] Finally, the benefit/cost trade-off parameter μ_ψ (utils/cent) denotes household ψ's *marginal utility of money* for OP, roughly defined to be the loss of benefit (utils) experienced by ψ for each cent increase in its electricity cost.[10] Here it is assumed that μ_ψ takes the explicit form

$$\mu_\psi = \frac{\gamma_\psi}{1-\gamma_\psi} \times \frac{1 \text{ util}}{1 \text{ cent}} \tag{9}$$

where $\gamma_\psi \in (0, 1)$ denotes the benefit/cost slider-knob control setting communicated to the DSO by household ψ at the start of the negotiation process N(OP).[11]

Let $p_\psi(t)$ (p.u.) and $q_\psi(t)$ (p.u.) denote household ψ's real and reactive TCL power usage levels during any sub-period $t \in \mathcal{K}$. Also, let the sequences of real and reactive TCL power usage levels for a household ψ during OP be denoted by the following $NK \times 1$ column vectors:

$$\mathcal{P}_\psi(\mathcal{K}) = [p_\psi(1), ..., p_\psi(NK)]^T$$

[9]Recall that the non-TCL power usage of each household ψ during each operating period OP is assumed to be fixed, i.e., not price sensitive. Consequently, the inclusion in (8) of the benefit and cost of non-TCL power usage would not affect the solution to household ψ's net benefit maximization problem. The benefit and cost of non-TCL power usage is therefore omitted for ease of exposition.

[10]In economics, marginal utility of money valuations for customers are formally expressed as the dual variable solutions for budget constraints in customer utility maximization problems.

[11]See Appendix D for a constructive definition of γ_ψ and additional discussion regarding the determination of μ_ψ from γ_ψ.

$$\mathcal{Q}_\psi(\mathcal{K}) = [q_\psi(1), ..., q_\psi(NK)]^T$$

The *discomfort* (utils) experienced by ψ during any sub-period $t \in \mathcal{K}$ is measured by the discrepancy between ψ's inside air temperature $T_\psi(p_\psi(t))$ (^oF) at the end of sub-period t and the *bliss temperature* TB_ψ (^oF) at which ψ attains maximum thermal comfort, multiplied by a conversion factor c_ψ $\big(\text{utils}/(^oF)^2\big)$. The *comfort* $u_\psi(p_\psi(t))$ (utils) attained by ψ during t is then measured as the deviation between ψ's maximum attainable comfort u_ψ^{max} (utils) and ψ's discomfort:

$$u_\psi(p_\psi(t), t) = u_\psi^{\mathsf{max}} - c_\psi\big[T_\psi(p_\psi(t), t) - TB_\psi\big]^2 \tag{10}$$

The total comfort (utils) attained by ψ during $\mathcal{K}$ is then given by

$$U_\psi(\mathcal{P}_\psi(\mathcal{K})) = \sum_{t\in\mathcal{K}} u_\psi(p_\psi(t), t) \tag{11}$$

As explained in Section III, the prices (cents/kWh) charged to ψ for its real TCL power usage $\mathcal{P}_\psi(\mathcal{K})$ during $\mathcal{K}$ are given by the price sequence $\boldsymbol{\pi}_\psi(\mathcal{K})$ determined by the negotiation process N(OP) between ψ and the DSO conducted during LAH(OP). Finally, let Δt denote the length of each sub-period t measured in hourly units; and let S_{base} denote the base power (kW) used to convert real power levels (kW) into per unit (p.u.) power levels by simple division.

Using the above notational conventions, the optimization problem of a household ψ for operating period OP can be expressed as follows:

$$\max_{\mathcal{P}_\psi(\mathcal{K})} \Big[U_\psi(\mathcal{P}_\psi(\mathcal{K})) - \mu_\psi \boldsymbol{\pi}_\psi(\mathcal{K})\mathcal{P}_\psi(\mathcal{K})S_{\mathsf{base}}\Delta t\Big] \tag{12}$$

subject to:

$$\begin{aligned} T_\psi(p_\psi(1), 1) \;&=\; \alpha_\psi^H T_\psi(0) \pm \alpha_\psi^P p_\psi(1) S_{\mathsf{base}}\Delta t \\ &\quad + (1-\alpha_\psi^H)T_o(1)\;; && \text{(13a)} \\ T_\psi(p_\psi(t), t) \;&=\; \alpha_\psi^H T_\psi(p_\psi(t-1), t-1) \\ &\quad \pm \alpha_\psi^P p_\psi(t) S_{\mathsf{base}}\Delta t \\ &\quad + (1-\alpha_\psi^H)T_o(t), \quad \forall t \in \mathcal{K}\backslash\{1\}\;; && \text{(13b)} \\ 0 \;\le\; p_\psi(t) \;&\le\; p_\psi^{\mathsf{max}}, \quad \forall t \in \mathcal{K} && \text{(13c)} \end{aligned}$$

Constraints (13a)-(13b) make use of a discrete linearized thermal model, adapted from [36]-[38], to model the fluctuation in household ψ's inside air temperature $T_\psi(t)$ during the successive sub-periods $t \in \mathcal{K}$.[12] For simplicity of exposition, the exogenously given ambient outside air temperature $T_o(t)$ for each sub-period t is assumed to be the same for each household ψ. The parameters α_ψ^H (unit-free) and α_ψ^P (^oF/kWh) are assumed to be positively valued. Constraint (13c) imposes an upper limit p_ψ^{max} (p.u.) on household ψ's real TCL power usage, assumed to represent the rated real power (p.u.) for household ψ's TCL devices.

Clearly, an optimal solution $\mathcal{P}_\psi(\mathcal{K})$ for household ψ's optimization problem (12) depends on the negotiated price-to-go sequence $\boldsymbol{\pi}_\psi(\mathcal{K})$. Let

$$\mathcal{X}_\psi(\mathcal{K}) = \{\mathcal{P}_\psi(\mathcal{K}) \in \mathbb{R}^{NK} | \mathcal{P}_\psi(\mathcal{K}) \text{ satisfies (13)}\} \tag{14}$$

[12]Temperature fluctuation, given by the terms preceded by the symbol $\pm$ in (13a) and (13b), takes a '+' sign for heating and a '-' sign for cooling.

An optimal solution for (12), given $\boldsymbol{\pi}_\psi(\mathcal{K})$, can then be expressed as follows:

$$\begin{aligned} \mathcal{P}_\psi(\boldsymbol{\pi}_\psi(\mathcal{K})) \;\in\; \underset{\mathcal{P}_\psi(\mathcal{K})\in\mathcal{X}_\psi(\mathcal{K})}{\operatorname{argmax}} \Big[&U_\psi(\mathcal{P}_\psi(\mathcal{K})) \\ &- \mu_\psi \boldsymbol{\pi}_\psi(\mathcal{K})\mathcal{P}_\psi(\mathcal{K})S_{\mathsf{base}}\Delta t\Big] \end{aligned} \tag{15}$$

The TCL devices owned by each household ψ are assumed to operate at a constant positive power factor $\text{PF}_\psi(t)$ (unit free) for each $t \in \mathcal{K}$.[13] Given this assumption, the TCL power-factor function f_ψ for each household ψ can be expressed as a collection $\{f_{\psi,t} | t \in \mathcal{K}\}$ of *TCL power-ratio functions* $f_{\psi,t}$ taking the linear form

$$q \;=\; f_{\psi,t}(p) \;=\; \eta_\psi(t)p\;,\; \forall t \in \mathcal{K} \tag{16}$$

where

$$\eta_\psi(t) \;=\; \sqrt{\frac{1}{[\text{PF}_\psi(t)]^2} - 1} \tag{17}$$

Note, by construction, that the unit-free coefficient $\eta_\psi(t)$ defined by (17) is non-negatively valued. Finally, let $\boldsymbol{H}_\psi(\mathcal{K})$ denote the *TCL power-ratio matrix* for household ψ for operating period OP; this $NK \times NK$ matrix is defined as follows:

$$\boldsymbol{H}_\psi(\mathcal{K}) = \text{diag}\big(\eta_\psi(1), \eta_\psi(2), ..., \eta_\psi(NK)\big) \tag{18}$$

VII. TES Design Illustration: DSO Modeling

A. DSO Modeling: Overview

Consider an unbalanced radial distribution network populated by households, as modeled in Sections V and VI. The independent DSO that is tasked with managing a consensus-based TES design for this distribution network operates at the linkage bus b^*, which is assumed to be the head bus 0 of the radial network.

As depicted in Fig. 1, each operating period OP during an operating day D is proceeded in time by a real-time market RTM(OP) followed by a look-ahead horizon LAH(OP). The negotiation process N(OP) between the DSO and the distribution system households takes place during LAH(OP).

The general goal of the DSO is to maximize household net benefit during OP, subject to network reliability constraints and the maintenance of household privacy. Since the distribution network has no distributed generation, the real and reactive power usage of the households must be serviced by wholesale power that is delivered at b^* in stepped-down voltage form. The price charged to the DSO for wholesale power is assumed to be $\text{LMP}(b^*, \text{OP})$ (cents/kWh), the locational marginal price determined in RTM(OP) for OP at b^*.

[13] Given a TCL real power level $p(\tau) > 0$ and a TCL reactive power level $q(\tau)$ at a time-point τ, the TCL power factor pf(τ) at τ is defined to be the ratio $p(\tau)/\sqrt{p(\tau)^2 + q(\tau)^2}$ in (0,1]. See Appendix E for further discussion of the assumption of a constant TCL power factor PF(t) for each sub-period t; i.e, pf(τ) = PF(t) at all time points $\tau \in t$.

B. DSO Optimization Problem in Centralized Form

Recall that each household $\psi \in \Psi$ is characterized by a vector $\psi = (u, \phi, i)$, where u denotes the households structural and physical attributes, and (ϕ, i) indicates the household is located on an external phase-ϕ line connected to the distribution network at bus i. Let $\mathcal{U}_{i,\phi}$ denote the set of all household attributes u such that (u, ϕ, i) denotes a household $\psi \in \Psi$. For each $i \in \mathcal{N}$ and $\phi \in \Phi$, let $p_i^{\phi}(t)$ and $q_i^{\phi}(t)$ denote the real and reactive load for phase ϕ at bus i during sub-period t:

$$p_i^{\phi}(t) = \sum_{u \in \mathcal{U}_{i,\phi}} [p_{\psi}(t) + p_{\psi}^{\mathsf{non}}(t)],\ \forall i \in \mathcal{N},\ \forall \phi \in \Phi \quad (19a)$$

$$q_i^{\phi}(t) = \sum_{u \in \mathcal{U}_{i,\phi}} [q_{\psi}(t) + q_{\psi}^{\mathsf{non}}(t)],\ \forall i \in \mathcal{N},\ \forall \phi \in \Phi \quad (19b)$$

Using the column vector expressions $\boldsymbol{p}_i(t)$, $\boldsymbol{q}_i(t)$, $\boldsymbol{p}(t)$, and $\boldsymbol{q}(t)$ given in Sections V-C and V-D, together with (16) and (19), it is seen that the power flow equation (7a) can equivalently be expressed as follows for any sub-period $t \in \mathcal{K}$:

$$\boldsymbol{v}(t, \boldsymbol{p}_{\Psi}(t)) = \boldsymbol{v}^{\mathsf{non}}(t) - 2\boldsymbol{s}(t, \boldsymbol{p}_{\Psi}(t)) \quad (20)$$

where

$$\boldsymbol{p}_{\Psi}(t) = \{p_{\psi}(t) \mid \psi \in \Psi\}$$
$$\boldsymbol{s}(t, \boldsymbol{p}_{\Psi}(t)) = \sum_{\psi \in \Psi} \big[\boldsymbol{h}_{\psi}(t, p_{\psi}(t))\big]$$
$$\boldsymbol{h}_{\psi}(t, p_{\psi}(t)) = \boldsymbol{r}_D(i, N_{\psi}^{\mathsf{ph}}) p_{\psi}(t) + \boldsymbol{x}_D(i, N_{\psi}^{\mathsf{ph}}) \eta_{\psi}(t) p_{\psi}(t)$$
$$N_{\psi}^{\mathsf{ph}} = \begin{cases} 1 & \text{if household } \psi \text{ connects to phase a} \\ 2 & \text{if household } \psi \text{ connects to phase b} \\ 3 & \text{if household } \psi \text{ connects to phase c} \end{cases}$$
$$\boldsymbol{v}^{\mathsf{non}}(t) = -[\boldsymbol{A}^T]^{-1} \boldsymbol{A}_0 \boldsymbol{v}_0(t) - 2\boldsymbol{s}^{\mathsf{non}}(t)$$
$$\boldsymbol{s}^{\mathsf{non}}(t) = \sum_{\psi \in \Psi} \big[\boldsymbol{r}_D(i, N_{\psi}^{\mathsf{ph}}) p_{\psi}^{\mathsf{non}}(t) + \boldsymbol{x}_D(i, N_{\psi}^{\mathsf{ph}}) q_{\psi}^{\mathsf{non}}(t)\big]$$

In (20), the $3N \times 1$ column vector $\boldsymbol{v}^{\mathsf{non}}(t)$ consists of the 3-phase squared voltage magnitudes for t at all non-head buses, assuming zero TCL; and the $3N \times 1$ column vector $\boldsymbol{v}_0(t)$ consists of the 3-phase squared voltage magnitudes for t at the head bus 0. Also, $\psi = (u, \phi, i)$ is the generic term for a household in the household set Ψ, and $\boldsymbol{r}_D(i, N_{\psi}^{\mathsf{ph}})$ and $\boldsymbol{x}_D(i, N_{\psi}^{\mathsf{ph}})$ are $3N \times 1$ column vectors; specifically, they are the $\{3(i-1) + N_{\psi}^{\mathsf{ph}}\}$-th columns of the $3N \times 3N$ matrices $\boldsymbol{R}_D$ and $\boldsymbol{X}_D$ defined as in (7b) and (7c).

Let the sequence of LMPs (cents/kWh) determined in RTM(OP) at the linkage bus b^* for the sub-periods $t \in \mathcal{K}$ comprising operating period OP be denoted by the $1 \times NK$ row vector $\mathbf{LMP}(\mathcal{K}) = [\mathrm{LMP}(b^*, 1), ..., \mathrm{LMP}(b^*, NK)]$. Also, let $\bar{P}$ (p.u.) denote an upper limit imposed on total demand during each sub-period $t \in \mathcal{K}$ for distribution network reliability, and let $\bar{\boldsymbol{P}}(\mathcal{K})$ denote the $NK \times 1$ column vector $[\bar{P}, \ldots, \bar{P}]^T$. In addition, let the $3N \times 1$ column vectors $\boldsymbol{v}_{\mathsf{min}}(t)$ and $\boldsymbol{v}_{\mathsf{max}}(t)$ denote lower and upper bounds (p.u.) imposed on the 3-phase squared voltage magnitudes during each sub-period $t \in \mathcal{K}$ for network reliability.

The *centralized DSO optimization problem* at the start of N(OP) is then expressed as follows:

$$\max_{\mathcal{P}(\mathcal{K}) \in \mathcal{X}(\mathcal{K})} \sum_{\psi \in \Psi} \Big[U_{\psi}(\mathcal{P}_{\psi}(\mathcal{K})) - \mu_{\psi} \mathbf{LMP}(\mathcal{K}) \mathcal{P}_{\psi}(\mathcal{K}) S_{\mathsf{base}} \Delta t\Big] \quad (21)$$

subject to the following demand and voltage network reliability constraints for each $t \in \mathcal{K}$:

$$\sum_{\psi \in \Psi} [p_{\psi}(t) + p_{\psi}^{\mathsf{non}}(t)] \le \bar{P} \quad (22a)$$

$$\boldsymbol{v}_{\mathsf{min}}(t) \le \boldsymbol{v}(t, \boldsymbol{p}_{\Psi}(t)) \le \boldsymbol{v}_{\mathsf{max}}(t) \quad (22b)$$

where

$$\mathcal{P}(\mathcal{K}) = \{\mathcal{P}_{\psi}(\mathcal{K}) \mid \psi \in \Psi\} = \{\boldsymbol{p}_{\Psi}(t)) \mid t \in \mathcal{K}\}$$
$$\mathcal{X}(\mathcal{K}) = \prod_{\psi \in \Psi} \mathcal{X}_{\psi}(\mathcal{K})$$

Finally, let the $(3N \cdot NK) \times 1$ column vectors $\boldsymbol{v}(\mathcal{P}(\mathcal{K}))$, $\boldsymbol{v}_{\mathsf{max}}(\mathcal{K})$, and $\boldsymbol{v}_{\mathsf{min}}(\mathcal{K})$ be defined as follows:

$$\boldsymbol{v}(\mathcal{P}(\mathcal{K})) = [\boldsymbol{v}(1, \boldsymbol{p}_{\Psi}(1))^T, \ldots, \boldsymbol{v}(NK, \boldsymbol{p}_{\Psi}(NK))^T]^T$$
$$\boldsymbol{v}_{\mathsf{max}}(\mathcal{K}) = [\boldsymbol{v}_{\mathsf{max}}(1)^T, \ldots, \boldsymbol{v}_{\mathsf{max}}(NK)^T]^T$$
$$\boldsymbol{v}_{\mathsf{min}}(\mathcal{K}) = [\boldsymbol{v}_{\mathsf{min}}(1)^T, \ldots, \boldsymbol{v}_{\mathsf{min}}(NK)^T]^T$$

Definition: Primal Problem. The centralized DSO optimization problem (21) can be expressed in a standard *nonlinear programming (NP)* form as follows:

$$\max_{\boldsymbol{x} \in \mathcal{X}} F(\boldsymbol{x}) \text{ subject to } \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{c} \quad (23)$$

where

$$\mathcal{X} = \mathcal{X}(\mathcal{K}) = \prod_{\psi \in \Psi} \mathcal{X}_{\psi}(\mathcal{K}) \subseteq \mathbb{R}^d$$
$$x_{\psi}(t) = p_{\psi}(t) \in \mathbb{R}$$
$$\boldsymbol{x}_{\psi} = \{x_{\psi}(t) \mid t \in \mathcal{K}\} = \mathcal{P}_{\psi}(\mathcal{K}) \in \mathbb{R}^{NK}$$
$$\boldsymbol{x} = \{\boldsymbol{x}_{\psi} \mid \psi \in \Psi\} = \mathcal{P}(\mathcal{K}) \in \mathbb{R}^d$$
$$F(\boldsymbol{x}) = \sum_{\psi \in \Psi} F_{\psi}(\boldsymbol{x}_{\psi})$$
$$F_{\psi}(\boldsymbol{x}_{\psi}) = \big[U_{\psi}(\boldsymbol{x}_{\psi}) - \mu_{\psi} \mathbf{LMP}(\mathcal{K}) \boldsymbol{x}_{\psi} \cdot S_{\mathsf{base}} \Delta t\big]$$
$$\boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} \sum_{\psi \in \Psi} [\boldsymbol{x}_{\psi} + \mathcal{P}_{\psi}^{\mathsf{non}}(\mathcal{K})] \\ \boldsymbol{v}(\boldsymbol{x}) \\ -\boldsymbol{v}(\boldsymbol{x}) \end{bmatrix}_{m \times 1}$$
$$\boldsymbol{c} = \begin{bmatrix} \bar{\boldsymbol{P}}(\mathcal{K}) \\ \boldsymbol{v}_{\mathsf{max}}(\mathcal{K}) \\ -\boldsymbol{v}_{\mathsf{min}}(\mathcal{K}) \end{bmatrix}_{m \times 1}$$

and: NH = number of households ψ; NK = number of sub-periods $t \in \mathcal{K}$; $d = NK \cdot NH$; N = number of non-head buses; and $m = ([1 + 6N] \cdot NK)$. Hereafter, problem (23) will be referred to as the *Primal Problem*. Depending on the context, a *solution* for this Primal Problem will variously be denoted by $\boldsymbol{x}^* = \{\boldsymbol{x}_{\psi}^* \mid \psi \in \Psi\} = \{\boldsymbol{\mathcal{P}}_{\psi}^*(\mathcal{K}) \mid \psi \in \Psi\} = \boldsymbol{\mathcal{P}}^*(\mathcal{K})$.

A critical point to note for later purposes is that the Primal Problem (23) does not directly depend on the retail price signals that the DSO communicates to the households under the

consensus-based TES design. What is sought below is a way to connect the optimal solution for the Primal Problem to the decentralized household optimal solutions determined as functions of the DSO's retail price signals.

C. DSO Optimization Problem in Distributed Control Form

The centralized DSO optimization problem (23) for operating period OP incorporates the local constraints $\mathcal{X}_\psi(\mathcal{K})$ for each household ψ as well as the network reliability constraints (22). Thus, to solve problem (21), the DSO would need a great deal of information about each household, a violation of household privacy.

Consequently, the DSO cannot directly solve the centralized optimization problem (21). Rather, as indicated in Fig.2, the DSO resorts to indirect control. Specifically, the DSO iteratively sets the price-to-go sequence $\boldsymbol{\pi}_\psi(\mathcal{K})$ for each household ψ's real TCL power usage to ensure that the resulting household real and reactive power usage levels for OP are consistent with the DSO's network reliability constraints for OP. As discussed in Section IV, each household ψ is required to continually submit its optimal real TCL power usage schedule $\mathcal{P}_\psi(\boldsymbol{\pi}_\psi(\mathcal{K}))$ for each possible price-to-go sequence $\boldsymbol{\pi}_\psi(\mathcal{K})$ sent by the DSO during period N(OP) until the negotiation process halts.[14]

The next section develops a specific analytical formulation for this distributed control method for the household TES design illustration.

VIII. TES Design Illustration: Solution Method

A. Solution Overview

This section replaces the centralized DSO optimization problem (21) with a specific formulation of the consensus-based TES design proposed in Section III. Thus, a centralized design requiring the DSO to exert direct control over household TCL power usage is replaced with a distributed design in which the DSO uses price signals to modify household TCL power usage decisions. A dual decomposition algorithm is developed to implement the negotiation process for this TES design, and the convergence properties of this design are established analytically.

B. TES Equilibrium

From the formulation (15) for each household ψ's optimization problem for an operating period OP, it is seen that ψ's choice of a TCL power-usage sequence for $\mathcal{K}$ depends on the price-to-go sequence $\boldsymbol{\pi}_\psi(\mathcal{K})$ for $\mathcal{K}$. The DSO can take advantage of this price dependence during N(OP) to ensure that household power usage does not violate any network reliability constraints.

Let $\boldsymbol{\pi}(\mathcal{K}) = \{\boldsymbol{\pi}_\psi(\mathcal{K}) \mid \psi \in \Psi\}$ denote a collection of price-to-go sequences communicated by the DSO to households during some iteration of the negotiation process N(OP). The real TCL power-usage sequences that households communicate back to the DSO in response to these communicated price-to-go sequences will then be denoted by $\boldsymbol{\mathcal{P}}(\boldsymbol{\pi}(\mathcal{K})) = \{\mathcal{P}_\psi(\boldsymbol{\pi}_\psi(\mathcal{K})) \mid \psi \in \Psi\}$.

