On the Optimality of Procrastination Policy for EV charging under Net Energy Metering

Minjae Jeon, Student Member, IEEE, Lang Tong, Fellow, IEEE, Qing Zhao, Fellow, IEEE

Abstract-We consider the problem of behind-the-meter EV charging by a prosumer, co-optimized with rooftop solar, electric battery, and flexible consumptions such as water heaters and HVAC. Under the time-of-use net energy metering tariff with the stochastic solar production and random EV charging demand, a finite-horizon surplus-maximization problem is formulated. We show that a procrastination threshold policy that delays EV charging to the last possible moment is optimal when EV charging is co-optimized with flexible demand, and the policy thresholds can be computed easily offline. When battery storage is part of the co-optimization, it is shown that the net consumption of the prosumer is a two-threshold piecewise linear function of the behind-the-meter renewable generation under the optimal policy, and the procrastination threshold policy remains optimal, although the thresholds cannot be computed easily. We propose a simple myopic solution and demonstrate in simulations that the performance gap between the myopic policy and an oracle upper bound appears to be 0.5-7.5%.

Index Terms—EV charging, distributed energy resources, stochastic dynamic programming, net energy metering

I. INTRODUCTION

We address the problem of co-optimizing behind the meter(BTM) electric vehicle (EV) charging, flexible consumptions, rooftop solar, and energy storage. This work is motivated by the increasing installation of the BTM DER and storage [1] to shift and flatten the aggregated household loads. With EV as a deferrable and interruptible load, co-optimizing BTM resources benefits not only individual prosumers but also the distribution grid operations in reduced power flow [2].

Most households with BTM DER in the U.S. are under some form of the net energy metering (NEM) tariff offered by a regulated utility or consumer choice aggregator, which bills the prosumer for its net energy consumption reading from the revenue meter, as shown in Fig. 1. Under that setting, we assume that an energy management system (EMS) controls flexible household demands, such as water heater, HVAC, EV charging, and BTM storage operations based on available renewable generations.

The co-optimization problem can be formulated as a continuous state and action space Markov decision process (MDP). Unfortunately, the solution to such a stochastic dynamic program (DP) is intractable without exploiting the special properties of the co-optimization. To this end, we focus on the particular structure of the NEM2.0 and beyond [2], where the purchasing (importing) rate is higher than the



Fig. 1: NEM scheme of the household with the BTM storage and EV. The direction of arrow indicates the direction of the power flow.

selling (exporting) rate. A striking property of such NEM tariffs is that it creates a net-zero zone in the household's net consumption. Our work in this paper builds upon these structural properties uncovered recently in [2], [3].

A. Related works

This work is closely related to [3], [4] which studied joint scheduling of EV and flexible loads [4], and joint scheduling of ESS and flexible loads [3]. A considerable amount of research has also been conducted on the joint scheduling of BTM ESS with an EV [5]–[10] and flexible demands [11]–[15], along with DER. However, only a few studies have specifically incorporated EV charging deadlines and the deferrable nature of EV into the co-optimization with storage. Moreover, to the best of our knowledge, no previous literature has explored an online scheduling of EV, storage, and consumption under NEM tariffs. In this article, we examine previous works that have investigated comparable optimization frameworks and techniques.

The co-optimization problem for the EMS with uncertainties (e.g. rooftop solar generation and outside temperature) is modeled as an one-shot optimization problem in [6], [7], [11]. However, this approach requires perfect forecasts of the uncertainties and can not provide an online policy. The authors of [6] aimed to maximize household surplus by co-optimizing flexible loads with BTM DER. In [7], the authors formulated a mixed integer linear programming (MILP) for scheduling of EV, storage, and PV power usage for the customer enrolled in a demand response program. Authors considered NEM 2.0 as the pricing tariff but they did not explicitly consider controlling household devices. In [11], MILP was formulated for the joint scheduling of controllable loads and storage with rooftop solar but EV was not considered.

Another method involves formulating co-optimization problem as an MDP and solving it using stochastic DP, which allows making online decision using the current DER generation [3]–[5], [13], [14], [16]. In [5], the optimal scheduling

Minjae Jeon, Lang Tong, and Qing Zhao({mj444, lt35, qz16}@cornell.edu) are with the School of Electrical and Computer Engineering, Cornell University, USA. This work was supported in part by the National Science Foundation under Grant 2218110 and 1932501.

of EV and battery with rooftop solar to minimize cost and incomplete EV charging demand was studied. However, the authors did not take into account flexible demand and selling excess DER generation to the grid. In [3], [13], the authors focused on finding optimal threshold exploiting threshold structure of the optimal policy for scheduling storage and loads. Although [3], [13] did not consider EV, they are among the closely related works that studied the scheduling of loads and storage to maximize expected household surplus with surplus renewables allowed to be sold to the grid. Additionally, [4], [16] explored optimal threshold policies for the joint scheduling of deferrable and non-deferrable loads. In [14], the authors formulated stochastic control problem into Lyapunov optimization and derived closedform scheduling of storage and loads. However, they did not consider controlling individual loads and deferrable loads such as EV.

Model predictive control (MPC) has been widely studied for joint scheduling under uncertainties as an alternate approach for solving MDP [8], [9], [17]. MPC can be adapted in real-time using the current available information but it's performance highly depends on the forecasting algorithm, and problem size, such as look ahead horizon, and state, action space. In [8], an MPC-based joint scheduling of rooftop solar, battery, and EV that optimizes user comfort and cost was studied. Selling surplus solar generation was allowed but the selling price was the same as the energy purchasing price. In [9], the same optimization problem using MPC-based method was studied but the work included storage battery degradation cost. An MPC-based algorithm was proposed in [17] to schedule EV and flexible loads. However, the storage wasn't scheduled, and selling to the grid was not allowed.