Definition: TES Equilibrium. Suppose an optimal solution $\boldsymbol{x}^* = \boldsymbol{\mathcal{P}}^*(\mathcal{K})$ for the Primal Problem (23) equals $\boldsymbol{\mathcal{P}}(\boldsymbol{\pi}^*(\mathcal{K}))$ for some collection $\boldsymbol{\pi}^*(\mathcal{K})$ of retail household price-to-go sequences for an operationg period OP. Then the pairing $(\boldsymbol{\mathcal{P}}^*(\mathcal{K}), \boldsymbol{\pi}^*(\mathcal{K}))$ will be called a *TES equilibrium for OP*.

For each sub-period $t \in \mathcal{K}$, let $\lambda_{\bar{P}}(t)$ denote the non-negative dual variable (utils/p.u.) associated with the peak demand constraint (22a). Also, let the $1 \times 3N$ row vectors $\boldsymbol{\lambda}_{v_{\mathsf{max}}}(t)$ and $\boldsymbol{\lambda}_{v_{\mathsf{min}}}(t)$ denote the non-negative dual variables (utils/p.u.) associated with the upper and lower 3-phase voltage inequality constraints (22b). The $1 \times m$ row vector $\boldsymbol{\lambda}$ whose components consist of all of these non-negative dual variables is then denoted by

$$\boldsymbol{\lambda} = [\boldsymbol{\lambda}_{\bar{P}}(\boldsymbol{\mathcal{K}}), \boldsymbol{\lambda}_{v_{\mathsf{max}}}(\boldsymbol{\mathcal{K}}), \boldsymbol{\lambda}_{v_{\mathsf{max}}}(\boldsymbol{\mathcal{K}})] \tag{24}$$

where the component row vectors for $\boldsymbol{\lambda}$ are given by

$$\begin{aligned}
\boldsymbol{\lambda}_{\bar{P}}(\mathcal{K}) &= [\lambda_{\bar{P}}(1), \ldots, \lambda_{\bar{P}}(NK)]_{1\times NK} \\
\boldsymbol{\lambda}_{v_{\mathsf{max}}}(\mathcal{K}) &= [\boldsymbol{\lambda}_{v_{\mathsf{max}}}(1), \ldots, \boldsymbol{\lambda}_{v_{\mathsf{max}}}(NK)]_{1\times(3N\cdot NK)} \\
\boldsymbol{\lambda}_{v_{\mathsf{min}}}(\mathcal{K}) &= [\boldsymbol{\lambda}_{v_{\mathsf{min}}}(1), \ldots, \boldsymbol{\lambda}_{v_{\mathsf{min}}}(NK)]_{1\times(3N\cdot NK)}
\end{aligned}$$

Finally, for later purposes, the dual variables corresponding to the upper and lower 3-phase voltage inequality constraints (22b) are also expressed in the following matrix form:

$$\boldsymbol{\Lambda}_{v_{\mathsf{max}}}(\mathcal{K}) = \begin{bmatrix} \boldsymbol{\lambda}_{v_{\mathsf{max}}}(1) \\ \vdots \\ \boldsymbol{\lambda}_{v_{\mathsf{max}}}(NK) \end{bmatrix}_{NK\times 3N}$$

$$\boldsymbol{\Lambda}_{v_{\mathsf{min}}}(\mathcal{K}) = \begin{bmatrix} \boldsymbol{\lambda}_{v_{\mathsf{min}}}(1) \\ \vdots \\ \boldsymbol{\lambda}_{v_{\mathsf{min}}}(NK) \end{bmatrix}_{NK\times 3N}$$

The *Lagrangian Function* L: $\mathcal{X} \times \mathbb{R}_+^m \to \mathbb{R}$ for the centralized DSO optimization problem (21), equivalently represented in the Primal Problem form (23), is then given by

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = F(\boldsymbol{x}) + \boldsymbol{\lambda}[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x})] \tag{25}$$

where

$$\boldsymbol{x} = \{\boldsymbol{x}_\psi \mid \psi \in \Psi\} = \mathcal{P}(\mathcal{K}) \tag{26}$$

Finally, for each $t \in \mathcal{K}$, let $\boldsymbol{x}_\Psi(t) = \{x_\psi(t) \mid \psi \in \Psi\}$ where, as previously defined in Section VII-C,

$$\{x_\psi(t) \mid \psi \in \Psi\} = \{p_\psi(t) \mid \psi \in \Psi\} = \boldsymbol{p}_\Psi(t) \tag{27}$$

Then the Lagrangian Function (25) can equivalently be expressed in more specific terms as follows:

$$\begin{aligned}
L(\boldsymbol{x}, \boldsymbol{\lambda}) = F(\boldsymbol{x}) & \\
&+ \boldsymbol{\lambda}_{\bar{\boldsymbol{P}}}(\mathcal{K})\big[\bar{\boldsymbol{P}}(\mathcal{K}) - \sum_{\psi\in\Psi}[\boldsymbol{x}_\psi + \mathcal{P}_\psi^{\mathsf{non}}(\mathcal{K})]\big] \\
&+ \sum_{t\in\mathcal{K}}\big[\boldsymbol{\lambda}_{v_{\mathsf{max}}}(t)[\boldsymbol{v}_{\mathsf{max}}(t) - \boldsymbol{v}(t, \boldsymbol{x}_\Psi(t))]\big]
\end{aligned} \tag{28}$$

[14] Note that this negotiation process replaces the use of bid functions in bid-based TES designs; it provides an alternative way for households to communicate their price-sensitive power-usage preferences to the DSO.

$$+\sum_{t\in\mathcal{K}}\left[\boldsymbol{\lambda}_{v_{\mathsf{min}}}(t)[-\boldsymbol{v}_{\mathsf{min}}(t)+\boldsymbol{v}(t,\boldsymbol{x}_{\Psi}(t))]\right]$$

Definition: Saddle Point. A point $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ in $\mathcal{X}\times\mathbb{R}^m_+$ is said to be a *saddle point* for the Lagrangian Function $L(\boldsymbol{x},\boldsymbol{\lambda})$ defined in (25) if the following condition holds:

$$L(\boldsymbol{x},\boldsymbol{\lambda}^*)\le L(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)\le L(\boldsymbol{x}^*,\boldsymbol{\lambda}) \tag{29}$$

for all $\boldsymbol{\lambda}\in\mathbb{R}^m_+$ and $x\in\mathcal{X}$.

Definition: Dual Problem. Let the *dual function* D:$\mathbb{M}\rightarrow\mathbb{R}$ for the Primal Problem (23) be defined as follows:

$$D(\boldsymbol{\lambda})=\max_{\boldsymbol{x}\in\mathcal{X}}L(\boldsymbol{x},\boldsymbol{\lambda}) \tag{30}$$

where

$$\mathbb{M}=\{\boldsymbol{\lambda}\in\mathbb{R}^m_+\mid D(\boldsymbol{\lambda})\text{ is a well-defined finite value}\} \tag{31}$$

Then the *Dual Problem* associated with the Primal Problem (23) is defined to be

$$\min_{\boldsymbol{\lambda}\in\mathbb{M}}D(\boldsymbol{\lambda}) \tag{32}$$

Proposition 1 (Classical): *A point $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ in $\mathcal{X}\times\mathbb{R}^m_+$ is a saddle point (29) for the Lagrangian Function $L(\boldsymbol{x},\boldsymbol{\lambda})$ defined in (25) if and only if:*

- **[P1.A]** *$\boldsymbol{x}^*$ is a solution for the Primal Problem (23) ;*
- **[P1.B]** *$\boldsymbol{\lambda}^*$ is a solution for the Dual Problem (32) ;*
- **[P1.C]** *$D(\boldsymbol{\lambda}^*)$ = $F(\boldsymbol{x}^*)$ (strong duality).*

Proof of Proposition 1: See Appendix G.

Proposition 2: *Suppose $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ in $\mathcal{X}\times\mathbb{R}^m_+$ is a saddle point for the Lagrangian Function $L(\boldsymbol{x},\boldsymbol{\lambda})$ defined in (25), where $\boldsymbol{x}^*$ = $\boldsymbol{\mathcal{P}}^*(\mathcal{K})$. Suppose, also, that $\boldsymbol{x}^*$ uniquely maximizes $L(\boldsymbol{x},\boldsymbol{\lambda}^*)$ with respect to $\boldsymbol{x}\in\boldsymbol{\mathcal{X}}$. Define $\boldsymbol{\pi}^*(\mathcal{K})=\{\boldsymbol{\pi}^*_{\psi}(\mathcal{K})\mid\psi\in\Psi\}$, where the price-to-go sequence $\boldsymbol{\pi}^*_{\psi}(\mathcal{K})$ for each household $\psi\in\Psi$ takes the following form:*

$$\begin{aligned}\boldsymbol{\pi}^*_{\psi}(\mathcal{K}) \;=\; &\boldsymbol{LMP}(\mathcal{K})+\frac{1}{\mu_{\psi}S_{\mathsf{base}}\Delta t}\Big[\boldsymbol{\lambda}^*_{\bar{P}}(\mathcal{K})\\ &-2\cdot\boldsymbol{r}_D(i,N^{\mathsf{ph}}_{\psi})^T\left[\boldsymbol{\Lambda}^*_{v_{\mathsf{max}}}(\mathcal{K})-\boldsymbol{\Lambda}^*_{v_{\mathsf{min}}}(\mathcal{K})\right]^T\\ &-2\cdot\boldsymbol{x}_D(i,N^{\mathsf{ph}}_{\psi})^T\left[\boldsymbol{\Lambda}^*_{v_{\mathsf{max}}}(\mathcal{K})-\boldsymbol{\Lambda}^*_{v_{\mathsf{min}}}(\mathcal{K})\right]^T\boldsymbol{H}_{\psi}(\mathcal{K})\Big]\end{aligned} \tag{33}$$

The pairing $(\boldsymbol{\mathcal{P}}^*(\mathcal{K}),\boldsymbol{\pi}^*(\mathcal{K}))$ then constitutes a TES equilibrium for OP.

Proof of Proposition 2: See Appendix G.

For later purposes, note that the price-to-go sequence (33) depends on the attributes of household ψ = (u,ϕ,i). Specifically, the right-hand side of (33) depends on $\psi's$ preference and physical attributes u: namely, ψ's marginal utility of money μ_{ψ}; and ψ's TCL power-ratio function (16) as characterized by the $NK\times NK$ TCL power-ratio matrix $\boldsymbol{H}_{\psi}(\mathcal{K})$ defined in (18). In addition, the right-hand side of (33) depends on ψ's phase and bus location attributes ϕ and i through the $1\times 3N$ row vectors $\boldsymbol{r}_D(i,N^{\mathsf{ph}}_{\psi})^T$ and $\boldsymbol{x}_D(i,N^{\mathsf{ph}}_{\psi})^T$.

Note, also, that the price-to-go sequence (33) depends on the extent to which network reliability constraints would be violated by household power usage choices if retail prices for OP were simply set equal to the elements of the $1\times NK$ price vector $\mathbf{LMP}(\mathcal{K})$, i.e., to the LMP values determined for OP by RTM(OP). As seen in (33), the extent to which deviations from $\mathbf{LMP}(\mathcal{K})$ are needed to avoid network reliability constraint violations depends on the non-negative magnitudes of the dual variables (24) associated with the peak demand and voltage magnitude constraints for the Primal Problem (23).

C. TES Equilibrium Solution Strategy

A dual decomposition algorithm DDA is presented below for practical implementation of the negotiation process N(OP) between the DSO and the households. As discussed more carefully in Appendix F, dual decomposition is a classical decentralized method that alternates the updating of primal and dual variables until convergence to optimal primal and dual variable solutions takes place within a specified tolerance level.

A key issue is whether any limit point $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ resulting from the specific dual decomposition algorithm DDA is guaranteed to determine a TES equilibrium for OP. The following propositions 3-5 establish sufficient conditions for this to be the case.

Proposition 3: *Suppose the Primal Problem (23) and the dual decomposition algorithm DDA satisfy the following three conditions:*

- **[P3.A]** *$\mathcal{X}$ is compact, and the objective function $F(\boldsymbol{x})$ and constraint function $\boldsymbol{g}(\boldsymbol{x})$ are continuous over $\mathcal{X}$.*
- **[P3.B]** *For every $\boldsymbol{\lambda}\in\mathbb{R}^m_+$, the Lagrangian Function $L(\boldsymbol{x},\boldsymbol{\lambda})$ defined in (25) achieves a finite maximum at a unique point $\boldsymbol{x}(\boldsymbol{\lambda})\in\mathcal{X}$, implying the dual function domain in (31) satisfies $\mathbb{M}=\mathbb{R}^m_+$.*
- **[P3.C]** *The sequence $(\boldsymbol{x}^y,\boldsymbol{\lambda}^y)$ determined by the dual decomposition algorithm DDA converges to a limit point $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ as the iteration time y approaches $+\infty$.*

Then the limit point $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ is a saddle point (29) for the Lagrangian Function (25), and this saddle point determines a TES equilibrium for OP.

Proof of Proposition 3: See Appendix H.

Proposition 4: *Suppose the the Primal Problem (23) and the Dual Function (30) satisfy the following four conditions:*

- **[P4.A]** *Conditions [P3.A] and [P3.B] both hold;*
- **[P4.B]** *The Lagrangian Function (25) has a saddle point $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ in $\mathcal{X}\times\mathbb{R}^m_+$;*
- **[P4.C]** ***Extended Lipschitz Continuity Condition:*** *There exists a real symmetric positive-definite $m\times m$ matrix $\boldsymbol{J}$ such that, for all $\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\in\mathbb{R}^m_+$,*

$$\left\langle\nabla D_+(\boldsymbol{\lambda}_1)-\nabla D_+(\boldsymbol{\lambda}_2),\boldsymbol{\lambda}_1-\boldsymbol{\lambda}_2\right\rangle\;\le\;||\boldsymbol{\lambda}_1-\boldsymbol{\lambda}_2||^2_{\boldsymbol{J}}$$

  where: $\nabla D_+(\boldsymbol{\lambda})$ denotes the gradient of the dual function $D(\boldsymbol{\lambda})$ in (30) for $\boldsymbol{\lambda}\in\mathbb{R}^m_{++}$ and the right-hand gradient of $D(\boldsymbol{\lambda})$ at boundary points of $\mathbb{R}^m_+$; $\langle\ ,\ \rangle$ denotes vector inner product; and $||\cdot||^2_{\boldsymbol{J}}=(\cdot)\boldsymbol{J}(\cdot)^T$
- **[P4.D]** The matrix $[\boldsymbol{I}-\boldsymbol{JB}]$ *is positive semi-definite, where $\boldsymbol{I}$ denotes an $m\times m$ identity matrix, and where $\boldsymbol{B}$ is the $m\times m$ diagonal positive-definite matrix defined in step **S4** of the dual decomposition algorithm DDA.*

Algorithm DDA: Dual Decomposition Method for Approximate Determination of a TES Equilibrium

S1: Initialization. At the initial iteration time $y = 0$, the DSO specifies positive scalar step-sizes β_1, β_2, and β_3. In addition, the DSO sets initial dual variable values as follows: $\boldsymbol{\lambda}^y_{\bar{P}}(\mathcal{K}) = \mathbf{0}$, $\boldsymbol{\lambda}^y_{v_{\mathsf{max}}}(\mathcal{K}) = \mathbf{0}$, and $\boldsymbol{\lambda}^y_{v_{\mathsf{min}}}(\mathcal{K}) = \mathbf{0}$.

S2: Set price-to-go sequences. The DSO sets the price-to-go sequence $\boldsymbol{\pi}^y_\psi(\mathcal{K})$ for each household $\psi \in \Psi$ as follows:

$$\begin{aligned}\boldsymbol{\pi}^y_\psi(\mathcal{K}) \;&=\; \mathbf{LMP}(\mathcal{K}) + \frac{1}{\mu_\psi S_{\mathsf{base}} \Delta t}\Big[\boldsymbol{\lambda}^y_{\bar{P}}(\mathcal{K}) \\ &- 2\cdot \boldsymbol{r}_D(i, N^{\mathsf{ph}}_\psi)^T \big(\boldsymbol{\Lambda}^y_{v_{\mathsf{max}}}(\mathcal{K}) - \boldsymbol{\Lambda}^y_{v_{\mathsf{min}}}(\mathcal{K})\big)^T \\ &- 2\cdot \boldsymbol{x}_D(i, N^{\mathsf{ph}}_\psi)^T \big(\boldsymbol{\Lambda}^y_{v_{\mathsf{max}}}(\mathcal{K}) - \boldsymbol{\Lambda}^y_{v_{\mathsf{min}}}(\mathcal{K})\big)^T \boldsymbol{H}_\psi(\mathcal{K})\Big]\end{aligned}$$

Note that $\boldsymbol{\pi}^y_\psi(\mathcal{K})$ reduces to $\mathbf{LMP}(\mathcal{K})$ if $y = 0$.

S3: Update primal variables.

$$\boldsymbol{x}^y = \underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}}\, L(\boldsymbol{x}, \boldsymbol{\lambda}^y)$$

This updating of primal variable values is implemented as follows. The DSO communicates to each household $\psi \in \Psi$ the price-to-go sequence $\boldsymbol{\pi}^y_\psi(\mathcal{K})$. Each household $\psi \in \Psi$ then adjusts its TCL power usage schedule to

$$\boldsymbol{x}^y_\psi = \mathcal{P}_\psi(\boldsymbol{\pi}^y_\psi(\mathcal{K}))$$

and communicates $\boldsymbol{x}^y_\psi$ back to the DSO. If this primal variable updating step triggers the *Stopping Rule* outlined in Section IV, the negotiation process stops. Otherwise, the process proceeds to step **S4**.

S4: Update dual variables. The DSO determines updated dual variable values as follows: For each $t \in \mathcal{K}$,

$$\begin{aligned}\lambda^{y+1}_{\bar{P}}(t) &= \Big[\lambda^y_{\bar{P}}(t) + \beta_1\big[\sum_{\psi\in\Psi}[x^y_\psi(t) + p^{\mathsf{non}}_\psi(t)] - \bar{P}\big]\Big]^+ \\ \boldsymbol{\lambda}^{y+1}_{v_{\mathsf{max}}}(t) &= \Big[\boldsymbol{\lambda}^y_{v_{\mathsf{max}}}(t) + \beta_2\big[\boldsymbol{v}^{\mathsf{non}}(t) - 2\boldsymbol{s}(t, \boldsymbol{x}^y(t)) - \boldsymbol{v}_{\mathsf{max}}(t)\big]^T\Big]^+ \\ \boldsymbol{\lambda}^{y+1}_{v_{\mathsf{min}}}(t) &= \Big[\boldsymbol{\lambda}^y_{v_{\mathsf{min}}}(t) + \beta_3\big[-\boldsymbol{v}^{\mathsf{non}}(t) + 2\boldsymbol{s}(t, \boldsymbol{x}^y(t)) + \boldsymbol{v}_{\mathsf{min}}(t)\big]^T\Big]^+\end{aligned}$$

where $[\ \cdot\ ]_+$ denotes projection on R^k_+ for appropriate dimension k, and β_1, β_2, and β_3 are the positive scalar step-sizes specified by the DSO in step **S1**. Expressed in more compact form,

$$\boldsymbol{\lambda}^{y+1} = \big[\boldsymbol{\lambda}^y + [\boldsymbol{g}(\boldsymbol{x}^y) - \boldsymbol{c}]^T \boldsymbol{B}\big]^+$$

where $\boldsymbol{B}$ is an $m\times m$ diagonal positive-definite matrix constructed as follows: the diagonal entries of $\boldsymbol{B}$ associated with $\boldsymbol{\lambda}_{\bar{P}}(\mathcal{K})$, $\boldsymbol{\lambda}_{v_{\mathsf{max}}}(\mathcal{K})$, and $\boldsymbol{\lambda}_{v_{\mathsf{min}}}(\mathcal{K})$ are repeated entries of β_1, β_2, and β_3, respectively.

S5: Update iteration time. The iteration time y is assigned the updated value $y + 1$ and the process loops back to step **S2**

Then the primal-dual point $(\boldsymbol{x}^y, \boldsymbol{\lambda}^y)$ determined by the dual decomposition algorithm DDA at iteration time y converges to a saddle point as $y \to +\infty$.

Proof of Proposition 4: See Appendix I.

The Extended Lipschitz Continuity Condition [P4.C] in Prop. 4 is expressed in a relatively complicated form. The following proposition provides sufficient conditions for [P4.C] that are easier to understand.

Proposition 5: *Suppose the Primal Problem (23) satisfies condition [P3.A] in Prop. 3 plus the following three additional conditions:*

- **[P5.A]** *$\mathcal{X}$ is a non-empty compact convex subset of $\mathbb{R}^d$.*
- **[P5.B]** *The objective function F:$\mathbb{R}^d \to \mathbb{R}$ restricted to $\mathcal{X} \subseteq \mathbb{R}^d$ has the quadratic form*

$$F(\boldsymbol{x}) \;=\; \frac{1}{2}\,\boldsymbol{x}^T \boldsymbol{W} \boldsymbol{x} + \boldsymbol{\rho}^T \boldsymbol{x} + \sigma \qquad (34)$$

  where $\boldsymbol{W}$ is any real symmetric negative-definite $d \times d$ matrix, $\boldsymbol{\rho}$ is any real $d \times 1$ column vector, and σ is any real positive scalar.

- **[P5.C]** *The constraint function g:$\mathbb{R}^d \to \mathbb{R}^m$ restricted to $\mathcal{X} \subseteq \mathbb{R}^d$ has the linear affine form*

$$g(\boldsymbol{x}) \;=\; \boldsymbol{C}\boldsymbol{x} + \boldsymbol{b} \qquad (35)$$

  where $\boldsymbol{C}$ is any real $m \times d$ matrix, and $\boldsymbol{b}$ is any real $m \times 1$ column vector.

Then the Extended Lipschitz Continuity Condition [P4.C] in Prop. 4 holds for $\boldsymbol{J} = \boldsymbol{C}\boldsymbol{H}^{-1}\boldsymbol{C}^T$, where $\boldsymbol{H} = -\boldsymbol{W}$.

Proof of Proposition 5: See Appendix J.

An important aspect of the negotiation process N(OP) implemented by means of the dual decomposition algorithm DDA is that the DSO does not directly communicate iterated dual variable solutions to the households. Rather, the DSO communicates iterated retail price sequences to the households, requesting only that these households communicate back to the DSO what power amounts they would be willing to procure at these retail prices.

Consequently, N(OP) is based on an empirically meaningful exchange of information that households should find readily

understandable. Moreover, N(OP) does not require the DSO to solve an optimization problem in order to determine an approximate TES equilibrium for OP. This greatly reduces computational requirements for the DSO.

IX. Case Study

A. Overview

This section explores the practical effectiveness of our proposed consensus-based TES design by means of a case study.

As detailed below, the distribution network for this case study is an unbalanced 123-bus radial network populated by households. A DSO is tasked with managing the power usage requirements of these households. The DSO's centralized optimization problem is formulated as a concave programming problem with a strictly concave quadratic objective function $\boldsymbol{F}(\boldsymbol{x})$ and a linear affine constraint function $\boldsymbol{g}(\boldsymbol{x})$ defined over a non-empty compact convex subset $\mathcal{X} \subseteq \mathbb{R}^d$. The DSO decentralizes this optimization problem by implementing a consensus-based TES design.