Recently, reinforcement learning (RL) based methods have been applied for home EMS co-optimization [10], [15], [18]. The RL-based approach has the advantage that of not requiring knowledge about uncertainties, but it is not free from the curse of dimensionality for high-dimensional discrete or continuous state/action space. For instance, [10], [18] discretized EV charging and battery operation, and applied Qlearning based method for joint scheduling of EV and loads under random price and PV generation. In [15], the authors used deep deterministic policy gradient (DDPG) algorithm for joint scheduling storage and loads in the continuous state/action space, but the algorithm converges slowly, which restricts the usage of the algorithm in an online setting.

B. Summary of results and contributions

 We demonstrate that the procrastination threshold policy, which involves delaying EV charging to the last possible moment, which is the optimal policy for the EV, controllable loads, and DER co-optimization problem under NEM [4], is the optimal strategy for the co-optimization including myopic battery operation with the additional procrastination thresholds. Co-optimization including myopic storage operation considers battery operations with immediate stage reward and ignore future renewables and storage decisions. Additional procrastination thresholds determine the priority rule of storage operations and EV charging based on the current remaining charging demand. The procrastination thresholds, which characterize the optimal policy, can be computed offline when the BTM DER is modeled as a sequence of independent random variables. Assuming a nonbinding battery SoC limit, the optimal policy for the cooptimization including non-myopic battery operations becomes myopic policy.

- The myopic storage decisions are piecewise linear function of renewable generation which is divided into 5 segments by 4 thresholds. The myopic battery operation stores energy in the battery solely using BTM DER generation, and storage discharges only to curtail the net consumption of the household.
- The empirical studies using the real-world data show that the procrastination threshold policy with myopic storage operations outperforms other non cooptimization policies and are in within 0.5-7% performance gap to an oracle policy.

The notations used in the paper are standard. Vectors are in boldface, with (x_1, \ldots, x_N) a column vector. We use 1 for a column of ones with appropriate size. $\mathbb{1}_A$ is an indicator function that maps to 1 if A is true, otherwise to zero.

The proofs of theorems and propositions are omitted due to space limitations, and they can be found at [19].

II. PROBLEM FORMULATION

We consider a sequential scheduling of EV charging (v_t) , controllable household loads (\mathbf{d}_t) , and storage operation (e_t) over a discrete-time and finite horizon of length T with intervals indexed by $\mathcal{T} = \{0, \ldots, T-1\}$.

A. EV charging, consumption, and storage model

a) BTM DER (r_t) : We model the BTM DER generation during interval t as an exogenous variable, represented by a sequence of independent random variables with a timedependent distribution $f_t(\cdot)$. Additionally, we assume that the renewable generation for that particular interval is already known at the start of the interval.

$$r_t \sim f_t(\cdot), \quad \forall t \in \mathcal{T}$$
 (1)

b) Remaining charging demand (y_t) and constraints: At the start of every interval, the outstanding charging demand (y_t) is measured. A charger with a maximum capacity \bar{v} and charging efficiency η provides v_t in interval t. We assume that the EV does not discharge it's battery and the total energy supplied to the EV will not exceed the amount originally requested at t = 0 (s_{req}).

$$y_{t+1} = y_t - \eta v_t, \quad \forall t \in \mathcal{T}.$$
 (2)

$$y_t \in [0, s_{\text{req}}], \quad \forall t \in \mathcal{T}.$$
 (3)

$$v_t \in [0, \min\{y_t/\eta, \bar{v}\}], \quad \forall t \in \mathcal{T}.$$
 (4)

Without loss of generality we assume $\eta = 1$.¹

¹Rescale y_t by $1/\eta$.

c) Household Consumption (\mathbf{d}_t) : Consumption vector $\mathbf{d}_t = (d_{t1}, \cdots, d_{tK})$ models the flexible consumption of K controllable devices within interval t which satisfy :

$$0 \leq \mathbf{d}_t \leq \bar{\mathbf{d}} := (\bar{d}_1, \dots, \bar{d}_K), \quad \forall t \in \mathcal{T}.$$
 (5)

The utility $U_t(\mathbf{d}_t)$ of consuming \mathbf{d}_t is linearly separable differentiable concave function with marginal utility $L_t(\mathbf{d}_t)$

$$U_t(\mathbf{d}_t) = \sum_{i=1}^K U_{ti}(d_{ti}), \ L_t(\mathbf{d}_t) := \nabla U_t = (L_{t1}, \dots, L_{tK}).$$

Our current model didn't consider uncontrollable loads. However, this can be handled by modeling it as a sequence of independent random variables with some distribution.

d) Battery storage operation (e_t) : The battery state of charge (SoC) is denoted as $s_t \in [0, B]$ where B represents maximum storage capacity of the battery. Battery charging and discharging rate is denoted as $e_t \in [-\underline{e}, \overline{e}]$, where $e_t \geq 0$ represents charging, and $e_t \leq 0$ represents discharging. Battery charging and discharging have inefficiency constants $\eta_c \in (0, 1]$ and $\eta_d \in (0, 1]$, respectively.

$$e_t \in [-\underline{e}, \overline{e}], \quad \forall t \in \mathcal{T}$$
 (6)

$$s_t \in [0, B], \quad \forall t \in \mathcal{T}$$
 (7)

$$s_{t+1} = s_t + \eta_c e_t \mathbb{1}_{e_t \ge 0} - \frac{e_t}{\eta_d} \mathbb{1}_{e_t \le 0}, \quad \forall t \in \mathcal{T}.$$
 (8)

e) Net consumption (z_t) : The net energy consumption of a household within interval t is defined as the net energy measured by the revenue meter during the billing period. This includes EV charging, controllable loads, storage operation and BTM DER generation within the billing period.

$$z_t := v_t + \mathbf{1}^T \mathbf{d}_t + e_t - r_t, \quad t \in \mathcal{T}.$$

We say that household is *net consuming* when $z_t > 0$, and *net producing* when $z_t < 0$.

B. NEM ToU tariff model

a) NEM payment $(P^{\pi_t}(z_t))$: Under the NEM tariff program, a household is billed or credited based on the net energy consumption during each billing period, which can range from 5 minutes to a day or a month. To simplify the notation, we matched the length of our decision intervals with the billing period. This enables us to index the billing period with t.

At interval t, given NEM tariff parameters $\pi_t = (\pi_t^0, \pi_t^-, \pi_t^+) \in \mathbb{R}^3_+$, household with net consumption z_t pays

$$P^{\pi_t}(z_t) := z_t(\mathbb{1}_{z_t \ge 0} \pi_t^+ + \mathbb{1}_{z_t \le 0} \pi_t^-) + \pi_t^0,$$

where π_t^+ is a *retail rate*, π_t^- is a *sell rate*, and π_t^0 is a fixed charge. A prosumer pays the retail rate for the net consumption, while credited by the sell rate for the net generation. The NEM tariff model was proposed in [2].².



Fig. 2: ToU scheme and decision horizon

b) NEM ToU tariff model: ToU tariff divides 24 hours into periods with different energy prices. For our analysis, we adopt a ToU tariff with two distinct energy prices : offpeak and on-peak. A typical ToU tariff sets the on-peak hours to a five-hour period in the late afternoon and early evening [20]. In this work, we assume that the decision horizon falls within the first off-peak period, the on-peak period, and the second off-peak period as shown in Fig.2³. Hence, once we know the connection time and the deadline, the scheduling horizon is well defined : $\mathcal{T} = \mathcal{T}_{off,1} \cup \mathcal{T}_{on} \cup \mathcal{T}_{off,2}$, with $T = |\mathcal{T}_{off,1}| + |\mathcal{T}_{on}| + |\mathcal{T}_{off,2}|$.

We assume that all the intervals in the same period have the same NEM parameters and NEM parameters for every interval is known *a priori*. Hence,

$$\pi_t = \begin{cases} \pi_{\text{on}} := (\pi_{\text{on}}^-, \pi_{\text{on}}^+), & t \in \mathcal{T}_{\text{on}}.\\ \pi_{\text{off}} := (\pi_{\text{off}}^-, \pi_{\text{off}}^+), & t \in \mathcal{T}_{\text{off},1} \cup \mathcal{T}_{\text{off},2}. \end{cases}$$
(9)

We also assume that NEM parameters satisfy $\pi_{\text{off}}^- < \pi_{\text{on}}^- < \pi_{\text{off}}^+ < \pi_{\text{on}}^+$ to avoid the BTM battery is used for the price arbitrage and simplify the problem.

C. Co-optimization of EV charging, consumption, and storage

We formulate the coordinated optimization problem into stochastic DP with state $x_t = (s_t, y_t, r_t, \pi_t)$. The initial SoC (s_0) and EV charging demand (s_{req}) is given.

The policy $\mu := (\mu_0, \dots, \mu_{T-1})$ is a sequence of functions that maps the current state to EV charging, consumption and storage operation :

$$\mu_t(x_t) := (v_t, e_t, \mathbf{d}_t), \quad t \in \mathcal{T}.$$
 (10)

The stage reward for the stochastic DP is the household surplus under the NEM tariff and the terminal reward is the sum of the salvage value of battery storage and the penalty for incomplete EV charging demand. The salvage value of battery storage is proportional to the terminal SoC, and the penalty for incomplete EV charging is proportional to the remaining charging demand at t = T.

$$g_t(x_t, v_t, e_t, \mathbf{d}_t) := \begin{cases} U_t(\mathbf{d}_t) - P^{\pi_t}(z_t), & t = 0, \dots, T-1. \\ \beta s_T - \alpha y_T, & t = T. \end{cases}$$

(

 $^{3}\mbox{We}$ refer period as the set of consecutive decision intervals with the same price.

²Fixed charge doesn't affect the optimal decision, so we assume $\pi_t^0 = 0$.

Co-optimization problem is :

$$\mathcal{P}: \max_{\mu} \quad \mathbb{E}_{\mathbf{r}} \left[\sum_{t=0}^{T-1} g_t(x_t, v_t, e_t, \mathbf{d}_t) - \alpha y_T + \beta s_T \right] \quad (11)$$

s.t. (1) - (10).

We denote the optimal value function as $V_t(x_t)$ that satisfies following Bellman Equation :

$$V_t(x_t) = \max_{v,e,\mathbf{d}} \{ g_t(x_t, v, e, \mathbf{d}) + \mathbb{E}[V_{t+1}(x_{t+1})] \}$$

s.t. (1) - (9) (12)

D. Model Assumption

A1. *Price - penalty - salvage value relation* : We assume that the NEM tariff parameters, the penalty for incomplete charging demand, and the battery salvage value satisfies

$$\pi_{\rm off}^- < \pi_{\rm on}^- < \eta_c\beta < \beta/\eta_d < \pi_{\rm off}^+ < \pi_{\rm on}^+ < \alpha$$

The intuition behind a high penalty for incomplete charging demand at the deadline is to allow EV charging from the BTM DER and grid purchases in all intervals, and to minimize incomplete charging demand. The penalty can be interpreted as the price of charging the EV for unfulfilled charging demand at the deadline.

An intuitive explanation for the price-salvage value relation is that 1) β/η_d should always be greater than π_t^- because we can always sell the energy stored in the battery to the grid at the price of π_t^- . To avoid depleting storage, it should be greater than π_t^- . 2) $\beta\eta_c$ should be less than π_t^+ for all t because we can always purchase energy from the grid at the price of π_t^+ . Consequently, to avoid the battery remaining fully charged at all time, $\beta\eta_c$ should be less than π_t^+ .

III. PROCRASTINATION THRESHOLD POLICY (CO-OPTIMIZATION WITHOUT BATTERY)

We first present the *procrastination threshold policy* which is the optimal EV owner's decision without a battery [4] to provide an intuition of the procrastination threshold policy. Here, \tilde{V}_t represents the value function of the MDP without storage.