The case study examines the performance of this consensus-based TES design for a single simulated 24-hour day D. The simulation is conducted using MATLAB R2019b, which integrates the YALMIP Toolbox [39] with the IBM ILOG CPLEX 12.9 solver [40]. Technical mathematical underpinnings for the case study are provided in Appendix K.

B. Unbalanced Radial Distribution Network

The standard IEEE 123-bus radial distribution network [41], shown in Fig.6, is modified for our case study in three ways. First, 345 households, each with non-TCL and TCL, are distributed across the 123 buses. Second, the distribution network is connected to a transmission network at its head bus 0. Third, power is supplied to the distribution network through this transmission-distribution interface. The base parameters for the distribution network are set as follows: $S_{\mathsf{base}} = 100$ (kVA); $\boldsymbol{v}_{\mathsf{min}}(t) = [0.95, 0.95, 0.95]^T$; $\boldsymbol{v}_{\mathsf{max}}(t) = [1.05, 1.05, 1.05]^T$; $\boldsymbol{v}_0(t) = [1.04, 1.04, 1.04]^T$; $V_{\mathsf{base}} = 4.16$ (kV); and $\bar{P} = 32$.

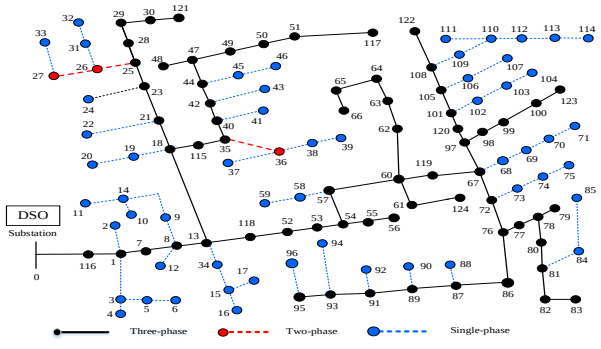

Fig. 6. IEEE 123-bus radial distribution network

C. Household Modeling

For simplicity, all households are assumed to have the same base parameters. The non-TCL profile of each household ψ during day D is shown in Fig. 7. The initial inside air temperature for each household ψ at the start of day D is $T_\psi(0) = 74$ (^oF). The ambient outside air temperature for each household ψ during day D is shown in Fig.8.

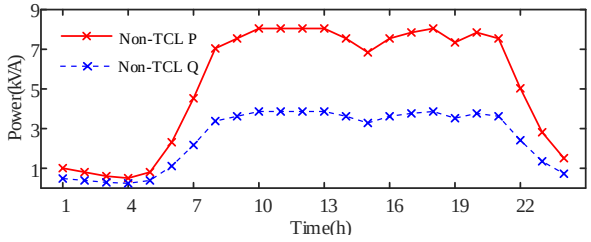

Fig. 7. Non-TCL real and reactive power profiles during day D

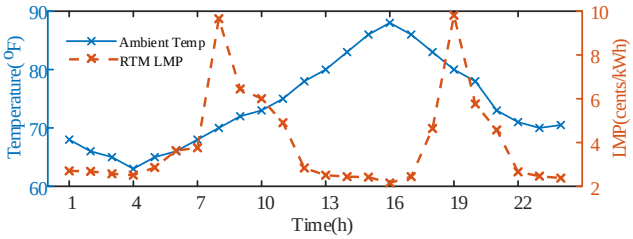

Fig. 8. Ambient outside air temperature and RTM LMPs during day D

The specific thermal dynamic parameter values set for each householdψ are: $\alpha_\psi^H = 0.96$ (unit-free); $\alpha_\psi^P = 0.7$ $(^oF/\text{kWh})$; $p_\psi^{\mathsf{max}} = 0.05$; and $\text{PF}_\psi(t) = 0.9$ for each sub-period t, which implies $\eta_\psi(t) = 0.48$ for each sub-period t.

As detailed in Appendix K, each household ψ has a strictly concave net-benefit objective function defined over a feasible choice set that is non-empty, compact, and convex. The specific preference parameter values set for each household ψ are: $c_\psi = 6.12$ $\left(\text{utils}/(^oF)^2\right)$; $u_\psi^{\mathsf{max}} = 1.20 \times 10^4$ (utils); $TB_\psi = 72$ (^oF); and $\mu_\psi = 1$ (utils/cent).

D. DSO, RTM, and N(OP) Modeling

A DSO operates at the radial network head bus 0 as a linkage entity that participates in both transmission and distribution system operations. An RTM operates over the transmission grid, and the DSO purchases power from this RTM in order to meet the power usage needs of distribution system households.

The simulated day D is partitioned into 24 operating hours OP. The duration of RTM(OP) and LAH(OP) for each operating hour OP are set to 1min and 59min; cf. Fig. 1. The number of sub-periods t partitioning each operating hour OP is set to $NK = 1$, and the length of this single sub-period t is set to Δt = 1h. The profile of RTM LMPs at the radial network head bus 0 determined in RTM(OP) for each operating hour OP of day D is depicted in Fig.8.

The objective of the DSO is to align local household goals and constraints with distribution network reliability constraints in a manner that respects household privacy. As detailed in Appendix K, in the absence of household privacy constraints

the DSO's centralized optimization problem for each operating period OP can be expressed as a concave programming problem with a strictly concave objective function and linear affine constraint function defined over a domain that is non-empty, compact, and convex. To avoid violation of household privacy, the DSO instead implements a consensus-based TES design that permits approximate implementation of the centralized optimal solution for each OP.

For each operating period OP during day D, the parameters of the dual decomposition algorithm DDA used to implement the negotiation process N(OP) are set as follows: β_1 = 15; β_2 = β_3 = 50,000; and $I_{\max}$ = 200.

E. Simulation Results

As detailed in Section VII-B, the DSO imposes two distinct types of network reliability constraints for each operating hour OP during day D: (i) an upper limit on the peak demand (kW) realized during OP; and (ii) lower and upper bounds imposed on the 3-phase squared voltage magnitudes (p.u.) realized during OP.

Consider, first, the case in which the DSO does not manage the power usage of its household customers. Rather, the DSO simply sets retail prices for all non-TCL and TCL household loads during day D equal to the RTM LMPs depicted in Fig. 8.

In this case, as shown in Fig.9, the peak demand is 2962kW. Consequently, as long as the upper limit on peak demand is set higher than this level, say at $\bar{P}$ = 3200kW, no peak demand limit violation occurs. On the other hand, lower and upper bounds on 3-phase squared voltage magnitudes (p.u.) are commonly set at 0.95 (p.u.) and 1.05 (p.u.). Given these bounds, it is seen in Fig. 9 that a voltage violation occurs at hour 17; specifically, the phase-a voltage magnitude drops to 0.9485 (p.u.).

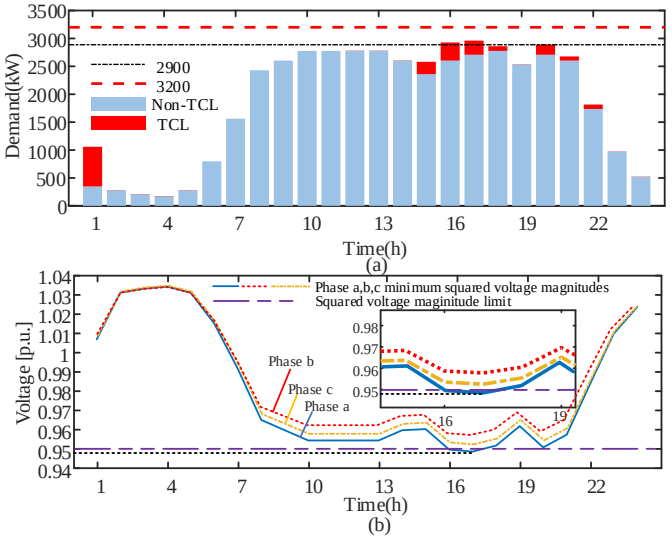

Fig. 9. **Unmanaged System Case (Peak Demand Limit 3200kW):** (a) Total household power demand (kW) and (b) 3-phase minimum squared voltage magnitudes (p.u.) across the distribution network during each hour of day D. The peak demand limit 3200kW is satisfied; but a phase-a violation of the lower voltage bound 0.95 (p.u.) occurs at hour 17.

Next, suppose one change is made to the Unmanaged System Case specifications: namely, the DSO now uses the consensus-based TES to manage household power usage. In particular, the DSO conducts a negotiation process N(OP) with households in advance of each operating hour OP during day D, implemented by means of the dual decomposition algorithm DDA. As explained in Section III-D, the negotiation process N(OP) continues until either there are no network reliability constraint violations or the number of negotiation rounds reaches the maximum permitted limit $I_{\max}$ = 200.

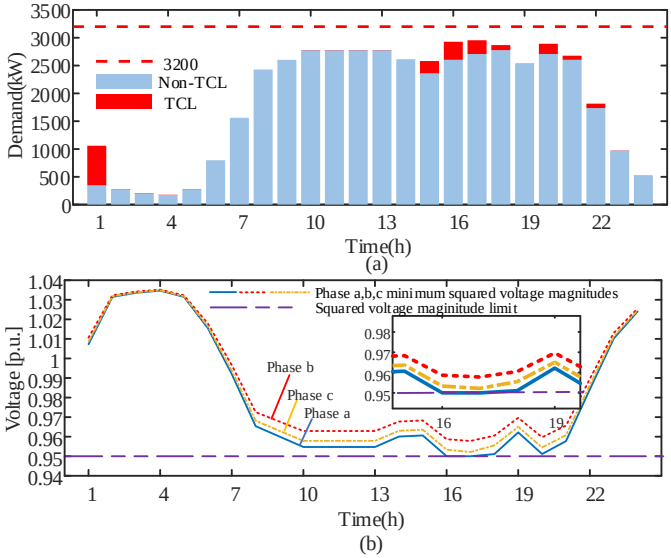

Fig. 10. **TES Management Case 1 (Peak Demand Limit 3200kW):** (a) Total household power demand (kW) and (b) 3-phase minimum squared voltage magnitudes (p.u.) across the distribution network during each hour of day D. The DSO-managed negotiation process ensures that no violations of the peak demand limit 3200kW or voltage magnitude bounds [0.95, 1.05] (p.u.) occur.

Fig.10 reports the total household power demand and 3-phase minimum squared voltage magnitudes that result for each hour for day D, given this change from no system management to TES management. As seen in Fig.10, all network reliability constraints are now satisfied. In particular, the switch to the use of the consensus-based TES design enables the DSO to eliminate the previously realized phase-a voltage constraint violation at hour 17 while still satisfying all other network reliability constraints.

Finally, suppose the peak demand limit is reduced from 3200kW to 2900kW, i.e., $\bar{P}$ = 2900kW. For the Unmanaged System Case, this change in peak demand limit has no effect on system operations. Consequently, as shown in Fig. 9, the peak demand 2962kW resulting for this case is now in violation of the reduced peak demand limit 2900kW; and the voltage magnitude violation for hour 17 continues to occur.

In contrast, given TES management, this reduction in peak demand limit from 3200kW to 2900kW results in a change in the DSO-conducted negotiation process with households. As seen in Fig. 10, the peak demand resulting for TES Management Case 1 (Peak Demand Limit 3200kW) does not satisfy the reduced peak demand limit 2900kW during some hours. Consequently, the DSO must now iteratively set retail prices for households in a different manner to ensure their power usage satisfies this reduced peak demand limit as well as the lower and upper voltage magnitude bounds.

The resulting demand outcomes for this TES Management Case 2 (Peak Demand 2900kW) are reported in Fig. 11. As seen, peak demand is maintained at or below the reduced peak demand limit 2900kW during all hours of day D. At the same time (not shown), the 3-phase minimum squared voltage magnitudes across the distribution network are maintained within their allowable limits [0.95, 1.05] (p.u.) during all hours of day D. For example, the smallest squared voltage magnitude across the distribution network during day D is 0.951 (p.u.).

As illustrated by these test cases, the core feature of the consensus-based TES design – namely, the DSO-managed negotiation process with distribution system customers – permits the DSO to protect against network reliability constraint violations, whatever form these constraints might take.

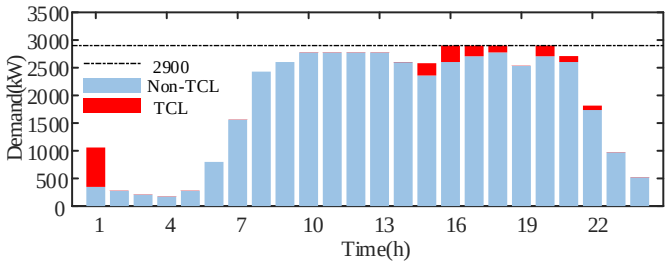

Fig. 11. **TES Management Case 2 (Peak Demand Limit 2900kW):** Total household power demand (kW) during each hour of day D. The DSO-managed negotiation process ensures that no peak demand limit violations occur.

F. Relationship Between Prices and Constraints

The retail price-to-go sequence for a household ψ in a TES equilibrium for an operating period OP, partitioned into subperiods $t \in \mathcal{K}$, is shown in Section VIII-B to take form $\boldsymbol{\pi}^*_\psi(\mathcal{K})$ in (33). This form is the summation of an initial price sequence, set by the DSO, that the DSO then modifies as necessary during the negotiation process N(OP) to ensure all network reliability constraints for OP are met.

As noted in Section VIII-B, the price-to-go sequence (33) for household ψ depends on $\psi's$ preference and physical attributes as well as ψ's network location. Specifically, the right-hand side of (33) depends on: (i) ψ's marginal utility of money μ_ψ; (ii) ψ's TCL power-ratio function (16) as characterized by the TCL power-ratio matrix $\boldsymbol{H}_\psi(\mathcal{K})$ defined in (18); and (iii) ψ's phase and bus location attributes ϕ and i through the terms $\boldsymbol{r}_D(i, N^{\mathsf{ph}}_\psi)^T$ and $\boldsymbol{x}_D(i, N^{\mathsf{ph}}_\psi)^T$.

For simplicity, this study assumes that the initial price-to-go sequence set by the DSO at the beginning of the negotiation process N(OP) is the sequence $\mathbf{LMP}(\mathcal{K})$ of LMPs determined in the real-time market RTM(OP) at the linkage bus b^*. This linkage bus, which connects the distribution network to a relatively large transmission network, is also the head bus 0 for the radial distribution network.

Any subsequent deviations from this initial price-to-go sequence that result from the negotiation process N(OP) are expressed in terms of dual variables for the network reliability constraints for OP. Specifically, these deviations are functions of the non-negative dual variables $\boldsymbol{\lambda}^*_{\bar{P}}(\mathcal{K})$, $\boldsymbol{\Lambda}^*_{v_{\min}}(\mathcal{K})$, and $\boldsymbol{\Lambda}^*_{v_{\max}}(\mathcal{K})$ corresponding to the peak demand constraint and the lower and upper voltage magnitude constraints for each $t \in \mathcal{K}$, where each of these constraints is expressed as an inequality constraint.

In a TES equilibrium for OP, the values of these dual variables must coincide, by definition, with the dual variable solutions for the Primal Problem (23). If strict inequality holds for a network reliability constraint in this Primal Problem solution, i.e., the constraint is *inactive*, then the corresponding dual variable solution must be zero. Thus, if strict inequality holds for all network reliability constraints in this Primal Problem solution, the retail price-to-go sequences communicated to households will simply coincide with the DSO's initially set RTM LMPs. The remainder of this section analyzes how TES equilibrium retail price outcomes deviate from RTM LMPs during the simulated day D for cases in which at least one network reliability constraint is active in the Primal Problem solution.

Consider, first, the TES equilibrium retail price outcomes for hour 17 that are reported in Fig. 12 for TES Management Case 1 with peak demand limit $\bar{P}$ = 3200kW. These retail price outcomes are seen to vary with respect to both bus location and phase. What explains this variation?

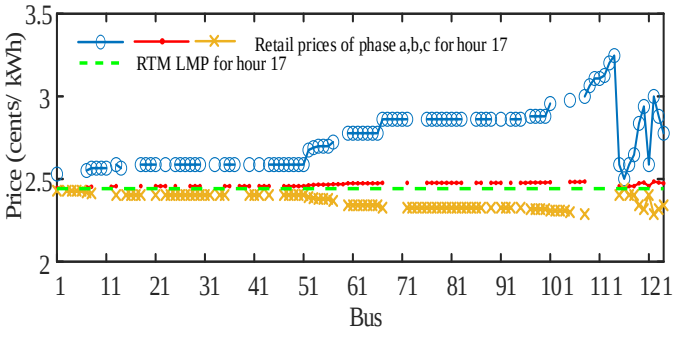

Fig. 12. TES Management Case 1 (Peak Demand Limit 3200kW): TES equilibrium retail prices for hour 17 of day D across the distribution network (buses 1-123), compared with the RTM LMP at bus 0 for hour 17 of day D.

As seen in Fig.10, during hour 17 the peak demand remains strictly below the peak demand limit 3200kW. Thus, the peak demand constraint is inactive, implying that the dual variable solution associated with this inactive peak demand constraint must be zero, i.e., $\boldsymbol{\lambda}^*_{\bar{P}}(\mathcal{K}) = \mathbf{0}$.

On the other hand, during hour 17 the minimum squared voltage magnitude across phases and buses reaches the lower bound 0.95 (p.u.), i.e., the lower-bound voltage constraint is active. Typically,[15] the dual variable solution $\boldsymbol{\Lambda}^*_{v_{\min}}(\mathcal{K})$ associated with this active voltage constraint will then be strictly positive. In this case, the TES equilibrium price-to-go sequence (33) determined for each household ψ for hour 17 will deviate from the RTM LMP at bus 0 for hour 17, i.e., the initial retail price commonly set by the DSO for each household ψ in the negotiation process for hour 17.

[15]By Lemma 1 in Appendix G, a non-negative dual variable solution for an inactive constraint must be 0, but the converse does not necessarily hold. That is, the dual variable solution corresponding to an active constraint is not necessarily strictly positive.

These findings have the following important implication. Even if all households populating a distribution network have identical benefit (comfort) functions and identical structural house attributes, this does *not* imply they should be charged the same retail power price. Rather, in the presence of active voltage reliability constraints, optimal pricing will typically require households associated with different marginal utility of money parameters, different power factors, different phases, and/or different bus locations to be charged different retail prices. This differential retail pricing reflects the roles played by *household preference attributes, power factors and network locations* in ensuring the satisfaction of these voltage reliability constraints.

Consider, next, the relationship between TES equilibrium retail price outcomes and network reliability constraints for the TES Management Case 2 with peak demand limit $\bar{P}$ = 2900kW. As discussed in Section IX-E, for this case the voltage reliability constraints are inactive; hence, all dual variable solutions associated with these voltage constraints are zero. It follows that the TES equilibrium price-to-go sequence (33) for each household ψ has the following reduced form:

$$\boldsymbol{\pi}^*_{\psi}(\mathcal{K}) = \mathbf{LMP}(\mathcal{K}) + \frac{1}{\mu_{\psi} S_{\text{base}} \Delta t} \boldsymbol{\lambda}^*_{\bar{P}}(\mathcal{K}) \tag{36}$$

On the other hand, as shown in Fig.11, the peak demand constraint is active for hours 16-18 and 20 during day D. The TES equilibrium retail price outcomes for these hours are reported in Fig.13. The retail price is strictly higher than the RTM LMP for each of these hours, indicating that the dual variable solution vector $\boldsymbol{\lambda}^*_{\bar{P}}(\mathcal{K})$ for the peak demand constraints during day D, appearing in (36) includes strictly positive values for these four hours.

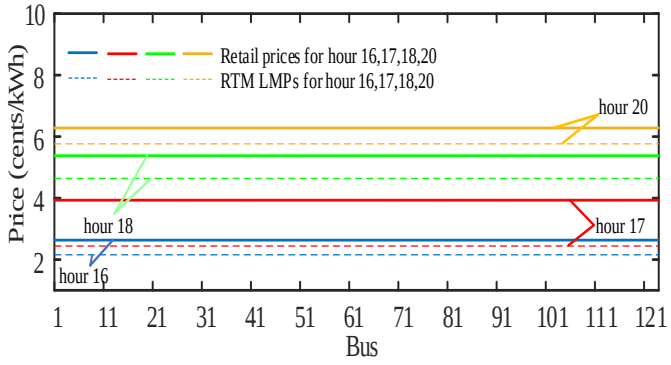

Fig. 13. TES Management Case 2 (Peak Demand Limit 2900kW): TES equilibrium retail prices for hours 16-18 and 20 of day D across the distribution network (buses 1-123), compared with the RTM LMP outcomes at bus 0 for these same hours.

The form of household ψ's TES equilibrium price-to-go sequence $\boldsymbol{\pi}^*_{\psi}(\mathcal{K})$ in (36) has the following important implication. Note that household ψ's marginal utility of money parameter μ_{ψ} appears in the denominator of the far-right term in (36). Consequently, even if voltage network constraints are inactive during day D, households with different marginal utility of money assessments will typically be charged different prices during each hour of day D for which the peak demand limit constraint is active.

As noted in IX-C, for the case study at hand the marginal utility of money parameter μ_{ψ} is commonly set to μ_{ψ} = 1 (utils/cent) for each household $\psi \in \Psi$. Consequently, as seen in Fig. 13, for this special case the TES equilibrium retail prices for TES Management Case 2 are the same for each household ψ during the operating hours 16-18 and 20 even though the peak demand limit is active for these hours.

G. Optimality Verification and Comparison

This subsection explores the following important question: Does the TES equilibrium determined by the consensus-based TES design closely approximate the optimal solution for the centralized DSO optimization problem (21)? An affirmative answer is provided for the case study developed in previous subsections. For concreteness, results are presented below for TES Management Case 1 (Peak Demand Limit 3200kW).

Fig. 14 compares the solutions obtained for total household TCL during each hour of day D using these two different methods. Fig. 15 provides a finer-grained comparison for total phase-a household TCL during hour 17 across the 123 buses comprising the entire distribution network. In each case, the resulting solutions are seen to be virtually indistinguishable.

This finding has two important implications. First, the consensus-based TES design achieves optimality while protecting the privacy of participating customers. In contrast, the centralized DSO optimization method requires extensive knowledge of customer attributes, including benefit (utility) functions and local feasibility constraints. Second, the consensus-based TES design is a decentralized solution method, which results in reduced computational requirements and improved scalability properties.[16] In contrast, the centralized DSO optimization method requires the DSO to solve a multiperiod optimization problem whose computational requirements dramatically increase with the number of participating customers.

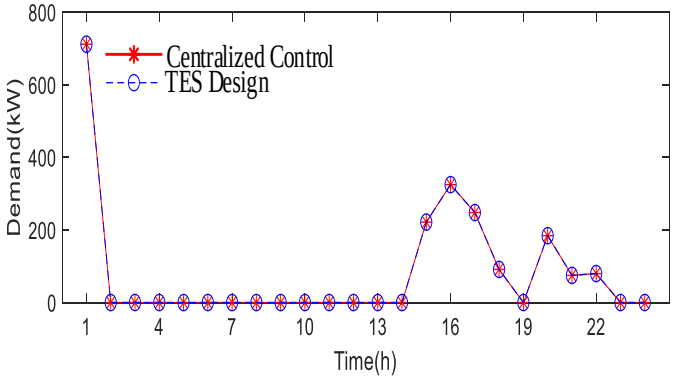

Fig. 14. Comparison of TES equilibrium and centralized DSO optimal solutions for total household TCL during day D.