Theorem 1 (Procrastination threshold policy). *NEM ToU* tariff parameters satisfying A1, and r_t is a sequence of independent random variables, optimal EV charging v_t^* and consumption d_t^* decisions are monotone increasing function of r_t , and :

$$\begin{split} v_t^* &= \begin{cases} h_{\tau_t}(y_t), & r_t < \Delta_t^+(y_t) \\ h_{w_{t+1}(\nu)}(y_t), & \Delta_t^+(y_t) \le r_t \le \Delta_t^-(y_t) \\ h_{\delta_t}(y_t), & r_t > \Delta_t^-(y_t). \end{cases} \\ d_{ti}^* &= \begin{cases} l_{ti}(\pi_t^+), & r_t < \Delta_t^+(y_t) \\ l_{ti}(\nu), & \Delta_t^+ \le r_t \le \Delta_t^-(y_t) \\ l_{ti}(\pi_t^-), & r_t > \Delta_t^-(y_t), \end{cases} \\ \forall \, i = 1, \dots, K, \end{split}$$

where $\nu \in [\pi_t^-, \pi_t^+]$ satisfies, $r_t = v_t^* + \sum_{i=1}^K d_{ti}^*$.

$$\begin{split} \bar{V}_t(y) &:= \mathbb{E}[\bar{V}_t(y, r_t, \pi_t)], \quad w_t(\pi) := (\partial \bar{V}_t)^{-1}(-\pi), \\ h_\theta(y) &:= \min\{\bar{v}, \max\{y - \theta, 0\}\}, \\ l_{ti}(\pi) &:= \min\{L_{ti}^{-1}(\pi), \bar{d}_i\}, \quad l_t(\pi) := \sum_{i=1}^K l_{ti}(\pi), \\ \Delta_t^-(y_t) &:= l_t(\pi_t^-) + h_{\delta_t}(y_t) \\ \Delta_t^+(y_t) &:= l_t(\pi_t^+) + h_{\tau_t}(y_t). \end{split}$$

The procrastination threshold policy at the interval t is characterized by two procrastination thresholds τ_t and δ_t . These thresholds represent the limit of procrastination for charging the EV. The limit of procrastination for charging EV with grid purchase is τ_t . If $y_t \leq \tau_t$, the controller will postpone charging the EV unless there is surplus BTM DER after fulfilling consumption, as shown in Fig.3. If $y_t \geq \tau_t$, energy will be purchased from the grid to charge the EV if DER is not sufficient, but EV charging using the grid energy will not exceed $y_t - \tau_t$. On the other hand, δ_t represents thresholds of using DER to charge the EV. Similarly, if $y_t \leq$ δ_t , EV is not charged even if there are surplus renewables.

An intuitive explanation for the optimality of procrastination charging behavior is that it increases the likelihood of completing the job using DER, which is a cheaper charging option. Let's consider special case-a charging session with a single price-to gain insights into the policy. In this case, the limit of procrastination for purchasing energy is the maximum charging capacity of remaining intervals : $\tau_t = (T - t - 1)\overline{v}$. Since there is no difference between purchasing at the current interval and in later intervals, it will postpone purchasing energy for charging the EV unless it's necessary, and anticipate future renewables to complete the job with lower price.

The procrastination threshold policy is optimal if the household is either net consuming or net producing, which corresponds to the BTM DER being less than Δ_t^+ , and greater than Δ_t^- , respectively. The optimal consumption of in these regions has the marginal value of π_t^+ and π_t^- .

In between these two thresholds, the total household load, which is the sum of EV charging and consumption, matches



Fig. 3: Procrastination threshold policy for $\delta_t = 0 < y_t < \tau_t$. In such y_t , controller will delay charging EV charging using the DER. So if renewable level is low, $r_t \leq \Delta_t^+$, EV is not charged, and EV is charged only when there are surplus renewables, $r_t \geq \Delta_t^+$.

the current BTM DER generation, meaning the household is in the net-zero zone. In this region, BTM DER is allocated to consumption and EV charging so that the marginal utility matches the marginal reward for charging EV, $\partial \bar{V}_t$. The marginal reward for charging EV is the expected price of energy that can be saved by charging at the current interval.

As the utility function and NEM payment is concave, it's easy to show that value function is also concave (Proof in the [19]). Then, the subdifferential of $\bar{V}_t(y)$ is non-empty, compact and monotone set. Hence, $h_t(\pi)$ is well defined. Also, due to the concavity of $\bar{V}_t(y)$ and $U_t(\mathbf{d})$, optimal EV charging and consumption is monotone increasing with respect to r_t as in the Fig.3.

Following proposition shows procrastination behavior of purchasing energy to charge EV.

Proposition 1 (Procrastination charging behavior). *NEM ToU tariff parameters satisfying A1, and DER is modeled as a sequence of independent random variables, the optimal charging decisions satisfy*

If
$$z_{t_0}^* v_{t_0}^* > 0$$
 for $t_0 \in \mathcal{T}_x$ then $v_t^* = \overline{v}, t_0 < \forall t \in \mathcal{T}_x$,
where \mathcal{T}_x is $\mathcal{T}_{\text{off},1}, \mathcal{T}_{\text{off},2}, \mathcal{T}_{\text{on}}$.

This proposition implies that once the controller is purchasing energy from the grid to charge the EV, it will charge at the maximum capacity in the remaining period, as if it had reached the limit of procrastination and devoting its full capacity.

IV. OPTIMAL PROSUMER DECISIONS

We now present a prosumer's decision for EV charging, flexible loads, and storage.

A. Myopic co-optimization

We will use the result from the previous section and consider a co-optimization with myopic storage operations to avoid solving infinite-dimensional optimization and to unveil the structure of the policies. The battery is myopic in the sense that its storage operations ignore future storage operations and renewable generations, and immediately generate stage reward, which are the salvage values of storage. Then, the Bellman equation for the myopic co-optimization becomes :

$$V_{t}^{M}(x_{t}) = \max_{v,e,\mathbf{d}} \quad g_{t}(x_{t},v,e,\mathbf{d}) + \left(\eta_{c}\mathbb{1}_{e_{t}\geq0} - \mathbb{1}_{e_{t}\leq0}/\eta_{d}\right)\beta e_{t} \\ + \mathbb{E}[V_{t+1}^{M}(s_{t},y_{t}-v,r_{t+1},\pi_{t+1})] \\ \text{s.t.} \quad -\min\{\underline{e},s_{t}\eta_{d}\} \leq e_{t} \leq \min\{\overline{e},(B-s_{t})/\eta_{c}\} \\ (4),(5).$$
(13)

where $V_t^M(x_t)$ is the optimal value function for the myopic co-optimization. For the rest of the paper, we will denote the maximal charging and discharging rate constrained by SoC limit as $\underline{e}' := \min\{\underline{e}, s_t\eta_d\}, \ \overline{e}' = \min\{\overline{e}, (B - s_t)/\eta_c\}.$

Theorem 2 (Myopic optimal policy). *NEM ToU tariff parameters satisfying A1, and* r_t *is a sequence of independent random variables, the optimal net consumption* z_t^* *of myopic*

co-optimization (13) is a piecewise linear monotone decreasing function of r_t :

$$z_t^* = \begin{cases} r_t - \Delta_t^+(y_t), & r_t < \Delta_t^+(y_t) \\ 0, & \Delta_t^+(y_t) \le r_t \le \Delta_t^-(y_t) \\ r_t - \Delta_t^-(y_t), & \Delta_t^-(y_t) < r_t. \end{cases}$$

The optimal EV charging v_t^* , consumption d_t^* and the storage e_t^* decisions are monotone increasing function of r_t , segmented by 6 thresholds on $r_t : \Delta_t^+(y_t), \Delta_{t,1}(y_t) - \Delta_{t,4}(y_t)$, and $\Delta_t^-(y_t)$, and they are decided by Algorithm 1.

$$\begin{split} \Delta_t^{+'}(y_t) &:= \max\{\Delta_t^+(y_t) - \underline{e}', 0\},\\ \Delta_{t,1}(y_t) &:= \max\{l_t(\beta/\eta_d) + h_{\sigma_t^+}(y_t) - \underline{e}', 0\},\\ \Delta_{t,2}(y_t) &:= l_t(\beta/\eta_d) + h_{\sigma_t^+}(y_t),\\ \Delta_{t,3}(y_t) &:= l_t(\beta\eta_c) + h_{\sigma_t^-}(y_t),\\ \Delta_{t,4}(y_t) &:= l_t(\beta\eta_c) + h_{\sigma_t^-}(y_t) + \overline{e}',\\ \Delta_t^{-'}(y_t) &:= \Delta_t^-(y_t) + \overline{e}', \end{split}$$

where $\tau_t > \sigma_t^+ > \sigma_t^- > \delta_t$ for all $t \in \mathcal{T}$.

If $r_t < \Delta_t^{+'}$ and $r_t > \Delta_t^{-'}$ (the first and the last case in **Algorithm** 1), the household is net consuming and net producing, and in these cases, consumption, and EV charging decisions follow procrastination threshold policy introduced in the section III.

In between these two thresholds, the household is in the net-zero zone. Here, two additional procrastination thresholds, σ_t^+ and σ_t^- , are introduced for the EV charging decisions. These thresholds determine the priority of EV charging and storage operations for allocating BTM DER. When $y_t < \sigma_t^-$, it implies that the remaining EV charging demand is small or deadline is still far away. Hence, priority is given to charging the storage over charging the EV. For instance, in the Fig. 4, $v_t^* > 0$ when $r_t > \Delta_{t,4}$ and $e_t^* = \bar{e}'$.

Similarly, for $y_t < \sigma_t^+$, the priority is given to the discharging storage over charging the EV, which means the EV will not charge using stored energy in the BTM battery. For instance, in the Fig. 4 $v_t^* = 0$ when $e_t^* < 0$.

In the net-zero zone, either the total household load (sum of consumption and EV charging) or storage operation



Fig. 4: Myopic optimal policy for $\delta_t = 0 < y_t < \sigma_t^- < \sigma_t^+$. For such y_t , charging and discharging the storage is prioritized over EV charging. Hence, the EV is charged only if storage is charged to maximum, and the EV is not charged when storage is discharging.

Algorithm 1 Myopic optimal decisions

Input: y_t, π_t, r_t **Output:** $v_t^*, e_t^*, \mathbf{d}_t^*$ For all i = 1, ..., K, myopic optimal decisions of each device are : (with a slight abuse of notation, $V_t(y)$:= $\mathbb{E}[V_t^M(y, r_t, \pi_t))],$ $w_t(\pi) := (\partial \bar{V}_t)^{-1}(-\pi))$ if $r_t < \Delta_t^{+'}(y_t)$ then $v_t^* \leftarrow h_{\tau_t}(y_t), \quad d_{ti}^* \leftarrow l_{ti}(\pi_t^+), \quad e_t^* \leftarrow -\underline{e}',$ else if $r_t \in \left[\Delta_t^{+'}(y_t), \Delta_{t,1}(y_t)\right)$ then $v_t^* \leftarrow h_{w_{t+1}(\nu)}(y_t), \quad d_{ti}^* \leftarrow l_{ti}(\nu), \quad e_t^* \leftarrow -\underline{e}',$ where $r_t + \underline{e}' = v_t^* + \sum_{i=1}^K d_{ti}^*,$ and $\nu \in [\beta/\eta_d, \pi_t^+],$ else if $r_t \in \left[\Delta_{t,1}(y_t), \Delta_{t,2}(y_t)\right)$ then $v_t^* \leftarrow h_{\sigma^+}(y_t), \quad d_{ti}^* \leftarrow l_{ti}(\beta/\eta_d), \quad e_t^* \leftarrow r_t - \Delta_{t,2},$ else if $r_t \in [\Delta_{t,2}(y_t), \Delta_{t,3}(y_t)]$ then $v_t^* \leftarrow h_{w_{t+1}(\nu)}(y_t), \quad d_{ti}^* \leftarrow l_{ti}(\nu), \quad e_t^* \leftarrow 0,$ where $r_t = v_t^* + \sum_{i=1}^K d_{ti}^*, \text{ and } \nu \in [\beta \eta_c, \beta/\eta_d]$ else if $r_t \in \left[\Delta_{t,3}(y_t), \Delta_{t,4}(y_t)\right)$ then $v_t^* \leftarrow h_{\sigma^-}(y_t), \quad d_{ti}^* \leftarrow l_{ti}(\beta \eta_c), \quad e_t^* \leftarrow r_t - \Delta_{t,3},$ else if $r_t \in \left[\Delta_{t,4}(y_t), \Delta_t^{-'}(y_t)\right)$ then $\begin{aligned} & v_t^* \leftarrow h_{w_{t+1}(\nu)}(y_t), \quad d_{ti}^* \leftarrow l_{ti}(\nu), \quad e_t^* \leftarrow \bar{e'}, \\ & \text{where } r_t - \bar{e'} = v_t^* + \sum_{i=1}^K d_{ti}^*, \text{ and } \nu \in [\pi_t^-, \beta\eta_c] \end{aligned}$ else $v_t^* \leftarrow h_{\delta_t}(y_t), \quad d_{ti}^* \leftarrow l_{ti}(\pi_t^-), \quad e_t^* \leftarrow \bar{e}'$ end if