[16]As will be discussed in Section X, the consensus-based TES design can be extended to incorporate aggregators as intermediaries between the DSO and various subsets of customers, thus further enhancing its scalability.

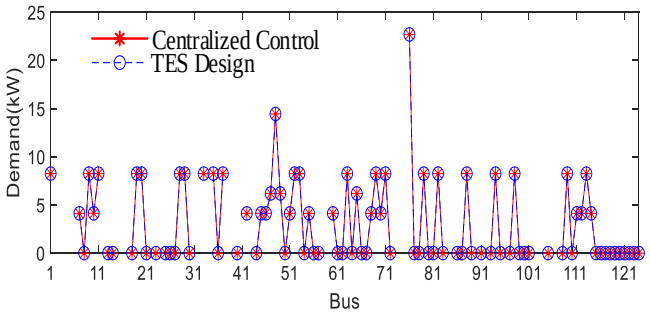

Fig. 15. Comparison of TES equilibrium and centralized DSO optimal solutions for total phase-a household TCL across the distribution network (123 buses) during operating hour 17.

X. Conclusion

This study develops a new consensus-based TES design for unbalanced distribution networks populated by customers with both fixed and price-sensitive power usage demands. The design is managed by a DSO. However, it is implemented as a distributed optimization problem, thus permitting alignment of system goals and network reliability constraints with local customer goals and constraints in a manner that respects customer privacy.

The core feature of this consensus-based TES design is a multi-round negotiation process N(OP) between the DSO and participant customers, to be held in advance of each operating period OP. At the start of N(OP), the DSO sets initial prices based on RTM LMPs. During each successive negotiation round, the DSO communicates updated price-to-go sequences to customers for OP; and the customers respond by communicating back to the DSO their optimal price-sensitive power usage levels for OP conditional on these prices and on private local constraints. The negotiation process terminates either when all network reliability constraints are satisfied by these customer power usage responses or when a stopping rule is activated.

A complete analytical formulation of the consensus-based TES design is developed for an unbalanced radial distribution network populated by welfare-maximizing households. Each operating day D is partitioned into operating periods OP of arbitrary duration, with look-ahead horizons LAH(OP). The negotiation process N(OP) for each operating period OP is then implemented during LAH(OP) by means of a newly developed dual decomposition algorithm DDA.

Making use of both classical and newly established results, sufficient conditions are established for the DDA to converge to a TES equilibrium whose power usage levels coincide with the optimal power usage solutions for a centralized full-information optimization problem that incorporates all network reliability constraints. Moreover, the TES equilibrium price-to-go sequences determined by the DDA are shown to have an informative additive structure that expresses deviations from initial prices in terms of the dual variable solutions associated with network reliability constraints.

A case study for an unbalanced 123-bus radial distribution network is presented to illustrate the capabilities of the consensus-based TES design and its DDA implementation. Numerical results are presented that demonstrate the convergence of the DDA to a TES equilibrium that closely approximates a centralized full-information optimal solution.

Future studies will seek to extend the capabilities of the consensus-based TES design in three main directions. First, the TES design will be generalized to permit consideration of customer-owned distributed generation as well as customer power usage levels. Particular attention will be focused on the inclusion of inverter-based distributed generation such as wind and solar power facilities. This extension will permit a more careful consideration of reactive power as an ancillary service product, supplied in return for appropriate compensation.

Second, the consensus-based TES design will be extended to permit the inclusion of aggregators operating as intermediaries between the DSO and its managed customers. The communication network would still take a radial form; however, the ability of aggregators to perform intermediate aggregation of customer responses could reduce communication times. For example, it could permit an efficient bundling of customers into distinct aggregator-managed subsets on the basis of their observable attributes or their historically observed behaviors. This bundling could enhance the ability of the DSO to ensure all network reliability constraints are met through the negotiation process in a practically reasonable amount of time.

Third, the negotiation process for the consensus-based TES design will be modified to permit more sophisticated specifications for the initial price-to-go sequences set by the DSO for customer price-sensitive demands. In the current study, these initial prices are simply set equal to RTM LMPs. However, as noted in Section III-D, this specification could expose customers to undesirable price volatility. Hence, a better alternative might be to set these initial prices equal to time and/or spatially averaged RTM LMPs.

However, an additional issue must also be considered. In order for the DSO to ensure its independent status, any net revenues or net costs that the DSO incurs through its operations must be allocated back to its managed customers. Consequently, the DSO's initial price-to-go sequences for customer price-sensitive demands should be set to ensure the DSO breaks even on average over time; that is, its operational revenues should match its operational costs on average over time. This break-even requirement could force a DSO to set initial prices at levels that deviate from RTM LMPs, even if setting initial prices equal to RTM LMPs would not result in any network reliability constraint violations.

Appendix A

Nomenclature

A. Acronyms, Parameters, and Other Exogenous Terms

$\bar{\boldsymbol{A}}$	Graph-based incidence matrix (p.u.) for the unbalanced radial network;
b^*	Linkage bus, i.e., the transmission grid bus at which the radial network links to the transmission network;

$\boldsymbol{B}$ Diagonal matrix with DDA step-sizes along its diagonal;

$BP(j)$ Bus immediately preceding bus j along radial network, for all $j \in \mathcal{N}$;

Bus 0 Radial network head bus, which is also the linkage bus b^*;

c_ψ Conversion factor (utils/$({}^oF)^2$) between comfort and inside air temperature for customer ψ;

d Number NK of sub-periods times number NH of households;

$\boldsymbol{D}_r$ Block diagonal matrix (p.u.) of line segment resistances;

$\boldsymbol{D}_x$ Block diagonal matrix (p.u.) of line segment reactances;

DDA Dual decomposition algorithm for implementation of N(OP);

DER Distributed energy resource;

DSO Distribution system operator;

$\boldsymbol{H}_\psi(\mathcal{K})$ TCL power-ratio matrix for customer ψ during $\mathcal{K}$;

I_{max} Maximum permitted number of negotiation process iterations;

ISO Independent system operator;

$\ell_j = (i,j)$ Line segment connecting buses i and j with i = BP(j) and $j \in \mathcal{N}$;

LAH(OP) Look-ahead horizon for RTM(OP);

LMP Locational marginal price;

LMP(b^*,t) RTM LMP (cents/kWh) at linkage bus b^* for sub-period t;

$\mathbf{LMP}(\mathcal{K})$ Vector of LMPs (cents/kWh) determined in RTM(OP) for $\mathcal{K}$;

m Number $([1+6N]\cdot NK)$ of explicit Primal Problem constraints;

$\bar{\boldsymbol{M}}$ Graph-based incidence matrix (p.u.) for the balanced radial network;

N Number of non-head buses for the radial network;

NH Number of households $\psi \in \Psi$;

NK Number of sub-periods t forming a partition of operating period OP;

N(OP) Negotiation process for OP;

N_ψ^{ph} Flag for phase $\phi \in \{a,b,c\}$ of the 1-phase line connecting customer ψ to a distribution network bus;

OP Operating period;

$\bar{P}$ Upper limit for peak demand (p.u.)

$\mathrm{PF}_\psi(t)$ Power factor (unit free) for the TCL device of customer ψ during sub-period t;

p_ψ^{max} Maximum real power level (p.u.) for customer ψ's TCL devices;

$p_\psi^{\mathsf{non}}(t)$ Non-TCL real power usage (p.u.) of customer ψ at start of sub-period t;

$\mathcal{P}_\psi^{\mathsf{non}}(\mathcal{K})$ Non-TCL real power sequence (p.u.) of customer ψ for $\mathcal{K}$;

$q_\psi^{\mathsf{non}}(t)$ Non-TCL reactive power usage (p.u.) of customer ψ at the start of sub-period t;

$\mathcal{Q}_\psi^{\mathsf{non}}(\mathcal{K})$ Non-TCL reactive power sequence (p.u.) of customer ψ for $\mathcal{K}$;

$\boldsymbol{R}_{ij}, \boldsymbol{X}_{ij}$ 3-phase resistance & reactance matrices (p.u.) for line segment (i,j);

$\bar{\boldsymbol{R}}_{ij}, \bar{\boldsymbol{X}}_{ij}$ 3-phase resistance & reactance matrices (p.u.) for line segment (i,j) after transformation;

RTM(OP) Real-time market for OP;

RTO Regional transmission operator;

S_{base} Base apparent power (kVA);

SCED(OP) Security-constrained economic dispatch for OP;

TB_ψ Inside air temperature $({}^oF)$ at which customer ψ achieves maximum comfort (bliss);

TES Transactive energy system;

TCL Thermostatically-controlled load;

$T_\psi(0)$ Inside air temperature $({}^oF)$ for customer ψ at the start of OP;

$T_o(t)$ Ambient outside air temperature $({}^oF)$ at the start of sub-period t;

u_ψ^{max} Customer ψ's maximum attainable thermal comfort (utils);

V_{base} Base voltage (kV);

$\boldsymbol{v_0}(t)$ 3-phase squared voltage magnitudes (p.u.) at the head bus 0 during sub-period t;

$\boldsymbol{v}^{\mathsf{non}}(t)$ 3-phase squared voltage magnitudes (p.u.) at all non-head buses for sub-period t, assuming no TCL;

$\boldsymbol{v}_{\mathsf{min}}(t), \boldsymbol{v}_{\mathsf{max}}(t)$ Lower and upper bounds (p.u.) for 3-phase squared voltage magnitudes during sub-period t;

α_ψ^H System inertia temperature parameter (unit-free) for customer ψ;

α_ψ^P Temperature parameter $({}^oF/\mathrm{kWh})$ for customer ψ;

$\beta_1, \beta_2, \beta_3$ Step sizes (unit-free) for the dual decomposition algorithm DDA;

$\triangle t$ Length (h) of each sub-period t;

$\eta_\psi(t)$ Ratio (unit free) of TCL reactive power to TCL real power for customer ψ during sub-period t;

γ_ψ Benefit/cost slider-knob control setting (unit free) communicated by customer ψ to the DSO for OP;

μ_ψ Customer ψ's marginal utility of money (utils/cent) for OP;

ϕ Circuit phase of a line segment ℓ_j, or of a 1-phase line connecting a household to a bus;

$\psi = (u, \phi, i)$ Designator for a customer with structural and preference attributes u located on a phase-ϕ line connected to bus i.

B. Sets and Sequences

$\mathcal{K} = (1, \ldots, NK)$	Sequence of sub-periods t that partition operating period OP;
$\mathcal{L}$	Set of all N distinct line segments (i, j) connecting adjacent buses i and j in $\{0\} \bigcup \mathcal{N}$;
$\mathcal{N} = \{1, \ldots, N\}$	Index set for all non-head buses of the radial network;
$\mathcal{N}_j$	Index set for all buses located strictly after bus j along the radial network, $0 \leq j < N$;
$\boldsymbol{\mathcal{P}}(\mathcal{K})$	Set of customer TCL power usage sequences during $\mathcal{K}$;
$\boldsymbol{\mathcal{P}}(\boldsymbol{\pi}(\mathcal{K}))$	Set of customer TCL power usage sequences during $\mathcal{K}$, given $\boldsymbol{\pi}(\mathcal{K})$;
$\mathcal{U}_{i,\phi}$	Set of attributes u such that (u, ϕ, i) denotes a customer $\psi \in \Psi$;
$\mathcal{X}_\psi(\mathcal{K})$	Set of customer ψ constraints for $\mathcal{K}$;
$\Phi = \{a, b, c\}$	Set of line phases ϕ;
$\boldsymbol{\pi}(\mathcal{K})$	Set of customer price-to-go sequences for $\mathcal{K}$;
Ψ	Set of all customers ψ.

C. Functions, & Variables

f_ψ	Power-factor function for price-sensitive demands communicated by customer ψ to the DSO for OP;
$L(\boldsymbol{x}, \boldsymbol{\lambda})$	Lagrangian Function for the DSO's centralized optimization problem;
NetBen_ψ	Customer ψ's net benefit function (utils) for OP;
$\boldsymbol{P}_{ij}(t), \boldsymbol{Q}_{ij}(t)$	3-phase real & reactive power flows (p.u.) over line segment (i, j) during sub-period t;
$\boldsymbol{P}(t), \boldsymbol{Q}(t)$	3-phase real & reactive power flows (p.u.) over all line segments during sub-period t;
$\boldsymbol{p}_i(t), \boldsymbol{q}_i(t)$	3-phase real & reactive power (p.u.) at bus i during sub-period t;
$\boldsymbol{p}(t), \boldsymbol{q}(t)$	3-phase real & reactive power (p.u.) at all non-head buses during sub-period t;
$p_\psi(t), q_\psi(t)$	TCL real & reactive power usage (p.u.) of customer ψ during sub-period t;
$\mathcal{P}_\psi(\mathcal{K})$	TCL real power usage sequence (p.u.) of customer ψ for $\mathcal{K}$;
$\mathcal{Q}_\psi(\mathcal{K})$	TCL reactive power usage sequence (p.u.) of customer ψ for $\mathcal{K}$;
$T_\psi(p_\psi(t), t)$	Inside air temperature (^oF) of customer ψ's home at the end of sub-period t, given $p_\psi(t)$;
$u_\psi(p_\psi(t), t)$	Comfort (utils) attained by customer ψ during sub-period t, given $p_\psi(t)$;
$U_\psi(\mathcal{P}_\psi(\mathcal{K}))$	Total comfort (utils) attained by customer ψ during OP, given $\mathcal{P}_\psi(\mathcal{K})$;
$\boldsymbol{v}(t, \boldsymbol{p}_\Psi(t))$	3-phase squared voltage magnitudes (p.u.) at all non-head buses for sub-period t;
$\boldsymbol{v}_i(t, \boldsymbol{p}_\Psi(t))$	3-phase squared voltage magnitudes (p.u.) at bus i for sub-period t;
$\boldsymbol{\lambda}$	Dual variables (utils/p.u.) for all network reliability constraints for $\mathcal{K}$;
$\lambda_{\bar{P}}(t)$	Dual variable (utils/p.u.) associated with demand limit for sub-period t;
$\boldsymbol{\lambda}_{\bar{P}}(\mathcal{K})$	Dual variables (utils/p.u.) associated with demand limits for $\mathcal{K}$;
$\boldsymbol{\lambda}_{v_{\max}}(t)$	Dual variables (utils/p.u.) associated with upper voltage limits for sub-period t;
$\boldsymbol{\Lambda}_{v_{\max}}(\mathcal{K})$	Dual variable matrix associated with upper voltage limits for $\mathcal{K}$;
$\boldsymbol{\lambda}_{v_{\min}}(t)$	Dual variables (utils/p.u.) associated with lower voltage limits for sub-period t;
$\boldsymbol{\Lambda}_{v_{\min}}(\mathcal{K})$	Dual variable matrix associated with lower voltage limits for $\mathcal{K}$;
$\pi_\psi(t)$	Retail price (cents/kWh) of customer ψ's price-sensitive demand for sub-period t;
$\boldsymbol{\pi}_\psi(\mathcal{K})$	Price-to-go sequence (cents/kWh) for customer ψ during $\mathcal{K}$.

APPENDIX B

Introduction of Virtual Circuits and Phases to Facilitate the Representation of Unbalanced Radial Distribution Networks

The number of phases associated with multi-phase line segments for unbalanced radial distribution networks can differ widely from one network to the next. This profusion of forms hinders the representation of these networks.

To resolve this problem, for multi-phase line segments with fewer than three phases we introduce "virtual" circuits with "virtual" phases whose self-impedance, mutual impedance, and bus loads are set to zero. By appropriate introduction of these virtual circuits and phases, each line segment in the original unbalanced radial distribution network can be represented as a 3-phase line segment. Since the power flow and voltage drop for any virtual circuits are 0, power flow solutions for the original network are not affected by these virtual extensions.

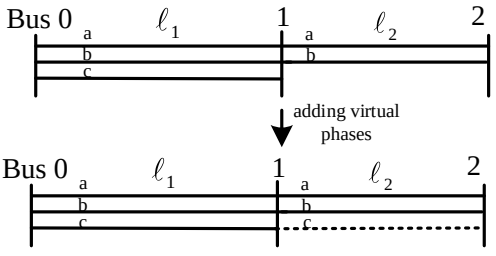

Fig. 16. Illustrative virtual extension of an unbalanced distribution network

A simple example will next be given to illustrate this virtual extension method for unbalanced radial distribution networks. Consider the simple network with three buses and two line

segments that is depicted in Fig.16. The line segment $\ell_2 = (1,2)$ connecting bus 1 to bus 2 has only two circuits, with phases a and b. A virtual circuit with phase c is added to line segment ℓ_2, and the self-impedance, mutual impedance, and bus load associated with this virtual circuit are each set to 0, as follows:

$$p_2^c(t) = q_2^c(t) = 0 \tag{37a}$$
$$z_{12}^{cc} = z_{12}^{ac} = z_{12}^{ca} = z_{12}^{bc} = z_{12}^{cb} = 0 \tag{37b}$$

The impedances of the line segments $\ell_1 = (0,1)$ and $\ell_2 = (1,2)$ are then represented in 3-phase form as follows:

$$\boldsymbol{Z}_{01} = \boldsymbol{R}_{01} + j\boldsymbol{X}_{01} = \begin{bmatrix} z_{01}^{aa} & z_{01}^{ab} & z_{01}^{ac} \\ z_{01}^{ba} & z_{01}^{bb} & z_{01}^{bc} \\ z_{01}^{ca} & z_{01}^{cb} & z_{01}^{cc} \end{bmatrix} \tag{38a}$$

$$\boldsymbol{Z}_{12} = \boldsymbol{R}_{12} + j\boldsymbol{X}_{12} = \begin{bmatrix} z_{12}^{aa} & z_{12}^{ab} & 0 \\ z_{12}^{ba} & z_{12}^{bb} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{38b}$$

From (2a)-(2c), we can write the power flow for the resulting unbalanced 3-phase radial distribution network as follows:

$$\boldsymbol{P}_{01}(t) = \boldsymbol{P}_{12}(t) + \boldsymbol{p}_1(t) \tag{39a}$$
$$\boldsymbol{Q}_{01}(t) = \boldsymbol{Q}_{12}(t) + \boldsymbol{q}_1(t) \tag{39b}$$
$$\boldsymbol{P}_{12}(t) = \boldsymbol{p}_2(t) \tag{39c}$$
$$\boldsymbol{Q}_{12}(t) = \boldsymbol{q}_2(t) \tag{39d}$$
$$\boldsymbol{v}_0(t) = \boldsymbol{v}_1(t) + 2(\bar{\boldsymbol{R}}_{01}\boldsymbol{P}_{01}(t) + \bar{\boldsymbol{X}}_{01}\boldsymbol{Q}_{01}(t)) \tag{39e}$$
$$\boldsymbol{v}_1(t) = \boldsymbol{v}_2(t) + 2(\bar{\boldsymbol{R}}_{12}\boldsymbol{P}_{12}(t) + \bar{\boldsymbol{X}}_{12}\boldsymbol{Q}_{12}(t)) \tag{39f}$$

where

$$\begin{aligned}\bar{\boldsymbol{R}}_{01} &= Re(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{R}_{01} + Im(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{X}_{01} \\ &= \begin{bmatrix} \bar{R}_{01}^{aa} & \bar{R}_{01}^{ab} & \bar{R}_{01}^{ac} \\ \bar{R}_{01}^{ba} & \bar{R}_{01}^{bb} & \bar{R}_{01}^{bc} \\ \bar{R}_{01}^{ca} & \bar{R}_{01}^{cb} & \bar{R}_{01}^{cc} \end{bmatrix} \\ \bar{\boldsymbol{X}}_{01} &= Re(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{X}_{01} - Im(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{R}_{01} \\ &= \begin{bmatrix} \bar{X}_{01}^{aa} & \bar{X}_{01}^{ab} & \bar{X}_{01}^{ac} \\ \bar{X}_{01}^{ba} & \bar{X}_{01}^{bb} & \bar{X}_{01}^{bc} \\ \bar{X}_{01}^{ca} & \bar{X}_{01}^{cb} & \bar{X}_{01}^{cc} \end{bmatrix} \\ \bar{\boldsymbol{R}}_{12} &= Re(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{R}_{12} + Im(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{X}_{12} \\ &= \begin{bmatrix} \bar{R}_{12}^{aa} & \bar{R}_{12}^{ab} & 0 \\ \bar{R}_{12}^{ba} & \bar{R}_{12}^{bb} & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \bar{\boldsymbol{X}}_{12} &= Re(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{X}_{01} - Im(\boldsymbol{a}\boldsymbol{a}^H) \odot \boldsymbol{R}_{01} \\ &= \begin{bmatrix} \bar{X}_{12}^{aa} & \bar{X}_{12}^{ab} & 0 \\ \bar{X}_{12}^{ba} & \bar{X}_{12}^{bb} & 0 \\ 0 & 0 & 0 \end{bmatrix}\end{aligned}$$

On the other hand, using the circuits and phases for the original network:

$$P_{12}^a(t) = p_2^a(t), P_{12}^b(t) = p_2^b(t) \tag{41a}$$
$$Q_{12}^a(t) = q_2^a(t), Q_{12}^b(t) = q_2^b(t) \tag{41b}$$
$$\begin{aligned} v_1^a(t) = v_2^a(t) &+ 2(\bar{R}_{12}^{aa}P_{12}^a(t) + \bar{R}_{12}^{ab}P_{12}^b(t) \\ &+ \bar{X}_{12}^{aa}Q_{12}^a(t) + \bar{X}_{12}^{ab}Q_{12}^b(t)) \end{aligned} \tag{41c}$$
$$\begin{aligned} v_1^b(t) = v_2^b(t) &+ 2(\bar{R}_{12}^{ba}P_{12}^a(t) + \bar{R}_{12}^{bb}P_{12}^b(t) \\ &+ \bar{X}_{12}^{ba}Q_{12}^a(t) + \bar{X}_{12}^{bb}Q_{12}^b(t)) \end{aligned} \tag{41d}$$

Comparing (39a), (39b), and (39e) with (41a)-(41d), it is seen that the power flow solution for the original unbalanced radial distribution network has not been affected by the introduction of a virtual circuit with a virtual phase. Moreover, for the virtual phase c:

$$P_{12}^c(t) = p_2^c(t) = 0 \tag{42a}$$
$$Q_{12}^c = q_2^c(t) = 0 \tag{42b}$$
$$v_2^c(t) = v_1^c(t) \tag{42c}$$

APPENDIX C

Incidence Matrix Construction for Single-Phase and Multi-Phase Radial Networks:

To illustrate the construction of the incidence matrices $\bar{\boldsymbol{M}}$ and $\bar{\boldsymbol{A}}$ introduced in Section V-D for single-phase and multi-phase radial networks, consider Fig. 17. This figure depicts a radial network with 4 buses (including the head bus 0) and 3 line segments. Specifically, $\mathcal{N} = \{1,2,3\}$, and $\mathcal{L} = \{\ell_1, \ell_2, \ell_3\}$ where $\ell_1 = (0,1)$, $\ell_2 = (0,2)$, and $\ell_3 = (0,3)$.