increases with renewables. When, the battery operation is a constant function of r_t (case 2, 4, and 6 in Algorithm 1), the total household load matches $r_t - e_t$, BTM DER adjusted by storage output. For instance, if $r_t \in [\Delta_t^{+'}(y_t), \Delta_{t,1}(y_t))$, the total household load matches sum of BTM DER and maximal storage discharging. The storage output adjusted renewables are allocated as the net-zero zone decisions of the co-optimization without storage.

One thing to note that, by introducing the storage, the household has a wider net zero zone compared to the cooptimization without storage. This implies increased internal usage of BTM DER within the household.

The optimal myopic storage operation is a piecewise linear function and it shows a complementarity property with the total household load, which is defined as $e_t^* z_t^* \leq 0$ [3]. This implies that energy is stored in the battery only using the BTM DER, and storage only discharges to reduce net consumption. This result is intuitive because, from A1, $\pi^- < \beta \eta_c < \pi^+$, so storage charging is beneficial only when using the BTM DER. Also, as $\pi^- < \beta/\eta_d < \pi^+$, discharging stored energy in the battery is beneficial only if it is used to reduce the net load.

B. Co-optimization with non-binding SoC assumption

Now, we make the following additional assumption on the battery SoC limit :

A2. Non-binding battery SoC limit : Battery SoC limit constraint, (7), is non binding for all $t \in \mathcal{T}$.

This assumption is restrictive but it can be valid if the storage capacity is large relative to the maximum storage charging and discharging rate, or if the maximum charging and discharging rate is relatively low. Under assumption A2, we can show that the optimal policy for the stochastic co-optimization problem (11) becomes myopic.

Theorem 3 (Optimal prosumer decisions). Under A1-A2, and BTM DER modeled as a sequence of independent random variables, optimal battery, consumption, EV charging decisions of Theorem 2 is the optimal solution of the problem (11).

C. Computation of the thresholds

We now comment on the computation of the thresholds that characterize the myopic optimal policy. The renewable thresholds, Δ_t 's, are comprised of thresholds on consumption and EV charging. Flexible loads are not coupled between time, hence given consumer utility function, NEM parameter, and storage parameters, we can compute the consumption thresholds offline.

The following proposition presents the characterization of procrastination thresholds.

Proposition 2 (Characterization of procrastination thresholds). *NEM ToU tariff parameters satisfying A1, and DER is modeled as a sequence of independent random variables, the optimal procrastination thresholds* $\tau_t, \sigma_t^+, \sigma_t^-, \delta_t$ *satisfy :*

$$\tau_{t} = \begin{cases} (T - t - 1)\bar{v}, & t \in \mathcal{T}_{\text{off},2} \cup \mathcal{T}_{\text{on}} \\ w_{t+1}(\pi_{\text{off}}^{+}), & t \in \mathcal{T}_{\text{off},1} \end{cases}$$
$$\delta_{t} = \begin{cases} 0, & t \in \mathcal{T}_{\text{off},1} \cup \mathcal{T}_{\text{off},2}, \\ 0, & t \in \mathcal{T}_{\text{on}} \wedge \mathcal{T}_{\text{off},2} = \emptyset \\ w_{t+1}(\pi_{\text{on}}^{-}), & t \in \mathcal{T}_{\text{on}} \wedge \mathcal{T}_{\text{off},2} \neq \emptyset \end{cases}$$
$$\sigma_{t}^{+} = w_{t+1}(\beta/\eta_{d})$$
$$\sigma_{t}^{-} = w_{t+1}(\beta\eta_{c})$$

 τ_t and δ_t satisfy following recursive relation within same pricing period :

$$\tau_t = \tau_{t+1} + \bar{v},$$

$$\delta_t = \delta_{t+1}.$$

In intervals with the purchasing procrastination threshold $\tau_t = (T - t - 1)\bar{v}$, energy is purchased for charging EV only if it's necessary to meet the charging requirement by the deadline. Combining the result of Proposition 1-2, the procrastination threshold policy always guarantees completion of the charging requirement by the deadline whenever it's possible ($s_{\text{req}} \leq T\bar{v}$).

Other thresholds that involve $w_{t+1}(\cdot)$ can be computed offline by solving dynamic programming equations. Such procrastination thresholds depend on the distribution of future renewable generation, and represent value of y_t such that expected value of charging is equal to the argument of $w_{t+1}(\cdot)$.



Fig. 5: Relative performance gap of the MO(blue), CCO(red), CSO(yellow), NCO(purple), CR(green) to the oracle policy for different $\pi_t^+ - \pi_t^-$ is plotted. The estimated mean of renewable is scaled by 0.5, 1, 1.5 for the figures from left to right.

V. NUMERICAL RESULTS

We implemented numerical simulations using real-world data, which involve a household with the BTM storage and EV under the utility's NEM tariff, to verify the performance of the myopic optimal policy presented in this paper and demonstrate the benefits of co-optimization.

We compared the performance of different policies using the relative expected accumulated surplus gap to the oracle policy. The oracle policy is the offline policy which is the outcome of solving convex optimization (11) with the realization of renewables, which will serve as the upper bound of the performance. If the oracle policy and comparing policy's expected accumulated reward is R_{OR} , and R_{μ} , respectively, the performance measure of the policy is

$$\frac{R_{OR} - R_{\mu}}{R_{OR}}$$

The myopic optimal policy (MO) in Theorem 2 is compared with 4 different policies with different level of cooptimization. The policies we considered are :

- 1) Cost reduction policy (CR)
- 2) Non co-optimization policy (NCO)
- 3) Consumption-storage co-optimized policy (CSO)
- 4) Charging-consumption co-optimized policy (CCO)

Policies are listed in the order of increasing levels of cooptimization. CR uses renewables and storage to reduce the costs, with the EV primarily charges in off-peak periods, while consumption is not optimized. NCO and CSO optimize the EV first, while NCO optimizes each device separately in the order of EV, consumption, and storage. CSO co-optimizes consumption and storage. CCO co-optimizes consumption and EV charging as Theorem 1, and then optimizes storage. Essentially, all the policies use the battery to reduce the cost and store surplus renewables.

A. Simulation setting

We implemented Monte Carlo simulations with random EV charging demand, BTM DER trajectory, and connection time. The distribution of EV charging demand was modeled using the charging request data from Adaptive Charging Network [21]. To model the distribution of rooftop solar generation, we used New York residential solar generation data from the Pecan Street [22]. Lastly, we assumed that the connection time of the EV is equally distributed.

For the NEM ToU tariff parameters, we used the ToU retail rate from the Pacific Gas and Electronic (PG&E)⁴ with onpeak hours from 4 PM to 9 PM. The sell rate was selected as a parameter to vary.

The BTM storage has a maximum capacity of B = 13.5 kWh and we assumed that $s_0 = 0.5B$. The maximum charging and discharging rate are $\bar{e} = \underline{e} = 3.2$ kW, and the charging/discharging inefficiency constants are $\eta_c = \eta_d = 0.95$. The chosen salvage value of storage, β , satisfies A1.

In our simulation, we considered a household with a single controllable load, modeling total consumption of the household. We adopted a quadratic concave utility function that has the form $U(d) = ad - 1/2bd^2$. The coefficients of the utility function were estimated based on the consumption data from Pecan Street, and the ToU retail rate.

As mentioned earlier, the EV charger has an efficiency of 1 and the maximum charging capacity is 3.6 kW.

B. Performance of myopic optimal policy

We present the simulation results of 100,000 Monte Carlo runs that show the performance gap of different policies when the retail rate is fixed and the sell rate is varied, assuming that $\pi_t^+ - \pi_t^-$ is kept constant at all intervals. The decision horizon T is fixed as 16 hours. The three different plots in Fig.5 show the relative surplus gap when the mean of the

⁴PG&E ToU tariff data can be found at PGE E ToU-B.

distribution of DER generation is scaled by 0.5, 1, 1.5 from left to right.

In the plots, we can observe that the MO has the smallest gap with the oracle policy ranging from 0.5-7.5% in all scenarios. This result shows the value of co-optimization. The performance gap increases as the mean of the renewable increases, because with a higher renewable level, more renewables will be stored in the battery, and it is more likely that the non-binding SoC limit assumption will be violated.

When π_t^- is lowered, the performance gap of MO is increasing. For instance, for 50% renewable means case, performance gap is increased from 0.6% to 3.8%. A low sell rate means that BTM DER generation has low value when it's exported to the grid. It also implies using renewable internally is cheaper in terms of opportunity cost. Hence, to complete charging requirement of EV using the renewables, procrastination thresholds $(\tau_t, \sigma_t^+, \sigma_t^-)$ are increased. This procrastinating behavior allocates more renewables to the storage, increasing the likelihood of violating the SoC limit non-binding assumption.

We can also observe that the gap with MO to other policies (CCO, CSO, NCO) are shrinking as the mean of the renewable distribution increases. The biggest gap between MO and NCO is reduced from 6.4% for the 50% renewable mean case to 1.9% for 150% renewable mean case. When the renewable level is high, the decisions of co-optimization policy get closer to the decisions of non co-optimization policies because there are sufficient renewables to be used for all devices. Hence, this result implies that the value of co-optimization is maximized when renewable level is low.

VI. CONCLUSION

This paper has presented a comprehensive analysis of co-optimizing BTM DER with storage, EV charging and controllable loads under NEM tariff. Our results suggest that under the NEM tariff, procrastination threshold policy for EV charging and myopic storage operations provide economic benefit to the household. Furthermore, the thresholds that characterize the procrastination threshold policy can be computed offline and the policy can be easily implemented in real-time.

It should be noted that the scope of this study was limited to co-optimizing a single EV with BTM DER and storage, and as such, application of our work to commercial or industrial EMS settings, which can have multiple EVs, is not trivial. There reside priority rules of different EVs, and high dimensionality of the control space can complicate the problem. This problem can be studied in future works. Also, our work assumes that we have information on the consumer utility and the BTM DER distribution. Modelfree reinforcement learning approach using the structure of the proposed policy can be studied in the future works to avoid modeling of unknown utility function and BTM DER generation distributions.