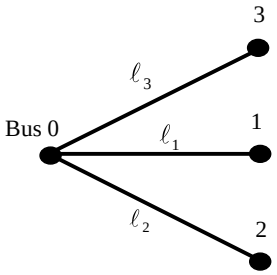

Fig. 17. A four-bus radial network

If the network depicted in Fig. 17 is a single-phase radial network, its incidence matrix $\bar{\boldsymbol{M}}$ takes the following form:

$$\bar{\boldsymbol{M}} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{43}$$

The first row $\boldsymbol{m}_0^T = [1\ 1\ 1]$ of the incidence matrix (43) indicates that bus 0 is the source bus for lines ℓ_1, ℓ_2, and ℓ_3.

Alternatively, suppose the network depicted in Fig. 17 is a multi-phase radial network. This network can be converted into a 3-phase radial network by an appropriate inclusion of virtual circuits and phases for each line segment with fewer than three circuits and phases. The incidence matrix $\bar{\boldsymbol{A}}$ for the resulting 3-phase radial network then takes the following

form:

$$\bar{\boldsymbol{A}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \tag{44}$$

The submatrix consisting of the first three rows of the incidence matrix (44) is denoted by:

$$\boldsymbol{A}_0^T = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \tag{45}$$

The form of (45) indicates that: (i) bus 0 is the source bus for the phase-a circuits of ℓ_1, ℓ_2 and ℓ_3; (ii) bus 0 is the source bus for the phase-b circuits of ℓ_1, ℓ_2 and ℓ_3; and (iii) bus 0 is the source bus for the phase-c circuits of ℓ_1, ℓ_2 and ℓ_3.

APPENDIX D

Determination of the benefit/cost trade-off coefficient μ_ψ from the slider-knob control setting γ_ψ

In economics, the marginal utility of money μ_ψ for a customer ψ is conceptually defined to be the change in the customer's utility (benefit) that would result from an incremental change in the customer's money income. Mathematically, μ_ψ is derived for ψ by: (i) formulating a constrained optimization problem for ψ expressed as the choice of a bundle of goods to maximize utility of good consumption (or usage) subject to a budget constraint; and (ii) defining μ_ψ to be the dual variable corresponding to ψ's budget constraint, evaluated at an optimal solution point.[17]

This study posits a practical way for a DSO to obtain marginal utility of money estimates for households equipped with smart thermostats for their TCL devices. As depicted in Fig. 18, each smart thermostat is assumed to have a knob control that varies along a linear slide ranging from "Comfort" to "Cost". The position of this knob control indicates the relative weight the household places on comfort versus cost.

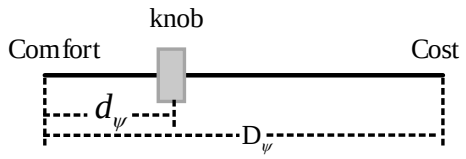

Fig. 18. Illustration of the slider-knob control setting along a linear slide

More precisely, if a household ψ selects a knob setting all the way to the left, this indicates that only comfort matters to ψ; cost is not relevant. Conversely, if ψ selects a knob setting all the way to the right, this indicates that only cost matters to ψ; comfort is not relevant. Intermediate knob settings indicate the household's relative preference for comfort versus cost.

Once household ψ has selected a knob setting on its smart thermostat, the thermostat internally translates this setting into a numerical estimate for μ_ψ as follows. Let the positive total length of the linear slide be denoted by D_ψ (inches) and the distance of the household ψ's slider knob from Comfort be denoted by d_ψ (inches). The unit-free slider-knob control setting γ_ψ for household ψ is then set equal to the relative distance of the slider knob from Comfort:

$$\gamma_\psi = \frac{d_\psi}{\mathrm{D}_\psi} \tag{46}$$

Finally, to avoid extreme cases in which a household ψ only cares about comfort or only cares about cost, it is assumed in this study that d_ψ lies in the open interval (0, D_ψ), implying that γ_ψ lies in the open interval (0, 1). The marginal utility of money μ_ψ (utils/cent) for household ψ is then constructed from γ_ψ in accordance with formula (9).

APPENDIX E

Assumed Constancy of Household Power Factors: Discussion

Section VI assumes the existence of a TCL power-factor function f_ψ for each household $\psi \in \Psi$. Specifically, function f_ψ postulates a linear relationship (16) between household ψ's TCL real power usage and TCL reactive power usage during each sub-period t of an operating period OP. Thus, household ψ's TCL power factor is assumed to be constant during t.

In principle, a household can measure its TCL power factor for each different sub-period t from historical data. However, this TCL power factor will depend on the internal circuits of the household's TCL devices as well as on the external environment. Consequently, accurate measurement could be difficult.

Utilities penalize industrial loads for low power factors; hence, most industrial loads are equipped with power factor correction devices. In general, residential loads are not penalized for low power factors since they are relatively small relative to industrial loads.

On the other hand, household TCL devices tend either to have intrinsically good power factors (close to 1.0) or to come equipped with built-in power factor correction capabilities. Thus, assuming a constant TCL power factor over relatively short time intervals for household TCL devices would appear to be an acceptable modeling strategy for the purposes of the present study.

APPENDIX F

Notes on the Dual Decomposition Method for Constrained Optimization Problems in Lagrangian form

Dual decomposition is a general approach to solving a problem by breaking it up into smaller problems that are then solved separately, either in parallel or sequentially. A brief description is given below for the dual decomposition method

[17]Given various regularity conditions, a dual variable solution for a constraint C with a constraint constant c can be expressed as the rate of change of the optimized objective function with respect to a change in c. For the example at hand, C is a budget constraint restricting total goods expenditures to be less than or equal to a given money income c.

applied in the current study to obtain the proof of Prop. 1 in Section VIII-B.

Consider an optimization problem with separable structure of the form

$$\max_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}},\boldsymbol{y}\in\boldsymbol{\mathcal{Y}}}[f(\boldsymbol{x})+g(\boldsymbol{y})] \tag{47a}$$

$$s.t.\ \boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{y}\leq\boldsymbol{c} \tag{47b}$$

where: $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ are subsets of $\mathbb{R}^d$ and $\mathbb{R}^m$; the component objective functions are differentiable real-valued functions taking the forms f:$\boldsymbol{\mathcal{X}}\rightarrow\mathbb{R}$ and g:$\boldsymbol{\mathcal{Y}}\rightarrow\mathbb{R}$; the constraint constant $\boldsymbol{c}$ is an $s\times 1$ element of $\mathbb{R}^s$; and $\boldsymbol{A}$ and $\boldsymbol{B}$ are $s\times n$ and $s\times m$ constraint matrices with real-valued elements. The *Lagrangian Function* for problem (47) can be represented as follows:

$$L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})=f(\boldsymbol{x})+g(\boldsymbol{y})+\boldsymbol{z}[\boldsymbol{c}-\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{y}] \tag{48}$$

where $\boldsymbol{z}\in\mathbb{R}^s_+$ is a non-negative $1\times s$ vector of dual variables[18] for constraint (47b).

Let ρ denote a positively-valued scalar step size, and let a convergence criterion CC(d) be given for any sequence of vectors in a Euclidean space $\mathbb{R}^d$ with a finite dimension d. The *dual decomposition method* for the approximate solution of problem (47) can then be expressed in four steps as follows:

1) *Initialization:* Set $k=0$, and set $\boldsymbol{z}^0$ to any value in $\mathbb{R}^s_+$.
2) *Update primal vectors:* Set
$$\boldsymbol{x}^k,\boldsymbol{y}^k \;=\; \underset{\boldsymbol{x}\in\boldsymbol{\mathcal{X}},\boldsymbol{y}\in\boldsymbol{\mathcal{Y}}}{\operatorname{argmax}}\ L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}^k) \tag{49}$$
Since $L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}^k)$ is decomposable in $\boldsymbol{x}$ and $\boldsymbol{y}$, the primal vectors $\boldsymbol{x}^k$ and $\boldsymbol{y}^k$ in (49) can equivalently be expressed as:
$$\boldsymbol{x}^k \;=\; \underset{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}}{\operatorname{argmax}}\ \big[f(\boldsymbol{x})-\boldsymbol{z}^k\boldsymbol{A}\boldsymbol{x}\big] \tag{50a}$$
$$\boldsymbol{y}^k \;=\; \underset{\boldsymbol{y}\in\boldsymbol{\mathcal{Y}}}{\operatorname{argmax}}\ \big[g(\boldsymbol{y})-\boldsymbol{z}^k\boldsymbol{B}\boldsymbol{y}\big] \tag{50b}$$
3) *Update dual vector:* Set
$$\boldsymbol{z}^{k+1}=[\boldsymbol{z}^k+\rho(\boldsymbol{A}\boldsymbol{x}^k+\boldsymbol{B}\boldsymbol{y}^k-\boldsymbol{c})^T]^+ \tag{51}$$
where $[\ \cdot\ ]^+$ denotes projection on $\mathbb{R}^s_+$.[19]
4) Stop if convergence of the primal and dual vectors has occurred. Otherwise, assign to k the value $k+1$ and return to Step 2.

APPENDIX G

This appendix provides proofs for Prop. 1 and Prop. 2 given in Section VIII-B, making use of the following classical lemma.[20]

Lemma 1 (Classical): *Suppose $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ in $\mathcal{X}\times\mathbb{R}^m_+$ is a saddle point (29) for the Lagrangian Function $L(\boldsymbol{x},\boldsymbol{\lambda})$ defined in (25). Then:*

- **[L1.A]** $[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}^*)]\geq\boldsymbol{0}$
- **[L1.B]** $\boldsymbol{\lambda}^*[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}^*)]=0$
- **[L1.C]** *$\boldsymbol{x}^*$ is a solution for the Primal Problem (23)*

Proof of Lemma 1: By assumption, $L(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)\leq L(\boldsymbol{x}^*,\boldsymbol{\lambda})$ for all $\boldsymbol{\lambda}\in\mathbb{R}^m_+$, which implies

$$[\boldsymbol{\lambda}^*-\boldsymbol{\lambda}][\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}^*)]\leq 0,\ \forall\boldsymbol{\lambda}\in\mathbb{R}^m_+ \tag{52}$$

Since the components of $\boldsymbol{\lambda}$ can take on arbitrarily large positive values, (52) implies that

$$[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}^*)]\geq\boldsymbol{0}\,, \tag{53}$$

which establishes [L1.A]. Letting $\boldsymbol{\lambda}=\boldsymbol{0}$ in (52),

$$\boldsymbol{\lambda}^*[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}^*)]\leq 0 \tag{54}$$

Since $\boldsymbol{\lambda}^*\geq\boldsymbol{0}$, conditions (53) and (54) together imply that

$$\boldsymbol{\lambda}^*[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}^*)]=0 \tag{55}$$

This establishes [L1.B]. Substituting (55) into the definition of the saddle point property (29) for $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$,

$$F(\boldsymbol{x})\leq F(\boldsymbol{x})+\boldsymbol{\lambda}^*[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x})]\leq F(\boldsymbol{x}^*) \tag{56}$$

for all $\boldsymbol{x}\in\boldsymbol{\mathcal{X}}$ that satisfy $[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}]\geq\boldsymbol{0}$. It follows that $\mathbf{x}^*$ is a solution for the Primal Problem (23). This establishes [L1.C]. Q.E.D.

Proposition 1 (Classical): *A point $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ in $\mathcal{X}\times\mathbb{R}^m_+$ is a saddle point (29) for the Lagrangian Function $L(\boldsymbol{x},\boldsymbol{\lambda})$ defined in (25) if and only if:*

- **[P1.A]** *$\boldsymbol{x}^*$ is a solution for the Primal Problem (23) ;*
- **[P1.B]** *$\boldsymbol{\lambda}^*$ is a solution for the Dual Problem (32) ;*
- **[P1.C]** *$D(\boldsymbol{\lambda}^*)=F(\boldsymbol{x}^*)$ (strong duality).*

Proof of Proposition 1: Necessity. Suppose that $(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)$ in $\mathcal{X}\times\mathbb{R}^m_+$ is a saddle point (29) for the Lagrangian Function $L(\boldsymbol{x},\boldsymbol{\lambda})$ defined in (25). Then it immediately follows from the Lemma 1 implication [L1.C] that [P1.A] in Prop. 1 holds. Define

$$\boldsymbol{\mathcal{X}}^o=\{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}\mid\boldsymbol{g}(\boldsymbol{x})\leq\boldsymbol{c}\} \tag{57}$$

Making use of the Lemma 1 implications [L1.A] and [L1.B], it then holds for any $\boldsymbol{\lambda}\in\mathbb{R}^m_+$ that:

$$\begin{aligned}
F(\boldsymbol{x}^*)&=\max_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}^o}F(\boldsymbol{x})\\
&\leq\max_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}^o}\big[F(\boldsymbol{x})+\boldsymbol{\lambda}^*[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x})]\\
&\leq\max_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}}\big[F(\boldsymbol{x})+\boldsymbol{\lambda}^*[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x})]\\
&=D(\boldsymbol{\lambda}^*)\\
&=L(\boldsymbol{x}^*,\boldsymbol{\lambda}^*)\\
&=F(\boldsymbol{x}^*)+\boldsymbol{\lambda}^*[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}^*)]\\
&=F(\boldsymbol{x}^*)\\
&\leq F(\boldsymbol{x}^*)+\boldsymbol{\lambda}[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x}^*)]\\
&\leq\max_{\boldsymbol{x}\in\boldsymbol{\mathcal{X}}}\big[F(\boldsymbol{x})+\boldsymbol{\lambda}[\boldsymbol{c}-\boldsymbol{g}(\boldsymbol{x})]\big]\\
&=D(\boldsymbol{\lambda})
\end{aligned}$$

[18]For constrained optimization problems expressible in Lagrangian form, dual variables are equivalent to Lagrange multipliers.

[19]For the particular projection set $\mathbb{R}^s_+$ at hand, this projection operation reduces to the following: Given any vector $\boldsymbol{w}=[w_1,\ldots w_s]\in\mathbb{R}^s$, $[\boldsymbol{w}]^+=[w_1^+,\ldots,w_s^+]$, where $w_j^+=\max\{w_j,0\}$ for $j=1,\ldots s$.

[20]Lemma 1 is a corollary of more generally established theorems in classical NP theory; see, e.g., Bertsekas [42]. For completeness, Lemma 1 is stated and proved here in full.

It follows that $\boldsymbol{\lambda}^*$ is a solution for the Dual Problem (32), and $D(\boldsymbol{\lambda}^*) = F(\boldsymbol{x}^*)$. Thus, [P1.B] and [P1.C] in Prop. 1 hold. Q.E.D.

Proof of Proposition 1: Sufficiency. Conversely, suppose conditions [P1.A]-[P1.C] in Prop. 1 hold. It then follows immediately from $[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*] \geq \boldsymbol{0}$ and $D(\boldsymbol{\lambda}^*) = F(\boldsymbol{x}^*)$ that $\boldsymbol{\lambda}^*[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*)] = \boldsymbol{0}$. Consequently, for all $\boldsymbol{\lambda} \in \mathbb{R}^m_+$,

$$L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = F(\boldsymbol{x}^*) \leq F(\boldsymbol{x}^*) + \boldsymbol{\lambda}[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*)] = L(\boldsymbol{x}^*, \boldsymbol{\lambda})$$

Also, for all $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$,

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{\lambda}^*) &\leq \max_{\boldsymbol{x} \in \boldsymbol{\mathcal{X}}} L(\boldsymbol{x}, \boldsymbol{\lambda}^*) \\ &= \mathrm{D}(\boldsymbol{\lambda}^*) \\ &= F(\boldsymbol{x}^*) \\ &= F(\boldsymbol{x}^*) + \boldsymbol{\lambda}^*[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*)] \\ &= L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \end{aligned}$$

Thus, $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ in $\mathcal{X} \times \mathbb{R}^m_+$ is a saddle point (29) for the Lagrangian Function $L(\boldsymbol{x}, \boldsymbol{\lambda})$ defined in (25). Q.E.D.

Proposition 2: *Suppose $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ in $\mathcal{X} \times \mathbb{R}^m_+$ is a saddle point for the Lagrangian Function $L(\boldsymbol{x}, \boldsymbol{\lambda})$ defined in (25), where $\boldsymbol{x}^* = \boldsymbol{\mathcal{P}}^*(\mathcal{K})$. Suppose, also, that $\boldsymbol{x}^*$ uniquely maximizes $L(\boldsymbol{x}, \boldsymbol{\lambda}^*)$ with respect to $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$. Define $\boldsymbol{\pi}^*(\mathcal{K}) = \{\boldsymbol{\pi}^*_\psi(\mathcal{K}) \mid \psi \in \Psi\}$, where the price-to-go sequence $\boldsymbol{\pi}^*_\psi(\mathcal{K})$ for each household $\psi \in \Psi$ takes the following form:*

$$\begin{aligned} \boldsymbol{\pi}^*_\psi(\mathcal{K}) \;=\; & \boldsymbol{LMP}(\mathcal{K}) + \frac{1}{\mu_\psi S_{\mathsf{base}} \Delta t} \Big[\boldsymbol{\lambda}^*_{\bar{P}}(\mathcal{K}) \\ & - 2 \cdot \boldsymbol{r}_D(i, N^{\mathsf{ph}}_\psi)^T \big[\boldsymbol{\Lambda}^*_{v_{\mathsf{max}}}(\mathcal{K}) - \boldsymbol{\Lambda}^*_{v_{\mathsf{min}}}(\mathcal{K}) \big]^T \\ & - 2 \cdot \boldsymbol{x}_D(i, N^{\mathsf{ph}}_\psi)^T \big[\boldsymbol{\Lambda}^*_{v_{\mathsf{max}}}(\mathcal{K}) - \boldsymbol{\Lambda}^*_{v_{\mathsf{min}}}(\mathcal{K}) \big]^T \boldsymbol{H}_\psi(\mathcal{K}) \Big] \end{aligned} \tag{58}$$

The pairing $(\boldsymbol{\mathcal{P}}^*(\mathcal{K}), \boldsymbol{\pi}^*(\mathcal{K}))$ then constitutes a TES equilibrium for OP.

Proof of Proposition 2:

By assumption, $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) \in \mathcal{X} \times \mathbb{R}^m_+$ is a saddle point for the Lagrangian Function $L(\boldsymbol{x}, \boldsymbol{\lambda})$ in (25) and $\boldsymbol{x}^*$ uniquely maximizes this Lagrangian Function over $\mathcal{X}$, given $\boldsymbol{\lambda}^*$. Thus,

$$\boldsymbol{x}^* = \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}^*) \tag{59}$$

The Lagrangian Function (25), given explicitly in (28), can equivalently be expressed in the following manner:

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{\lambda}) &= \sum_{\psi \in \Psi} \Big[F_\psi(\boldsymbol{x}_\psi) - \boldsymbol{\lambda}_{\bar{P}}(\mathcal{K}) \boldsymbol{x}_\psi \\ &+ \sum_{t \in \mathcal{K}} 2[\boldsymbol{\lambda}_{v_{\mathsf{max}}}(t) - \boldsymbol{\lambda}_{v_{\mathsf{min}}}(t)] \boldsymbol{h}_\psi(t, x_\psi(t)) \Big] + C(\mathcal{K}) \\ &= \sum_{\psi \in \Psi} \Big[F_\psi(\boldsymbol{x}_\psi) - \boldsymbol{\lambda}_{\bar{P}}(\mathcal{K}) \boldsymbol{x}_\psi \\ &+ 2 \cdot \boldsymbol{r}_D(i, N^{\mathsf{ph}}_\psi)^T \big(\boldsymbol{\Lambda}_{v_{\mathsf{max}}}(\mathcal{K}) - \boldsymbol{\Lambda}_{v_{\mathsf{min}}}(\mathcal{K}) \big)^T \boldsymbol{x}_\psi \\ &+ 2 \cdot \boldsymbol{x}_D(i, N^{\mathsf{ph}}_\psi)^T \big(\boldsymbol{\Lambda}_{v_{\mathsf{max}}}(\mathcal{K}) - \boldsymbol{\Lambda}_{v_{\mathsf{min}}}(\mathcal{K}) \big)^T \boldsymbol{H}_\psi(\mathcal{K}) \boldsymbol{x}_\psi \Big] \\ &+ C(\mathcal{K}) \end{aligned} \tag{60}$$

where

$$\begin{aligned} C(\mathcal{K}) &= \boldsymbol{\lambda}_{\bar{\boldsymbol{P}}}(\mathcal{K}) [\bar{\boldsymbol{P}}(\mathcal{K}) - \boldsymbol{P}^{\mathsf{non}}_\psi(\mathcal{K})] \\ &+ \sum_{t \in \mathcal{K}} \boldsymbol{\lambda}_{v_{\mathsf{max}}}(t) [\boldsymbol{v}_{\mathsf{max}}(t) - \boldsymbol{v}^{\mathsf{non}}(t)] \\ &+ \sum_{t \in \mathcal{K}} \boldsymbol{\lambda}_{v_{\mathsf{min}}}(t) [-\boldsymbol{v}_{\mathsf{min}}(t) + \boldsymbol{v}^{\mathsf{non}}(t)] \end{aligned}$$

Note that $C(\mathcal{K})$ does not depend on $\boldsymbol{x}$. Using form (60) for $L(\boldsymbol{x}, \boldsymbol{\lambda})$, it follows that:

$$\begin{aligned} \boldsymbol{x}^* &= \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}^*) \\ &= \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \sum_{\psi \in \Psi} \Big[F_\psi(\boldsymbol{x}_\psi) - \boldsymbol{\lambda}^*_{\bar{P}}(\mathcal{K}) \boldsymbol{x}_\psi \\ &+ 2 \cdot \boldsymbol{r}_D(i, N^{\mathsf{ph}}_\psi)^T \big[\boldsymbol{\Lambda}^*_{v_{\mathsf{max}}}(\mathcal{K}) - \boldsymbol{\Lambda}^*_{v_{\mathsf{min}}}(\mathcal{K}) \big]^T \boldsymbol{x}_\psi \\ &+ 2 \cdot \boldsymbol{x}_D(i, N^{\mathsf{ph}}_\psi)^T \big[\boldsymbol{\Lambda}^*_{v_{\mathsf{max}}}(\mathcal{K}) - \boldsymbol{\Lambda}^*_{v_{\mathsf{min}}}(\mathcal{K}) \big]^T \boldsymbol{H}_\psi(\mathcal{K}) \boldsymbol{x}_\psi \Big] \end{aligned} \tag{61}$$

Substituting into expression (61) the definition for $F_\psi(\boldsymbol{x}_\psi)$ from (23) and the form of the price-to-go sequence $\boldsymbol{\pi}^*_\psi(\mathcal{K})$ given by (33), problem (61) can equivalently be expressed as

$$\boldsymbol{x}^* = \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \sum_{\psi \in \Psi} \big[U_\psi(\boldsymbol{x}_\psi) - \mu_\psi \boldsymbol{\pi}^*_\psi(\mathcal{K}) \boldsymbol{x}_\psi S_{\mathsf{base}} \Delta t \big] \tag{62}$$

Note that problem (62) can be solved in a distributed manner by having each household $\psi \in \Psi$ separately solve

$$\boldsymbol{x}^*_\psi = \operatorname*{argmax}_{\boldsymbol{x}_\psi \in \mathcal{X}_\psi(\mathcal{K})} \big[U_\psi(\boldsymbol{x}_\psi) - \mu_\psi \boldsymbol{\pi}^*_\psi(\mathcal{K}) \boldsymbol{x}_\psi S_{\mathsf{base}} \Delta t \big] \tag{63}$$

However, problem (63) is precisely the optimization problem (12) formulated for each household $\psi \in \Psi$ in Section VI, with $\boldsymbol{x}_\psi = \mathcal{P}_\psi(\mathcal{K})$. Thus, using the notation introduced in (15) for the solution of (12), $\boldsymbol{x}^*_\psi = \boldsymbol{\mathcal{P}}_\psi(\boldsymbol{\pi}^*_\psi(\mathcal{K}))$ for each $\psi \in \Psi$, which in turn implies that $\boldsymbol{x}^* = \boldsymbol{\mathcal{P}}(\boldsymbol{\pi}^*(\mathcal{K}))$.

By the Lemma 1 implication [L1.C], $\boldsymbol{x}^*$ is also a solution for the Primal Problem (23) formulated in Section VII-B. Thus, using the notation introduced in Section VII-B for any solution to this Primal Problem, $\boldsymbol{x}^* = \mathcal{P}^*(\mathcal{K})$.

It follows from the above analysis that $(\boldsymbol{\mathcal{P}}^*(\mathcal{K}), \boldsymbol{\pi}^*(\mathcal{K}))$ is a TES equilibrium for operating period OP; that is, $\boldsymbol{\mathcal{P}}^*(\mathcal{K}) = \boldsymbol{\mathcal{P}}(\boldsymbol{\pi}^*(\mathcal{K}))$. This establishes Prop. 2. Q.E.D.

APPENDIX H

This appendix provides a proof of Prop. 3 in Section VIII-C, making use of three additional lemmas.

Lemma 2 (Classical): *The dual function D:$\mathbb{M} \to \mathbb{R}$ defined by (30) and (31) is convex over any convex subset of $\mathbb{M}$.*

Proof of Lemma 2: Let $\boldsymbol{\lambda}'$ and $\boldsymbol{\lambda}''$ be elements of $\mathbb{M}$ such that the line segment connecting $\boldsymbol{\lambda}'$ and $\boldsymbol{\lambda}''$ lies entirely in $\mathbb{M}$, and let a be any point in the interval $[0, 1]$. Using the definition of $D(\boldsymbol{\lambda})$, and the linearity of L with respect to $\boldsymbol{\lambda}$, it follows that

$$\begin{aligned} D(a\boldsymbol{\lambda}' + [1-a]\boldsymbol{\lambda}'') &= \max_{\boldsymbol{x} \in \mathcal{X}} L(\boldsymbol{x}, a\boldsymbol{\lambda}' + [1-a]\boldsymbol{\lambda}'') \\ &= \max_{\boldsymbol{x} \in \mathcal{X}} \big[aL(\boldsymbol{x}, \boldsymbol{\lambda}') + [1-a] L(\boldsymbol{x}, \boldsymbol{\lambda}'') \big] \\ &\leq a \cdot D(\boldsymbol{\lambda}') + [1-a] \cdot D(\boldsymbol{\lambda}'') \end{aligned}$$

Q.E.D.

Lemma 3 (Classical): *Suppose X is a convex subset of $\mathbb{R}^k$ for some $k \geq 1$, and suppose f:$\mathbb{R}^k \to \mathbb{R}$ is convex over X and continuously differentiable in a neighborhood of a point $x^* \in X$. Then, $\boldsymbol{x}^*$ is a solution for*

$$\min_{\boldsymbol{x} \in X} f(\boldsymbol{x}) \tag{64}$$

if and only if

$$\nabla f(\boldsymbol{x}^*)^T[\boldsymbol{x}-\boldsymbol{x}^*] \geq 0,\ \forall \boldsymbol{x} \in X \tag{65}$$

Proof of Lemma 3:[21]

Suppose $\boldsymbol{x}^* \in X$ satisfies (65). Since $f(\boldsymbol{x})$ is convex over X and differentiable at $\boldsymbol{x}^* \in X$, it follows that

$$f(\boldsymbol{x}) - f(\boldsymbol{x}^*) \geq \nabla f(\boldsymbol{x}^*)^T[\boldsymbol{x}-\boldsymbol{x}^*] \geq 0,\ \forall \boldsymbol{x} \in X\ , \tag{66}$$

which implies $\boldsymbol{x}^*$ is a solution for (64.

The converse will be established using proof by contradiction. Suppose $\boldsymbol{x}^* \in X$ solves (64) but condition (65) fails to hold. Then there must exist some $\boldsymbol{x}' \in X$ such that

$$\nabla f(\boldsymbol{x}^*)^T(\boldsymbol{x}'-\boldsymbol{x}^*) < \boldsymbol{0} \tag{67}$$

By convexity of X, the point $\boldsymbol{z}(\alpha) = \alpha\boldsymbol{x}' + (1-\alpha)\boldsymbol{x}^*$ lies in X for any $\alpha \in [0,1]$. Note, also, that

$$\frac{d}{d\alpha} f(\boldsymbol{z}(\alpha))\Big|_{\alpha=0} = \nabla f(\boldsymbol{x}^*)^T(\boldsymbol{x}'-\boldsymbol{x}^*) < 0 \tag{68}$$

Thus, by continuous differentiability of $f(\boldsymbol{x})$ in a neighborhood of $\boldsymbol{x}^*$, the derivative of $f(\boldsymbol{z}(\alpha))$ with respect to α must be strictly negative for all sufficiently small positive values for α. It follows that $f(\boldsymbol{z}(\alpha)) < f(\boldsymbol{z}(0)) = f(\boldsymbol{x}^*)$ for all sufficiently small values of $\alpha > 0$, which contradicts the presumption that $\boldsymbol{x}^*$ solves (64). Q.E.D.

Lemma 4: *Suppose conditions [P3.A] and [P3.B] in Prop. 3 hold. Then:*

- **[L4.A]** *The dual function $D(\boldsymbol{\lambda})$ is continuously differentiable over the interior $\mathbb{R}^m_{++}$ of its domain $\mathbb{R}^m_+$ and right continuously differentiable at any boundary point of its domain $\mathbb{R}^m_+$;*
- **[L4.B]** *For each $\boldsymbol{\lambda} \in \mathbb{M}$,*

$$\nabla \boldsymbol{D}_+(\boldsymbol{\lambda}) = [\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}(\boldsymbol{\lambda}))]^T \tag{69}$$

  where $\nabla_+\boldsymbol{D}(\boldsymbol{\lambda})$ in (69) denotes the gradient of $D(\boldsymbol{\lambda})$ at each interior point $\boldsymbol{\lambda}$ in $\mathbb{R}^m_{++}$ and the right-hand gradient of $D(\boldsymbol{\lambda})$ at each boundary point $\boldsymbol{\lambda}$ in $\mathbb{R}^m_+$.
- **[L4.C]** *Strong duality holds for the primal and dual problems (23) and (32).*

Proof of Lemma 4: Given conditions [P3.A] and [P.3B], implications [L4.A] and [L4.B] follow from Bertsekas [42, Prop. 6.1.1]. Finally, implication [L4.C] follows from Lemma 2, [L4.A], Lemma 3, Bertsekas [43, Sol. 8.1], and Bertsekas [42, Prop. 5.1.5]. Q.E.D.

Proposition 3: *Suppose the Primal Problem (23) and the dual decomposition algorithm DDA satisfy the following three conditions:*

- **[P3.A]** *$\mathcal{X}$ is compact, and the objective function $F(\boldsymbol{x})$ and constraint function $\boldsymbol{g}(\boldsymbol{x})$ are continuous over $\mathcal{X}$.*
- **[P3.B]** *For every $\boldsymbol{\lambda} \in \mathbb{R}^m_+$, the Lagrangian Function $L(\boldsymbol{x},\boldsymbol{\lambda})$ defined in (25) achieves a finite maximum at a unique point $\boldsymbol{x}(\boldsymbol{\lambda}) \in \mathcal{X}$, implying the dual function domain in (31) satisfies $\mathbb{M} = \mathbb{R}^m_+$.*
- **[P3.C]** *The sequence $(\boldsymbol{x}^y, \boldsymbol{\lambda}^y)$ determined by the dual decomposition algorithm DDA converges to a limit point $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ as the iteration time y approaches $+\infty$.*

Then the limit point $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ is a saddle point (29) for the Lagrangian Function (25), and this saddle point determines a TES equilibrium for OP.

Proof of Proposition 3:

By the Cauchy Convergence Criterion, if $(\boldsymbol{x}^y, \boldsymbol{\lambda}^y)$ converges to a limit point $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$, then:

$$\lim_{y\to\infty} \left[\boldsymbol{\lambda}^y - \boldsymbol{\lambda}^{y+1}\right] = \boldsymbol{0}\,; \tag{70a}$$

$$\lim_{y\to\infty} \left[\boldsymbol{x}^y - \boldsymbol{x}^{y+1}\right] = \boldsymbol{0} \tag{70b}$$

Thus, combining the updating of dual variables in **S4** of the dual decomposition algorithm DDA with (70), the following condition must hold in the limit:

$$\boldsymbol{\lambda}^* = \left[\boldsymbol{\lambda}^* + [\boldsymbol{g}(\boldsymbol{x}^*) - \boldsymbol{c}]^T\boldsymbol{B}\right]^+ \tag{71}$$

Since $\boldsymbol{\lambda}^* \in \mathbb{R}^m_+$ and $\boldsymbol{B}$ is a positive definite diagonal matrix, equation (71) is equivalent to:

$$[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*)]_i \geq 0, \text{if } [\boldsymbol{\lambda}^*]_i = 0 \tag{72a}$$

$$[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*)]_i = 0, \text{if } [\boldsymbol{\lambda}^*]_i > 0 \tag{72b}$$

where $[\,\cdot\,]_i$ denotes the i-th element in $[\,\cdot\,]$.

Lemma 2 and condition [P3.C] imply that the dual function $D(\boldsymbol{\lambda})$ is convex over $\mathbb{R}^m_+$. Given this convexity, Lemma 3 and implications [L4.A] and [L4.B] in Lemma 4 imply that $\boldsymbol{\lambda}^*$ is the optimal solution for the Dual Problem (32) if and only if:

$$(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*)[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*)] \geq 0, \forall \boldsymbol{\lambda} \in \mathbb{R}^m_+ \tag{73}$$

From (72), for all $\boldsymbol{\lambda} \in \mathbb{R}^m_+$ we have:

$$[\boldsymbol{\lambda} - \boldsymbol{\lambda}^*]_i[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*)]_i \geq 0, \text{if } [\boldsymbol{\lambda}^*]_i = 0 \tag{74a}$$

$$[\boldsymbol{\lambda} - \boldsymbol{\lambda}^*]_i[\boldsymbol{c} - \boldsymbol{g}(\boldsymbol{x}^*)]_i = 0, \text{if } [\boldsymbol{\lambda}^*]_i > 0 \tag{74b}$$

It follows that (73) holds. Hence, we can conclude that $\boldsymbol{\lambda}^*$ is the optimal solution for the Dual Problem.

According to steps **S2** and **S3** of the dual decomposition algorithm DDA, and form (58) for the DSO's offered price-to-go sequence in the limit, we have $\boldsymbol{x}^* = \{\boldsymbol{x}^*_\psi = \mathcal{P}_\psi\big(\boldsymbol{\pi}^*_\psi(\mathcal{K})\big) \mid \psi \in \Psi\}$. Making use of (60) and assumption [P3.B], $\boldsymbol{x}^*$ can equivalently be expressed as:

$$\boldsymbol{x}^* = \boldsymbol{x}(\boldsymbol{\lambda}^*) = \underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmax}}\, L(\boldsymbol{x}, \boldsymbol{\lambda}^*) \tag{75}$$

By the strong duality implication [L4.C] in Lemma 4, it then follows that $\boldsymbol{x}^* = \boldsymbol{x}(\boldsymbol{\lambda}^*)$ is the optimal solution for the Primal Problem.

From Prop. 1, it follows that $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ is a saddle point (29) for the Lagrangian Function (25). Finally, by Prop. 2, this saddle point determines a TES equilibrium for OP. Q.E.D.

[21] Lemma 3 is an implication of [42, Prop. 2.1.2].

APPENDIX I

This appendix provides a proof of Prop. 4 in Section VIII-C, making use of two additional lemmas.

Lemma 5: *Suppose the assumptions of Prop. 4 hold. Then for all* $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^m$*:*

$$\begin{aligned} &D(\boldsymbol{\lambda}_1) - D(\boldsymbol{\lambda}_2) \\ &\quad \leq (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)[\nabla D_+(\boldsymbol{\lambda}_2)]^T + \frac{1}{2}||\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2||_{\boldsymbol{J}}^2 \end{aligned} \tag{76}$$

Proof of Lemma 5: Define a function w:$\mathbb{R}_+^m \rightarrow \mathbb{R}$ by $w(\boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{\lambda}\boldsymbol{J}\boldsymbol{\lambda}^T - D(\boldsymbol{\lambda})$, where $D(\boldsymbol{\lambda})$ is the dual function defined by (30). By assumption, condition [P4.A] holds. It thus follows from Lemma 4 that the right-hand gradient $\nabla w_+(\boldsymbol{\lambda})$ is well-defined at each $\boldsymbol{\lambda} \in \mathbb{R}_+^m$. It will next be shown that $w(\boldsymbol{\lambda})$ is convex over $\mathbb{R}_+^m$ if and only if $\nabla w_+(\boldsymbol{\lambda})$ is *monotonic*; that is, if and only if:

$$\big\langle \nabla w_+(\boldsymbol{\lambda}_1) - \nabla w_+(\boldsymbol{\lambda}_2), \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 \big\rangle \geq 0, \forall \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^m \tag{77}$$

where $\big\langle\ ,\ \big\rangle$ denotes vector inner product.

Suppose, first, that w is convex over $\mathbb{R}_+^m$. Then for all $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^m$ it follows that:

$$w(\boldsymbol{\lambda}_1) \geq w(\boldsymbol{\lambda}_2) + (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\nabla w_+(\boldsymbol{\lambda}_2)^T \tag{78a}$$

$$w(\boldsymbol{\lambda}_2) \geq w(\boldsymbol{\lambda}_1) + (\boldsymbol{\lambda}_2 - \boldsymbol{\lambda}_1)\nabla w_+(\boldsymbol{\lambda}_1)^T \tag{78b}$$

Adding together these two inequalities, and rearranging terms, then gives (77).

Conversely, suppose (77) holds for all $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^m$. Define $\boldsymbol{\lambda}_3(\tau) = \boldsymbol{\lambda}_2 + \tau[\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2]$ for $0 \leq \tau \leq 1$. Then:

$$\begin{aligned} w(\boldsymbol{\lambda}_1) &= w(\boldsymbol{\lambda}_2) + \int_0^1 (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\nabla w_+(\boldsymbol{\lambda}_3(\tau))^T d\tau \\ &= w(\boldsymbol{\lambda}_2) + (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\nabla w_+(\boldsymbol{\lambda}_2)^T \\ &\quad + \int_0^1 \big\langle \nabla w_+(\boldsymbol{\lambda}_3(\tau)) - \nabla w_+(\boldsymbol{\lambda}_2), \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 \big\rangle d\tau \\ &= w(\boldsymbol{\lambda}_2) + (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\nabla w_+(\boldsymbol{\lambda}_2)^T \\ &\quad + \int_0^1 \frac{1}{\tau}\big\langle \nabla w_+(\boldsymbol{\lambda}_3(\tau)) - \nabla w_+(\boldsymbol{\lambda}_2), \boldsymbol{\lambda}_3(\tau) - \boldsymbol{\lambda}_2 \big\rangle d\tau \\ &\geq w(\boldsymbol{\lambda}_2) + (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\nabla w_+(\boldsymbol{\lambda}_2)^T \end{aligned}$$

It follows that w is convex over $\mathbb{R}_+^m$.

By assumption, condition [P4.C] holds. Rearranging terms in [P4.C] gives

$$\big\langle (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\boldsymbol{J} - [\nabla D_+(\boldsymbol{\lambda}_1) - \nabla D_+(\boldsymbol{\lambda}_2)], \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 \big\rangle \geq 0 \tag{79}$$

Inequality (79) implies that (77) holds, i.e., $\nabla w_+(\boldsymbol{\lambda})$ is monotonic. It follows that w is convex over $\mathbb{R}_+^m$. Thus, using the symmetry of $\boldsymbol{J}$:

$$\begin{aligned} &w(\boldsymbol{\lambda}_1) = \frac{1}{2}\boldsymbol{\lambda}_1\boldsymbol{J}\boldsymbol{\lambda}_1^T - D(\boldsymbol{\lambda}_1) \\ &\geq w(\boldsymbol{\lambda}_2) + (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\nabla w_+(\boldsymbol{\lambda}_2)^T \\ &= \frac{1}{2}\boldsymbol{\lambda}_2\boldsymbol{J}\boldsymbol{\lambda}_2^T - D(\boldsymbol{\lambda}_2) + < \boldsymbol{\lambda}_2\boldsymbol{J} - \nabla D_+(\boldsymbol{\lambda}_2), \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 > \\ &= -D(\boldsymbol{\lambda}_2) - < \nabla D_+(\boldsymbol{\lambda}_2), \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 > \\ &\quad - \frac{1}{2}||\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2||_{\boldsymbol{J}}^2 + \frac{1}{2}\boldsymbol{\lambda}_1\boldsymbol{J}\boldsymbol{\lambda}_1^T \end{aligned} \tag{80}$$

Multiplying each side of (80) by -1 with reversed inequality sign, and rearranging terms, it is seen that (76) holds. Q.E.D.

Lemma 6 (Classical): *Given any real symmetric positive-definite* $n \times n$ *matrix* $\boldsymbol{A}$*, there exists a real invertible symmetric* $n \times n$ *matrix* $\boldsymbol{Q}$ *such that* $\boldsymbol{Q}^2 = \boldsymbol{A}$*. The matrix* $\boldsymbol{Q}$ *is called the symmetric square root of* $\boldsymbol{A}$*, denoted by* $\boldsymbol{A}^{\frac{1}{2}}$*, and the inverse of* $\boldsymbol{Q}$ *is denoted by* $\boldsymbol{A}^{-\frac{1}{2}}$*.*

Proof of Lemma 6:[22] Let $\mu_1, \ldots, \mu_n$ denote the n real positive eigenvalues of $\boldsymbol{A}$, and let $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ denote the corresponding non-zero real orthogonal eigenvectors of $\boldsymbol{A}$ expressed as $n \times 1$ column vectors with normalized unit lengths. For each $k = 1, \ldots, n$, let $\mu_k^{\frac{1}{2}}$ denote the positive square root of μ_k. Define:

$$\boldsymbol{Q} = \sum_{k=1}^{n} \big[\mu_k^{\frac{1}{2}} \boldsymbol{z}_k \boldsymbol{z}_{\boldsymbol{k}}^T\big] \tag{81}$$

Then

$$\begin{aligned} \boldsymbol{Q}^2 &= \big[\sum_{i=1}^{n}\sum_{k=1}^{n} \mu_i^{\frac{1}{2}}\mu_k^{\frac{1}{2}} \boldsymbol{z}_i \boldsymbol{z}_i^T \boldsymbol{z}_k \boldsymbol{z}_k^T\big] \\ &= \sum_{k=1}^{n} \big[\mu_k \boldsymbol{z}_k \boldsymbol{z}_{\boldsymbol{k}}^T\big] \\ &= \boldsymbol{A} \end{aligned} \tag{82}$$

Proposition 4: *Suppose the Primal Problem (23) and the Dual Function (30) satisfy the following four conditions:*

- **[P4.A]** *Conditions [P3.A] and [P3.B] both hold;*
- **[P4.B]** *The Lagrangian Function (25) has a saddle point* $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ *in* $\mathcal{X} \times \mathbb{R}_+^m$*;*
- **[P4.C]** ***Extended Lipschitz Continuity Condition:*** *There exists a real symmetric positive-definite* $m \times m$ *matrix* $\boldsymbol{J}$ *such that, for all* $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^m$*,*

  $$\big\langle \nabla D_+(\boldsymbol{\lambda}_1) - \nabla D_+(\boldsymbol{\lambda}_2), \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 \big\rangle \ \leq\ ||\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2||_{\boldsymbol{J}}^2$$

  where: $\nabla D_+(\boldsymbol{\lambda})$ *denotes the gradient of the dual function* $D(\boldsymbol{\lambda})$ *in (30) for* $\boldsymbol{\lambda} \in \mathbb{R}_{++}^m$ *and the right-hand gradient of* $D(\boldsymbol{\lambda})$ *at boundary points of* $\mathbb{R}_+^m$*:* $\big\langle\ ,\ \big\rangle$ *denotes vector inner product; and* $||\cdot||_{\boldsymbol{J}}^2 = (\cdot)\boldsymbol{J}(\cdot)^T$
- **[P4.D]** The matrix $[\boldsymbol{I} - \boldsymbol{J}\boldsymbol{B}]$ *is positive semi-definite, where* $\boldsymbol{I}$ *denotes an* $m \times m$ *identity matrix, and where* $\boldsymbol{B}$ *is the* $m \times m$ *diagonal positive-definite matrix defined in step* ***S4*** *of the dual decomposition algorithm DDA.*

Then the primal-dual point $(\boldsymbol{x}^y, \boldsymbol{\lambda}^y)$ *determined by the dual decomposition algorithm DDA at iteration time* y *converges to a saddle point as* $y \rightarrow +\infty$*.*

Proof of Proposition 4: By assumption [P4.A] and Lemma 4, the domain of the dual function (30) is given by $\mathbb{M} = \mathbb{R}_+^m$. For any $\boldsymbol{u} \in \mathbb{R}_+^m$, define $\boldsymbol{\lambda}(\boldsymbol{u}) \in \mathbb{R}_+^m$ as follows:

$$\boldsymbol{\lambda}(\boldsymbol{u}) = \boldsymbol{u}\boldsymbol{B}^{\frac{1}{2}} \tag{83}$$

where: $\boldsymbol{B}$ denotes the diagonal positive-definite $m \times m$ matrix defined in step $\boldsymbol{S4}$ of the dual decomposition algorithm DDA; and $\boldsymbol{B}^{\frac{1}{2}}$ denotes the invertible symmetric square root of $\boldsymbol{B}$.

[22] This proof is adapted from Bertsekas [42, Props. A.17(c) & Prop. A.21].

Then, the Dual Problem (32) can equivalently be expressed as follows:

$$\min_{\boldsymbol{u}\in\mathbb{R}^m_+} h(\boldsymbol{u}) \tag{84}$$

where $h(\boldsymbol{u}) = D(\boldsymbol{\lambda}(\boldsymbol{u})) = D(\boldsymbol{u}\boldsymbol{B}^{\frac{1}{2}})$.

By assumption [P4.A], Lemma 4, and step **S4** of the dual decomposition algorithm DDA,

$$\boldsymbol{\lambda}^{y+1} = \big[\boldsymbol{\lambda}^y - \nabla D_+(\boldsymbol{\lambda}^y)\boldsymbol{B}\big]^+ \tag{85}$$

For each iteration time y, define $\boldsymbol{u}^y = \boldsymbol{\lambda}^y\boldsymbol{B}^{-\frac{1}{2}}$ and $\boldsymbol{u}^{y+1} = \boldsymbol{\lambda}^{y+1}\boldsymbol{B}^{-\frac{1}{2}}$, where $\boldsymbol{B}^{-\frac{1}{2}}$ denotes the inverse of $\boldsymbol{B}^{\frac{1}{2}}$. Then from Lemma 4 and (83),

$$\nabla h_+(\boldsymbol{u}^y) = \nabla D_+(\boldsymbol{\lambda}^y)\boldsymbol{B}^{\frac{1}{2}}, \forall \boldsymbol{u}^y \in \mathbb{R}^m_+ \tag{86}$$

where $\nabla h_+(\boldsymbol{u})$ denotes the gradient of $h(\boldsymbol{u})$ for $\boldsymbol{u} \in \mathbb{R}^m_{++}$ and the right hand gradient of $h(\boldsymbol{u})$ at boundary points of $\mathbb{R}^m_+$.

Substituting $\boldsymbol{\lambda}^{y+1} = \boldsymbol{u}^{y+1}\boldsymbol{B}^{\frac{1}{2}}$, $\boldsymbol{\lambda}^y = \boldsymbol{u}^y\boldsymbol{B}^{\frac{1}{2}}$, and $\nabla h_+(\boldsymbol{u}^y) = \nabla D_+(\boldsymbol{\lambda}^y)\boldsymbol{B}^{\frac{1}{2}}$ into (85), the latter relation can equivalently be expressed as:

$$\begin{aligned}\boldsymbol{u}^{y+1} &= \big[\Big(\boldsymbol{u}^y - \nabla h_+(\boldsymbol{u}^y)\Big)\boldsymbol{B}^{\frac{1}{2}}\big]^+\boldsymbol{B}^{-\frac{1}{2}} \\ &= \big[\boldsymbol{u}^y - \nabla h_+(\boldsymbol{u}^y)\big]^+\end{aligned} \tag{87}$$

The second equality in (87) holds since $\boldsymbol{B}$ is a positive definite diagonal matrix.

Lemma 5, together with the diagonal positive-definite form of $\boldsymbol{B}$, then implies that:

$$\begin{aligned}h(\boldsymbol{u}^{y+1}) \le h(\boldsymbol{u}^y) &+ (\boldsymbol{u}^{y+1} - \boldsymbol{u}^y)\nabla h_+(\boldsymbol{u}^y)^T \\ &+ \frac{1}{2}||\boldsymbol{u}^{y+1} - \boldsymbol{u}^y||^2_{\boldsymbol{JB}}\end{aligned} \tag{88}$$

where $||\cdot||^2_{\boldsymbol{JB}} = (\cdot)\boldsymbol{JB}(\cdot)^T$. From condition [P4.D], it follows that

$$||\boldsymbol{u}^{y+1} - \boldsymbol{u}^y||^2_{\boldsymbol{JB}} \le ||\boldsymbol{u}^{y+1} - \boldsymbol{u}^y||^2 \tag{89}$$

Combining (88) and (89), it follows that:

$$\begin{aligned}h(\boldsymbol{u}^{y+1}) \le h(\boldsymbol{u}^y) &+ (\boldsymbol{u}^{y+1} - \boldsymbol{u}^y)\nabla h_+(\boldsymbol{u}^y)^T \\ &+ \frac{1}{2}||\boldsymbol{u}^{y+1} - \boldsymbol{u}^y||^2\end{aligned} \tag{90}$$

By Lemma 1 and [P4.A], the dual function $D(\boldsymbol{\lambda})$ is convex over $\boldsymbol{\lambda} \in \mathbb{R}^m_+$; and $\boldsymbol{\lambda} = \boldsymbol{u}\boldsymbol{B}^{\frac{1}{2}}$ is a linear function of $\boldsymbol{u}$. Consequently, $h(\boldsymbol{u})$ is convex over $\boldsymbol{u} \in \mathbb{R}^m_+$. Making use of a well-known theorem for differentiable convex functions (see [44, Sec. 3.1.3]), it follows that:

$$h(\boldsymbol{u}^y) \le h(\boldsymbol{u}) + (\boldsymbol{u}^y - \boldsymbol{u})\nabla h_+(\boldsymbol{u}^y)^T, \forall \boldsymbol{u} \in \mathbb{R}^m_+ \tag{91}$$

Combining (90) and (91), it follows:

$$\begin{aligned}h(\boldsymbol{u}^{y+1}) \le h(\boldsymbol{u}) &+ (\boldsymbol{u}^{y+1} - \boldsymbol{u})\nabla h_+(\boldsymbol{u}^y)^T \\ &+ \frac{1}{2}||\boldsymbol{u}^{y+1} - \boldsymbol{u}^y||^2, \forall \boldsymbol{u} \in \mathbb{R}^m_+\end{aligned} \tag{92}$$

From (87), we have:

$$\boldsymbol{u}^{y+1} = \arg\min_{\boldsymbol{u}\in\mathbb{R}^m_+} q(u) \tag{93}$$

where

$$q(u) = \frac{1}{2}\big|\big|\boldsymbol{u} - \big[\boldsymbol{u}^y - \nabla h_+(\boldsymbol{u}^y)\big]\big|\big|^2 \tag{94}$$

Making use of Lemma 3 applied to the convex function $q(u)$ in (94), for $\forall \boldsymbol{u} \in \mathbb{R}^m_+$ we have:

$$0 \le (\boldsymbol{u} - \boldsymbol{u}^{y+1})[\boldsymbol{u}^{y+1} - \boldsymbol{u}^y + \nabla h_+(\boldsymbol{u}^y)]^T \tag{95}$$

It follows that:

$$(\boldsymbol{u}^{y+1} - \boldsymbol{u})\nabla h_+(\boldsymbol{u}^y)^T \le (\boldsymbol{u}^{y+1} - \boldsymbol{u}^y)(\boldsymbol{u} - \boldsymbol{u}^{y+1})^T \tag{96}$$

Combining (92) and (96):

$$\begin{aligned}h(\boldsymbol{u}^{y+1}) \;\le\; h(\boldsymbol{u}) &+ (\boldsymbol{u}^{y+1} - \boldsymbol{u}^y)(\boldsymbol{u} - \boldsymbol{u}^{y+1})^T \\ &+ \frac{1}{2}||\boldsymbol{u}^{y+1} - \boldsymbol{u}^y||^2, \forall \boldsymbol{u} \in \mathbb{R}^m_+\end{aligned} \tag{97}$$

Setting $\boldsymbol{u} = \boldsymbol{u}^y$ in (97),

$$h(\boldsymbol{u}^{y+1}) \le h(\boldsymbol{u}^y) - \frac{1}{2}||\boldsymbol{u}^{y+1} - \boldsymbol{u}^y||^2 \le h(\boldsymbol{u}^y) \tag{98}$$

Thus, $h(\boldsymbol{u}^y)$ is non-increasing in the iteration time y.

Also, setting $\boldsymbol{u}$ equal to $\boldsymbol{u}^* = \boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\lambda}^*$ in (97), and manipulating and rearranging terms, one obtains:

$$\begin{aligned}h(\boldsymbol{u}^{y+1}) &\le h(\boldsymbol{u}^*) + (\boldsymbol{u}^{y+1} - \boldsymbol{u}^y)(\boldsymbol{u}^* - \boldsymbol{u}^{y+1})^T \\ &\quad + \frac{1}{2}||\boldsymbol{u}^{y+1} - \boldsymbol{u}^y||^2 \\ &= h(\boldsymbol{u}^*) + \frac{1}{2}\big[||\boldsymbol{u}^y - \boldsymbol{u}^*||^2 - ||\boldsymbol{u}^{y+1} - \boldsymbol{u}^*||^2\big]\end{aligned} \tag{99}$$

Thus,

$$h(\boldsymbol{u}^{y+1}) - h(\boldsymbol{u}^*) \le \frac{1}{2}\big[||\boldsymbol{u}^y - \boldsymbol{u}^*||^2 - ||\boldsymbol{u}^{y+1} - \boldsymbol{u}^*||^2\big] \tag{100}$$

Summing each side of (100) over $y = 0, 1, 2, ..., k-1$,

$$\begin{aligned}\sum_{y=0}^{k-1}\big[h(\boldsymbol{u}^{y+1}) - h(\boldsymbol{u}^*)\big] &\le \frac{1}{2}\big[||\boldsymbol{u}^0 - \boldsymbol{u}^*||^2 - ||\boldsymbol{u}^k - \boldsymbol{u}^*||^2\big] \\ &\le \frac{1}{2}||\boldsymbol{u}^0 - \boldsymbol{u}^*||^2\end{aligned} \tag{101}$$

From (98), we have:

$$k[h(\boldsymbol{u}^k) - h(\boldsymbol{u}^*)] \le \sum_{y=0}^{k-1}\big[h(\boldsymbol{u}^{y+1}) - h(\boldsymbol{u}^*)\big] \tag{102}$$

Hence:

$$h(\boldsymbol{u}^k) - h(\boldsymbol{u}^*) \le \frac{1}{2k}||\boldsymbol{u}^0 - \boldsymbol{u}^*||^2 \tag{103}$$

It then follows from the definition of $h(\boldsymbol{u})$, $\boldsymbol{u}^k$, and $\boldsymbol{u}^*$ that:

$$D(\boldsymbol{\lambda}^k) - D(\boldsymbol{\lambda}^*) \le \frac{1}{2k}||\boldsymbol{\lambda}^0 - \boldsymbol{\lambda}^*||^2_{\boldsymbol{B}^{-1}} \tag{104}$$

However, by Prop. 1 and assumptions [P4.A] and [P4.B], λ^* is an optimal solution for the dual minimization problem (32) with domain $\mathbb{M} = \mathbb{R}^m_+$. Hence:

$$0 \le [D(\boldsymbol{\lambda}^k) - D(\boldsymbol{\lambda}^*)] = [h(\boldsymbol{u}^k) - h(\boldsymbol{u}^*)] \tag{105}$$

From (103), (104), and (105),

$$D(\boldsymbol{\lambda}^k) \to D(\boldsymbol{\lambda}^*) \;\text{ as }\; k \to \infty \tag{106}$$

$$h(\boldsymbol{u}^k) \to h(\boldsymbol{u}^*) \;\text{ as }\; k \to \infty \tag{107}$$

This implies the dual variable $\boldsymbol{\lambda}^k$ determined by the dual decomposition algorithm DDA at the kth iteration time must

converge to the solution set $\boldsymbol{\Lambda}^*$ for the dual problem (32) as $k \to \infty$, where $\boldsymbol{\Lambda}^*$ includes $\boldsymbol{\lambda}^*$ by [P4.B] and Prop. 1.

Moreover, from the Cauchy Convergence Criterion, the convergent scalar sequence $(h(\boldsymbol{u}^k))$ must satisfy the Cauchy condition

$$|h(\boldsymbol{u}^{k+1}) - h(\boldsymbol{u}^k)| \to 0 \quad \text{as} \quad k \to \infty \tag{108}$$

Thus, it follows from (98) and (108) that

$$\frac{1}{2}||\boldsymbol{u}^{k+1} - \boldsymbol{u}^k||^2 \leq |h(\boldsymbol{u}^{k+1}) - h(\boldsymbol{u}^k)| \to 0 \tag{109}$$

as the iteration time $k \to \infty$. Consequently, the dual variable $\boldsymbol{\lambda}^k = \boldsymbol{u}^k \boldsymbol{B}^{\frac{1}{2}}$ must converge to a single point $\boldsymbol{\lambda}' = \boldsymbol{u}' \boldsymbol{B}^{\frac{1}{2}}$ within the solution set $\boldsymbol{\Lambda}^*$; it cannot cycle among the points in $\boldsymbol{\Lambda}^*$.

By [P3.B], an implication of assumption [P4.A], step **S3** of the dual decomposition algorithm DDA determines the following unique solution $\boldsymbol{x}^k$ for every iteration time k:

$$\boldsymbol{x}^k = \underset{\boldsymbol{x} \in \mathcal{X}}{\operatorname{argmax}}\, L(\boldsymbol{x}, \boldsymbol{\lambda}^k) \tag{110}$$

The sequence $(\boldsymbol{x}^k, \boldsymbol{\lambda}^k)$ determined by DDA thus converges to a limit point $(\boldsymbol{x}', \boldsymbol{\lambda}')$ as the iteration time $k \to \infty$. From Prop/ 3, it follows that $(\boldsymbol{x}', \boldsymbol{\lambda}')$ is a saddle point for the Lagrangian Function. Q.E.D.

APPENDIX J

This appendix provides a proof of Prop. 5 in Section VIII-C.[23]

Definition: Subgradient and Subdifferential. For any convex function $f{:}\mathbb{R}^d \to \mathbb{R}$, a vector $\boldsymbol{\nu} \in \mathbb{R}^d$ is defined to be a *subgradient* of f at a point $\boldsymbol{x} \in \mathbb{R}^d$ if:

$$f(\boldsymbol{z}) \geq f(\boldsymbol{x}) + (\boldsymbol{z} - \boldsymbol{x})^T \boldsymbol{\nu} \tag{111}$$

The set of all subgradients of a convex function f at a point $\boldsymbol{x} \in \mathbb{R}^d$ is called the *subdifferential* of f at $\boldsymbol{x}$, and is denoted by $\partial f(\boldsymbol{x})$.

Proposition 5: *Suppose the Primal Problem (23) satisfies condition [P3.A] in Prop. 3 plus the following three additional conditions:*

- **[P5.A]** *$\mathcal{X}$ is a non-empty compact convex subset of $\mathbb{R}^d$.*
- **[P5.B]** *The objective function $F{:}\mathbb{R}^d \to \mathbb{R}$ restricted to $\mathcal{X} \subseteq \mathbb{R}^d$ has the quadratic form*

$$F(\boldsymbol{x}) \;=\; \frac{1}{2}\,\boldsymbol{x}^T \boldsymbol{W} \boldsymbol{x} + \boldsymbol{\rho}^T \boldsymbol{x} + \sigma \tag{112}$$

  where $\boldsymbol{W}$ is any real symmetric negative-definite $d \times d$ matrix, $\boldsymbol{\rho}$ is any real $d \times 1$ column vector, and σ is any real positive scalar.
- **[P5.C]** *The constraint function $g{:}\mathbb{R}^d \to \mathbb{R}^m$ restricted to $\mathcal{X} \subseteq \mathbb{R}^d$ has the linear affine form*

$$g(\boldsymbol{x}) \;=\; \boldsymbol{C}\boldsymbol{x} + \boldsymbol{b} \tag{113}$$

  where $\boldsymbol{C}$ is any real $m \times d$ matrix, and $\boldsymbol{b}$ is any real $m \times 1$ column vector.

Then the Extended Lipschitz Continuity Condition [P4.C] in Prop. 4 holds for $\boldsymbol{J} = \boldsymbol{C}\boldsymbol{H}^{-1}\boldsymbol{C}^T$, where $\boldsymbol{H} = -\boldsymbol{W}$.

[23]The proof of Prop. 5 is based on Giselsson [45, Lemma 12] and [45, Thm. 13].

Proof of Proposition 5: Given assumption [P5.C], the explicit constraints $g(\boldsymbol{x}) \leq \boldsymbol{c}$ for the Primal Problem (23) can equivalently be expressed as

$$\boldsymbol{C}\boldsymbol{x} \leq [\boldsymbol{c} - \boldsymbol{b}] \equiv \boldsymbol{e} \tag{114}$$

where $\boldsymbol{e}$ is a real $m \times 1$ column vector. The Lagrangian Function (25) can then equivalently be expressed as $L(\boldsymbol{x}, \boldsymbol{\lambda}) = F(\boldsymbol{x}) + \boldsymbol{\lambda}[\boldsymbol{e} - \boldsymbol{C}\boldsymbol{x}]$.

Let the extended indicator function $I_{\mathcal{X}}{:}\mathbb{R}^d \to \{0, \infty\}$ be defined as follows:

$$I_{\mathcal{X}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x} \in \mathcal{X} \\ +\infty & \text{otherwise} \end{cases} \tag{115}$$

By assumption [P5.A], $\mathcal{X}$ is a non-empty convex subset of $\mathbb{R}^d$. It is then straightforward to verify that $\boldsymbol{I}_{\mathcal{X}}(\boldsymbol{x})$ is a convex function over $\mathbb{R}^d$; and it follows from [42, Prop. B.24] that the subdifferential $\partial \boldsymbol{I}_{\mathcal{X}}(\boldsymbol{x})$ of $\boldsymbol{I}_{\mathcal{X}}(\boldsymbol{x})$ is a nonempty, compact, and convex subset of $\mathbb{R}^d$.

For each $\boldsymbol{\lambda} \in \mathbb{R}^m_+$, assumptions [P5.A]-[P5.C] ensure that the function $Q_{\boldsymbol{\lambda}}(\boldsymbol{x}) = L(\boldsymbol{x}, \boldsymbol{\lambda})$ is a real-valued twice continuously differentiable function of $\boldsymbol{x} \in \mathcal{X}$ whose Hessian matrix is given by the real symmetric negative-definite matrix $\boldsymbol{W}$. Thus, $Q_{\boldsymbol{\lambda}}(\boldsymbol{x})$ is a continuous strictly-concave function over the non-empty compact convex set $\mathcal{X}$, which implies that condition [P3.B] in Prop. 3 holds. It then follows from [P3.A] and [P3.B] that Lemma 4 holds.

Using (115), the function $\boldsymbol{x}(\boldsymbol{\lambda})$ defined in [P3.B] can be represented as follows: for each $\boldsymbol{\lambda} \in \mathbb{R}^m_+$,

$$\boldsymbol{x}(\boldsymbol{\lambda}) = \underset{\boldsymbol{x} \in \mathbb{R}^d}{\arg\min} \big[-F(\boldsymbol{x}) + I_{\mathcal{X}}(\boldsymbol{x}) + \boldsymbol{\lambda}[\boldsymbol{C}\boldsymbol{x} - \boldsymbol{e}] \big] \tag{116}$$

Given assumption [P5.B], the gradient $\nabla \boldsymbol{F}_+(\boldsymbol{x})$ of $F(\boldsymbol{x})$ takes the form $\boldsymbol{W}\boldsymbol{x} + \boldsymbol{\rho}$ for each $\boldsymbol{x} \in \mathcal{X}$.[24] Thus, the first-order necessary conditions for $\boldsymbol{x}(\boldsymbol{\lambda})$ to satisfy (116) at any specific points $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ in $\mathbb{R}^m_+$ are:

$$0 \in -\nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_1)) + \partial I_{\mathcal{X}}(\boldsymbol{x}(\boldsymbol{\lambda}_1)) + \boldsymbol{C}^T \boldsymbol{\lambda}_1^T \tag{117a}$$

$$0 \in -\nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_2)) + \partial I_{\mathcal{X}}(\boldsymbol{x}(\boldsymbol{\lambda}_2)) + \boldsymbol{C}^T \boldsymbol{\lambda}_2^T \tag{117b}$$

Let $\xi(\boldsymbol{x}(\boldsymbol{\lambda}_1)) \in \partial I_{\mathcal{X}}(\boldsymbol{x}(\boldsymbol{\lambda}_1))$ and $\xi(\boldsymbol{x}(\boldsymbol{\lambda}_2)) \in \partial I_{\mathcal{X}}(\boldsymbol{x}(\boldsymbol{\lambda}_2))$ denote the particular subgradients that result in equalities in (117a) and (117b), respectively. The existence of these subgradients is assured by [P3.B]. This gives:

$$0 = -\nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_1)) + \xi(\boldsymbol{x}(\boldsymbol{\lambda}_1)) + \boldsymbol{C}^T \boldsymbol{\lambda}_1^T \tag{118a}$$

$$0 = -\nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_2)) + \xi(\boldsymbol{x}(\boldsymbol{\lambda}_2)) + \boldsymbol{C}^T \boldsymbol{\lambda}_2^T \tag{118b}$$

Taking the scalar product of (118a) with $\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)$, and the scalar product of (118b) with $\boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)$:

$$\begin{aligned} \big\langle \nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_1)), \boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1) \big\rangle - \big\langle \boldsymbol{C}^T \boldsymbol{\lambda}_1^T, \boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1) \big\rangle \\ = \big\langle \xi(\boldsymbol{x}(\boldsymbol{\lambda}_1)), \boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1) \big\rangle \end{aligned} \tag{119a}$$

$$\begin{aligned} \big\langle \nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_2)), \boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2) \big\rangle - \big\langle \boldsymbol{C}^T \boldsymbol{\lambda}_2^T, \boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2) \big\rangle \\ = \big\langle \xi(\boldsymbol{x}(\boldsymbol{\lambda}_2)), \boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2) \big\rangle \end{aligned} \tag{119b}$$

[24]In this study the derivative of a scalar with respect to a column (row) vector is represented as a column (row) vector.

where $\langle\ ,\ \rangle$ denotes vector inner product. Summing the two equalities (119a) and (119b):

$$\begin{aligned}&\big\langle \nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_2) - \nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_1), \boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)\big\rangle \\ &\quad + \big\langle \boldsymbol{C}^T(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)^T, \boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)\big\rangle \\ &\qquad = \big\langle \xi(\boldsymbol{x}(\boldsymbol{\lambda}_2)) - \xi(\boldsymbol{x}(\boldsymbol{\lambda}_1)), \boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)\big\rangle\end{aligned} \tag{120}$$

Since $I_{\mathcal{X}}(\boldsymbol{x})$ is a convex function, it follows from (111 that

$$I_{\mathcal{X}}(\boldsymbol{x}(\boldsymbol{\lambda}_1)) \geq I_{\mathcal{X}}(\boldsymbol{x}(\boldsymbol{\lambda}_2)) + \xi(\boldsymbol{x}(\boldsymbol{\lambda}_2))^T\big[\boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)\big] \tag{121a}$$

$$I_{\mathcal{X}}(\boldsymbol{x}(\boldsymbol{\lambda}_2)) \geq I_{\mathcal{X}}(\boldsymbol{x}(\boldsymbol{\lambda}_1)) + \xi(\boldsymbol{x}(\boldsymbol{\lambda}_1))^T\big[\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)\big] \tag{121b}$$

Summing the two inequalities (121a) and (121b), we have:

$$\big\langle \xi(\boldsymbol{x}(\boldsymbol{\lambda}_2)) - \xi(\boldsymbol{x}(\boldsymbol{\lambda}_1)), \boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)\big\rangle \leq 0 \tag{122}$$

Combining (120) and (122):

$$\begin{aligned}&\big\langle \nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_2) - \nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_1), \boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)\big\rangle \\ &\leq \big\langle \boldsymbol{C}^T(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)^T, \boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)\big\rangle\end{aligned} \tag{123}$$

Let $\boldsymbol{H}$ denote the real symmetric positive-definite $d \times d$ matrix given by $\boldsymbol{H} = -\boldsymbol{W}$. It follows from [P5.B] that

$$\nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_2) - \nabla \boldsymbol{F}_+(\boldsymbol{x}(\boldsymbol{\lambda}_1) = \boldsymbol{H}\big[\boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)\big] \tag{124}$$

Substituting (124) into (123), one obtains

$$\begin{aligned}||\,[\boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)]^T||_{\boldsymbol{H}}^2 &\leq \big\langle \boldsymbol{C}^T(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)^T, \boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)\big\rangle \\ &= \big\langle \boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{C}^T[\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2]^T, \boldsymbol{H}^{\frac{1}{2}}[\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1]\big\rangle \\ &\leq ||\,\boldsymbol{H}^{-\frac{1}{2}}\boldsymbol{C}^T[\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2]^T\,|| \cdot ||\,\boldsymbol{H}^{\frac{1}{2}}[\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)]^T\,|| \\ &= ||(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\boldsymbol{C}\boldsymbol{H}^{-\frac{1}{2}}|| \cdot ||\,[\boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)]^T||_{\boldsymbol{H}}\end{aligned} \tag{125}$$

where $||u|| = \sqrt{<u,u>}$, $||\cdot||_{\boldsymbol{H}} = \sqrt{(\cdot)\boldsymbol{H}(\cdot)^T}$, $||\cdot||_{\boldsymbol{H}}^2 = (\cdot)\boldsymbol{H}(\cdot)^T$, and the inequality holds by the Cauchy-Schwartz Inequality. Thus, we can conclude from (125) that:

$$\begin{aligned}||[\boldsymbol{x}(\boldsymbol{\lambda}_1) - \boldsymbol{x}(\boldsymbol{\lambda}_2)]^T||_{\boldsymbol{H}} &\leq ||(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\boldsymbol{C}\boldsymbol{H}^{-\frac{1}{2}}|| \\ &= ||\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2||_{\boldsymbol{C}\boldsymbol{H}^{-1}\boldsymbol{C}^T}\end{aligned} \tag{126}$$

Finally, using Lemma 4 together with (126,

$$\begin{aligned}&\big\langle \nabla D_+(\boldsymbol{\lambda}_1) - \nabla D_+(\boldsymbol{\lambda}_2), \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2\big\rangle \\ &= \big\langle [\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)]^T\boldsymbol{C}^T, \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2\big\rangle \\ &= \big\langle [\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)]^T, (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\boldsymbol{C}\big\rangle \\ &= \big\langle [\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)]^T, (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\boldsymbol{C}\boldsymbol{H}^{-1}\big\rangle_{\boldsymbol{H}} \\ &\leq ||[\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)]^T||_{\boldsymbol{H}}||(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\boldsymbol{C}\boldsymbol{H}^{-1}||_{\boldsymbol{H}} \\ &= ||[\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)]^T||_{\boldsymbol{H}}||\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2||_{\boldsymbol{C}\boldsymbol{H}^{-1}\boldsymbol{C}^T} \\ &\leq ||\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2||_{\boldsymbol{C}\boldsymbol{H}^{-1}\boldsymbol{C}^T}^2\end{aligned} \tag{127}$$

where:

$$\begin{aligned}&\big\langle [\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)]^T, (\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\boldsymbol{C}\boldsymbol{H}^{-1}\big\rangle_{\boldsymbol{H}} \\ &= [\boldsymbol{x}(\boldsymbol{\lambda}_2) - \boldsymbol{x}(\boldsymbol{\lambda}_1)]^T\boldsymbol{H}\big[(\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2)\boldsymbol{C}\boldsymbol{H}^{-1}\big]^T\end{aligned}$$

In (127) the first inequality follows from the Cauchy-Schwartz Inequality and the second inequality follows from (126). Relation (127) implies that condition [P4.C] holds with $\boldsymbol{J} = \boldsymbol{C}\boldsymbol{H}^{-1}\boldsymbol{C}^T$. Q.E.D.

Comment Regarding Proposition 5: Since the matrix $\boldsymbol{H}$ depends on private household information, so does $\boldsymbol{J} = \boldsymbol{C}\boldsymbol{H}^{-1}\boldsymbol{C}^T$. Thus, the DSO cannot directly guarantee condition [P4.D] in Prop. 4 is satisfied by setting the elements of the step-size matrix $\boldsymbol{B}$ for the dual decomposition algorithm DDA in a manner that ensures $[\boldsymbol{I} - \boldsymbol{J}\boldsymbol{B}]$ is positive semi-definite. However, a DSO can learn from experience how to adjust the elements of $\boldsymbol{B}$ to ensure [P4.D] holds, hence to ensure convergence of DDA.

APPENDIX K

This appendix analyzes the relationship between the conditions assumed in Prop. 3, Prop. 4, and Prop. 5 in Section VIII-C and the structural form of the Primal Problem (23) assumed for the case study presented in Section IX. Specifically, it is shown that the case study satisfies all of the conditions assumed for these three propositions as long as the positive step-sizes $\{\beta_1, \beta_2, \beta_3\}$ appearing in step **S4** of the dual decomposition algorithm DDA are set to ensure the positive semi-definiteness of the matrix $I - JB$ appearing in the Prop. 4 condition [P4.D].

From (13), each household constraint set $\mathcal{X}_\psi(\mathcal{K})$ is compact and convex; and the number NH of households is assumed to be finite. Thus, the set product $\mathcal{X} = \mathcal{X}(\mathcal{K}) = \prod_{\psi\in\Psi}\mathcal{X}_\psi(\mathcal{K})$ is also compact and convex. In addition, it follows from (23) that $F(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are continuous functions of $\boldsymbol{x} \in \mathcal{X}$. Conditions [P3.A] and [P5.A] thus hold for the case study.

For concreteness, the '$\pm$' sign in the household constraints (13a) and (13b) for the case study is assumed to be a minus sign '-'. These constraints can then be expressed in the following compact $NK \times 1$ column vector form:

$$\boldsymbol{T}_\psi(\mathcal{P}_\psi(\mathcal{K})) = \boldsymbol{a}_\psi - \boldsymbol{G}_\psi\mathcal{P}_\psi(\mathcal{K}) \tag{128}$$

with $NK \times 1$ column vectors

$$\begin{aligned}\boldsymbol{T}_\psi(\mathcal{P}_\psi(\mathcal{K})) &= [T_\psi(p_\psi(1),1), ..., T_\psi(p_\psi(NK),NK)]^T \\ \boldsymbol{a}_\psi &= \boldsymbol{Y}_\psi\boldsymbol{T}_\psi(0) + \boldsymbol{E}_\psi\boldsymbol{T}_o \\ \boldsymbol{T}_\psi(0) &= [T_\psi(0), T_\psi(0), ..., T_\psi(0)]^T \\ \boldsymbol{T}_o &= [T_o(1), T_o(2), ..., T_o(NK)]^T\end{aligned}$$

and with square $NK \times NK$ matrices:

$$\boldsymbol{Y}_\psi = \begin{bmatrix}\alpha_\psi^H & & & \\ & (\alpha_\psi^H)^2 & & \\ & & \ddots & \\ & & & (\alpha_\psi^H)^{NK}\end{bmatrix}$$

$$\boldsymbol{E}_\psi = \begin{bmatrix}1-\alpha_\psi^H & & & \\ \alpha_\psi^H(1-\alpha_\psi^H) & 1-\alpha_\psi^H & & \\ \vdots & \ddots & \ddots & \\ (\alpha_\psi^H)^{NK-1}(1-\alpha_\psi^H) & \cdots & \ddots & 1-\alpha_\psi^H\end{bmatrix}$$

$$\boldsymbol{G}_\psi = S_{base}\Delta t\begin{bmatrix}\alpha_\psi^P & & & \\ \alpha_\psi^H\alpha_\psi^P & \alpha_\psi^P & & \\ \vdots & \ddots & \ddots & \\ (\alpha_\psi^H)^{NK-1}\alpha_\psi^P & \cdots & \alpha_\psi^H\alpha_\psi^P & \alpha_\psi^P\end{bmatrix}$$

Also, the expression $U_\psi(\mathcal{P}_\psi(\mathcal{K}))$ defined in (11) can be written in the compact form:

$$\begin{aligned}&U_\psi(\mathcal{P}_\psi(\mathcal{K})) = u_\psi^{max}\cdot NK \\ &- c_\psi[\boldsymbol{T}_\psi(\mathcal{P}(\mathcal{K})) - \boldsymbol{T}\boldsymbol{B}_\psi]^T[\boldsymbol{T}_\psi(\mathcal{P}(\mathcal{K})) - \boldsymbol{T}\boldsymbol{B}_\psi]\end{aligned} \tag{129}$$

with $NK \times 1$ column vector $\boldsymbol{TB}_\psi = [TB_\psi, ..., TB_\psi]^T$. Substituting $\boldsymbol{x}_\psi = \mathcal{P}_\psi(\mathcal{K})$ and (128) into (129):

$$U_\psi(\boldsymbol{x}_\psi) = u_\psi^{max} \cdot NK \\ - c_\psi[-\boldsymbol{G}_\psi \boldsymbol{x}_\psi + \boldsymbol{a}_\psi - \boldsymbol{TB}_\psi]^T \cdot [-\boldsymbol{G}_\psi \boldsymbol{x}_\psi + \boldsymbol{a}_\psi - \boldsymbol{TB}_\psi] \tag{130}$$

Using the definition of the objective function $F_\psi(\boldsymbol{x}_\psi)$ in the Primal Problem (23), the Hessian matrix of $F_\psi(\boldsymbol{x}_\psi)$ for the case study is given by:

$$\nabla^2 F_\psi(\boldsymbol{x}_\psi) = \nabla^2 U_\psi(\boldsymbol{x}_\psi) = -2c_\psi \boldsymbol{G}_\psi^T \boldsymbol{G}_\psi \tag{131}$$

Since $\alpha_\psi^P > 0$, the matrix $\boldsymbol{G}_\psi$ has a positive determinant, implying $\boldsymbol{G}_\psi$ has full rank. Since $c_\psi > 0$, it follows that $\nabla^2 F_\psi(\boldsymbol{x}_\psi)$ in (131) is a symmetric negative-definite matrix, hence $F_\psi(\boldsymbol{x})$ is strictly concave over $\boldsymbol{x} \in \mathcal{X}$. Defining $F(\boldsymbol{x}) = \sum_{\psi \in \Psi} F_\psi(\boldsymbol{x}_\psi)$, it follows that $\nabla^2 F(\boldsymbol{x})$ is also a symmetric negative-definite matrix, hence $F(\boldsymbol{x})$ is strictly concave over $\boldsymbol{x} \in \mathcal{X}$. Thus, condition [P5.B] holds for the case study. Also, using the definition of the constraint function $g(\boldsymbol{x})$ in the Primal Problem (23), it is also seen that condition [P5.C] holds for the case study. Condition [P3.B] then follows from conditions [P5.A]-[P5.C].

Moreover, it then follows directly from Prop. 5 that condition [P4.C] is satisfied for the case study. Condition [P4.D] in Prop. 4 will then be satisfied for the case study as long as the positive step-sizes $\{\beta_1, \beta_2, \beta_3\}$ in step **S4** of the dual decomposition algorithm DDA are selected in a manner to ensure the positive semi-definiteness of the matrix $I - JB$ appearing in condition [P4.D].

By construction, the case study satisfies the following feasibility condition:

$$\mathcal{Z} \equiv \{\boldsymbol{x} \in \mathcal{X} \mid \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{c}\} \neq \emptyset \tag{132}$$

Given [P5.A]-[P5.C] and condition (132), the Primal Problem (23) for the case study reduces to the maximization of the continuous strictly-concave objective function $F(\boldsymbol{x})$ over the non-empty, compact, and convex subset $\mathcal{Z} \subseteq \mathbb{R}^d$. Thus, this Primal Problem has a unique finite solution $\boldsymbol{x}^* \in \mathcal{Z}$.

Moreover, by [P5.C], $\boldsymbol{g}(\boldsymbol{x})$ is linear affine. It follows from [46, Prop.5.3.1] that the optimal solution set for the Dual Problem (32) for the case study contains at least one point $\boldsymbol{\lambda}^*$ in $\mathbb{R}_+^m$; and strong duality holds for $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$. Thus, Prop. 1 implies that $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ is a saddle point solution for the case study; i.e., condition [P4.B] holds.

Finally, Prop. 3 and Prop. 4 imply that [P3.C] must hold for the case study, given that [P4.B] holds for the case study.

REFERENCES

[1] GridWise Architectural Council, *GridWise Transactive Energy Framework Version 1.0*, January, 2015.

[2] H.T. Nguyen, S. Battula, R.R. Takkala, Z. Wang, and L. Tesfatsion, "An integrated transmission and distribution test system for evaluation of transactive energy designs," *Appl. Energy*, vol. 240, pp. 666-679, 2019.

[3] K. Kok, *The PowerMatcher: Smart Coordination for the Smart Electricity Grid*, SIKS Dissertation Series No. 2013-17, Dutch Research School for Information and Knowledge Systems, TNO, The Netherlands, 2013. http://dare.ubvu.vu.nl/handle/1871/43567

[4] L. Hurwicz, "The design of mechanisms for resource allocation," *The American Economic Review*, vol. 63, no. 2, pp. 1-30, 1973.

[5] Royal Swedish Academy of Sciences, *Mechanism Design Theory: Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel*, compiled by the Prize Committee, 2007.

[6] K. KoK, and S. Widergren, "A society of devices: Integrating intelligent distributed resources with transactive energy," *IEEE Power and Energy Mag.*, vol. 14, no. 3, pp. 34-45, 2016.

[7] R. Deng, Z. Yang, M.Y. Chow, and J. Chen, "A survey on demand response in smart grids: Mathematical models and approaches," *IEEE Trans. Ind. Informat.*, vol. 11, no. 3, pp. 570-582, June 2015.

[8] E. Koch and M.A. Piette, "Direct versus facility centric load control for automated demand," presented at the Grid-Interop Forum, Denver, CO, USA, 2009.

[9] S. Chen, Q. Chen, and Y. Xu, "Strategic bidding and compensation mechanism for a load aggregator with direct thermostat control capabilities," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 2327-2336, May 2018.

[10] N. O'Connell, P. Pinson, H. Madsen, and M. OMalley, "Benefits and challenges of electrical demand response: A critical review," *Renew. Sust. Energ. Rev.*, vol. 39, pp. 686-699, 2014.

[11] L. Gan and S.H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial network," in *18th Power Systems Computation Conference (PSCC)*, 2014.

[12] E. Celebi and J.D. Fuller, "Time-of-use pricing in electricity markets under different market structures," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1170-1181, Aug. 2012.

[13] K. Herter, "Residential implementation of critical-peak pricing of electricity," *Energy Policy*, vol. 35, pp. 2121-2130, Apr. 2007

[14] A.G. Thomas and L. Tesfatsion, "Braided Cobwebs: Cautionary Tales for Dynamic Pricing in Retail Electric Power Markets," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6870-6882, 2018.

[15] H. Hao, C.D. Corbin, K. Kalsi, and R.G. Pratt, "Transactive control of commercial buildings for demand response," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 774-783, Jan. 2017.

[16] F. Rahimi and A. Ipakchi, "Using a transactive energy framework: providing grid services from smart buildings," *IEEE Electrification Mag.*, vol. 4, no. 4, pp. 23-29, 2016.

[17] S. Chen and C.C. Liu, "From demand response to transactive energy: State of the art," *J. Mod. Power Syst. Clean Energy*, vol. 5, no. 1, pp. 10-19, 2017.

[18] E. Mengelkamp, J. Grttner, K. Rock, S. Kessler, L. Orsini, and C. Weinhardt, "Designing microgrid energy markets: A case study: The Brooklyn microgrid," *Appl. Energy*, vol. 210, pp. 870-880, Jan. 2018.

[19] S. Battula, L. Tesfatsion, and Z. Wang, "A customer-centric approach to bid-based transactive energy system design," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4996-5008, 2020.

[20] N. Atamturk and M. Zafar, "Transactive energy: A surreal vision or a necessary and feasible solution to grid problems," California Public Utilities Commission, Policy Planning Division, San Francisco, CA, USA, Tech. Rep., 2014.

[21] D.J. Hammerstrom, R. Ambrosio, T.A. Carlon *et al.*, "Pacific northwest gridwise test bed demonstration projects," Pacific Northwest Nat. Lab., Richland, WA, USA, Tech. Rep. PNNL-17167, 2007.

[22] S.E. Widergren, K. Subbarao, J.C. Fuller *et al.*, "AEP Ohio grid smart demonstration project real-time pricing demonstration analysis," Pacific Northwest Nat. Lab., Richland, WA, USA, Tech. Rep. PNNL-23192, 2014.

[23] S. Li, W. Zhang, J. Lian, and K. Kalsi, "Market-based coordination of thermostatically controlled loads-Part I: A mechanism design formulation," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1170-1178, Mar. 2016.

[24] Z. Liu, Q. Wu, K. Ma, M. Shahidehpour, Y. Xue, and S. Huang, "Two stage optimal scheduling of electric vehicle charging based on transactive control," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2948-2958, May 2019.

[25] S. Vandael, B. Claessens, M. Hommelberg, T. Holvoet, and G. Deconinck, "A scalable three-step approach for demand side management of plug-in hybrid vehicles," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 720-728, Jun. 2013.

[26] S. Weckx, R. D'Hulst, B. Claessens, and J. Driesensam, "Multiagent charging of electric vehicles respecting distribution transformer loading and voltage limits," *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2857-2867, Nov. 2014.

[27] J. Hu, G. Yang, H.W. Bindner, and Y. Xue, "Application of network-constrained transactive control to electric vehicle charging for secure grid operation," *IEEE Trans. Sustain. Energy*, vol. 8, no. 2, pp. 505-515, Apr. 2017.

[28] FERC, "Participation of Distributed Energy Resource Aggregations in Markets Operated by Regional Transmission Organizations and Independent System Operators,", U.S. Federal Energy Regulatory Commission

(FERC), Docket No. RM18-9-000, Order No. 2222, Final Rule, September 17, 2020.
[29] ERCOT. Nodal Protocols - Section 6: Adjustment Period and Real-time Operations. Electric Reliability Council of Texas; 2020.
[30] S. Battula, L. Tesfatsion, and T.E. McDermott, "An ERCOT test system for market design studies," *Appl. Energy*, vol. 275, October. 2020.
[31] M.E. Baran and F.F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 725-734, Jan. 1989.
[32] M.E. Baran and F.F. Wu, "Network reconfiguration in distribution system for loss reduction and load balancing," *IEEE Trans. Power Del.*, vol. 4, no. 2, pp. 1401-1407, Apr. 1989.
[33] B.A. Robbins and A.D. Domínguez-Garca, "Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2903-2913, Jul. 2016.
[34] Q. Zhang, K. Dehghanpour and Z. Wang, "Distributed CVR in unbalanced distribution networks with PV penetration," *IEEE Trans. Smart Grid.*, vol. 10, no. 5, pp. 5308-5319, Sep. 2019.
[35] H. Zhu and H.J. Liu, "Fast local voltage control under limited reactive power: optimality and stability analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3794-3803, Sep. 2016.
[36] M. Ilic, J.W. Black and J.L. Watz, "Potential benefits of implementing load control," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, NewYork, NY, USA, 2002, pp. 177-182.
[37] Z. Chen, L. Wu and Y. Fu,"Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1822-1831, Dec. 2012.
[38] Y. Hong, J. Lin, C. Wu, and C. Chuang, "Multi-objective air-conditioning control considering fuzzy parameters using immune clonal selection programming," *IEEE Trans. Smart Grid*, vol. 3, pp. 1603-1610, Dec. 2012.
[39] J. Lfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. CACSD Conf.*, Taipei, Taiwan, 2004.
[40] C. Bliek, P. Bonami, and A. Lodi, "Solving mixed-integer quadratic programming problems with IBM-CPLEX: a progress report," in *Proc. 26th RAMP Symposium*, Tokyo, Japan, Oct. 2014.
[41] W.H. Kersting, "Radial distribution test feeders," in *Proc. IEEE Power Eng. Soc. Win. Meeting*, pp. 908-912, 2001.
[42] D. P. Bertsekas, *Nonlinear Programming*, Second Edition, Athena Scientific, Belmont, MA, USA, 1999.
[43] D. P. Bertsekas, *Convex Analysis and Optimization: Chapter 8 Solutions*, Athena Scientific, Belmont, MA, USA, 2003. http://www.athenasc.com
[44] S. Boyd and L. Vandenberghe, *Convex Optimization*, Seventh Printing, Cambridge University Press, Cambridge, UK, 2009.
[45] P. Giselsson, "Improving Fast Dual Ascent for MPC - Part I: The Distributed Case," *arXiv: 1312.3012*, 2013.
[46] D. P. Bertsekas, *Convex Optimization Theory*, Athena Scientific, Belmont, MA, USA, 2009.

Rui Cheng (S'19) is currently a Ph.D. student in the Department of Electrical & Computer Engineering at Iowa State University. He received the B.S. degree in electrical engineering from Hangzhou Dianzi University in 2015 and the M.S. degree in electrical engineering from North China Electric Power University in 2018. From 2018 to 2019, he was an Electrical Engineer with State Grid Corporation of China, Hangzhou, China. His research interests include power distribution systems, transactive energy markets, and applications of optimization and machine learning methods to power systems.

Leigh Tesfatsion (M'05-SM'18) is Research Professor of Economics, Professor Emerita of Economics, and Courtesy Research Professor of Electrical & Computer Engineering, all at Iowa State University. She received the Ph.D. degree in economics from the University of Minnesota, Mpls., in 1975, with a minor in mathematics. Her principal current research areas are electric power market design and the development of Agent-based Computational Economics (ACE) platforms for the performance testing of these designs. She is the recipient of the 2020 David A. Kendrick Distinguished Service Award from the Society for Computational Economics (SCE). She has served as guest editor and associate editor for a number of journals, including the *IEEE Transactions on Power Systems*, the *IEEE Transactions on Evolutionary Computation*, the *Journal of Energy Markets*, the *Journal of Economic Dynamics and Control*, the *Journal of Public Economic Theory*, and *Computational Economics*.

Zhaoyu Wang (S'13-M'15-SM'20) is the Harpole-Pentair Assistant Professor with Iowa State University. He received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiaotong University, and the M.S. and Ph.D. degrees in electrical and computer engineering from Georgia Institute of Technology. His research interests include optimization and data analytics in power distribution systems and microgrids. He is the principal investigator for a multitude of projects focused on these topics and funded by the National Science Foundation, the Department of Energy, National Laboratories, PSERC, and the Iowa Economic Development Authority. He is Chair of the IEEE Power and Energy Society (PES) and PSOPE Award Subcommittee, Co-Vice Chair of the PES Distribution System Operation and Planning Subcommittee, and Vice Chair of the PES Task Force on Advances in Natural Disaster Mitigation Methods. He is an editor of *IEEE Transactions on Power Systems*, *IEEE Transactions on Smart Grid*, *IEEE Open Access Journal of Power and Energy*, *IEEE Power Engineering Letters*, and *IET Smart Grid*. He is also a recipient of the National Science Foundation (NSF) CAREER Award, the IEEE PES Outstanding Young Engineer Award, and the Harpole-Pentair Young Faculty Award Endowment.