REFERENCES

- G. Barbose, S. Elmallah, and W. Gorman, "Behind-the-meter solar + storage : Market data and trends," Lawrence Berkeley National Laboratory, July 2021.
- [2] A. S. Alahmed and L. Tong, "On net energy metering X: Optimal prosumer decisions, social welfare, and cross-subsidies," *IEEE Transactions on Smart Grid*, 2022.
- [3] A. S. Alahmed, L. Tong, and Q. Zhao, "Co-optimizing distributed energy resources in linear complexity under net energy metering," *arXiv preprint arXiv:2208.09781*, 2022.
- [4] M. Jeon, L. Tong, and Q. Zhao, "Co-optimizing consumption and EV charging under net energy metering," arXiv preprint arXiv:2211.11088, 2022.
- [5] F. Hafiz, A. R. de Queiroz, and I. Husain, "Coordinated control of PEV and PV-based storages in residential systems under generation and load uncertainties," *IEEE Transactions on Industry Applications*, vol. 55, no. 6, pp. 5524–5532, 2019.
- [6] M. A. A. Pedrasa, T. D. Spooner, and I. F. MacGill, "Coordinated scheduling of residential distributed energy resources to optimize smart home energy services," *IEEE Transactions on Smart Grid*, vol. 1, pp. 134–143, 9 2010.
- [7] O. Erdinc, N. G. Paterakis, T. D. Mendes, A. G. Bakirtzis, and J. P. Catalão, "Smart household operation considering bi-directional EV and ESS utilization by real-time pricing-based DR," *IEEE Transactions on Smart Grid*, vol. 6, pp. 1281–1291, 5 2015.
- [8] S. Seal, B. Boulet, V. R. Dehkordi, F. Bouffard, and G. Joos, "Centralized MPC for home energy management with EV as mobile energy storage unit," *IEEE Transactions on Sustainable Energy*, pp. 1–10, 1 2023.
- [9] M. Yousefi, A. Hajizadeh, M. N. Soltani, and B. Hredzak, "Predictive home energy management system with photovoltaic array, heat pump, and plug-in electric vehicle," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 430–440, 2020.
- [10] X. Xu, Y. Jia, Y. Xu, Z. Xu, S. Chai, and C. S. Lai, "A multi-agent reinforcement learning-based data-driven method for home energy management," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3201–3211, 2020.
- [11] T. Hubert and S. Grijalva, "Modeling for residential electricity optimization in dynamic pricing environments," *IEEE Transactions on Smart Grid*, vol. 3, pp. 2224–2231, 2012.
- [12] T. Li and M. Dong, "Real-time residential-side joint energy storage management and load scheduling with renewable integration," *IEEE Transactions on Smart Grid*, vol. 9, pp. 283–298, 1 2018.
- [13] Y. Xu and L. Tong, "Optimal operation and economic value of energy storage at consumer locations," *IEEE Transactions on Automatic Control*, vol. 62, pp. 792–807, 2 2017.
- [14] Y. Guo, M. Pan, and Y. Fang, "Optimal power management of residential customers in the smart grid," *IEEE Transactions on Parallel* and Distributed Systems, vol. 23, pp. 1593–1606, 2012.
- [15] L. Yu, W. Xie, D. Xie, Y. Zou, D. Zhang, Z. Sun, L. Zhang, Y. Zhang, and T. Jiang, "Deep reinforcement learning for smart home energy management," *IEEE Internet of Things Journal*, vol. 7, pp. 2751–2762, 4 2020.
- [16] T. T. Kim and H. V. Poor, "Scheduling power consumption with price uncertainty," *IEEE Transactions on Smart Grid*, vol. 2, pp. 519–527, 9 2011.
- [17] C. Chen, J. Wang, Y. Heo, and S. Kishore, "MPC-based appliance scheduling for residential building energy management controller," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1401–1410, 2013.
- [18] D. Wu, G. Rabusseau, V. François-lavet, D. Precup, and B. Boulet, "Optimizing home energy management and electric vehicle charging with reinforcement learning," *Proceedings of the 16th Adaptive Learning Agents*, 2018.
- [19] M. Jeon, L. Tong, and Q. Zhao, "On the optimality of procrastination policy for EV charging under net energy metering," *arXiv preprint*, 2023.
- [20] A. Faruqui, R. Hledik, and S. Sergici, "A survey of residential timeof-use (ToU) rates," Available at www.shorturl.at/chTU3 (2023/1/31).
- [21] Z. J. Lee, T. Li, and S. H. Low, "ACN-Data: Analysis and Applications of an Open EV Charging Dataset," in *Proceedings of the Tenth International Conference on Future Energy Systems*, ser. e-Energy '19, Jun. 2019.
- [22] "Pecan street dataset," Available at www.pecanstreet.org/dataport/ (2022/11/01).

APPENDIX A

Proofs

The optimality of procrastination threshold policy for the co-optimization without storage (Theorem 1) was proved [4]. We will prove the Proposition 1 which wasn't stated in [4].

A. Proof of Proposition 1

Proof. By the proposition 2 in [4], τ_t and τ_{t+1} satisfy :

$$\tau_t = \tau_{t+1} + \bar{v}$$

For some $t_0 \in \mathcal{T}_x$, given y_{t_0} , if $z_{t_0}^* v_{t_0}^* > 0$, by the procrastination threshold policy, $y_{t_0} > \tau_{t_0}$ and,

$$v_{t_0}^* = \min\{\bar{v}, y_{t_0} - \tau_{t_0}\}.$$

Then, $y_{t_0+1} = y_{t_0} - v_{t_0}^* \ge \tau_{t_0} = \tau_{t_0+1} + \bar{v}$. Hence, by the monotonicity of v_t^* with respect to r_t , $v_{t_0+1}^* = \bar{v}$ for all r_{t_0+1} .

The recursive relation of τ_t holds within the same period, therefore, $v_t^* = \bar{v}$ for all $t \in \mathcal{T}_0$ such that $t > t_0$.

B. Concavity of the optimal value function

We first prove the concavity of the optimal value function of myopic co-optimization. Here, we are only interested in concavity of the value function with respect to y_t . So, let's simplify the notation as $V_t^M(y_t)$.

Lemma 1. Optimal value function of myopic cooptimization, $V_t^M(y_t)$, that satisfies Bellman equation (??) is concave function of y_t .

Proof. At t = T - 1, there is no uncertainty and Bellman equation becomes

$$V_{T-1}^{M}(y_{T-1}) = \max_{v,e,\mathbf{d}} \{ g_{T-1}(x_{T-1}, v, e, \mathbf{d}) - \alpha(y_{T-1} - v) \\ + \beta (\mathbb{1}_{e \ge 0} \eta_c - \mathbb{1}_{e \le 0} / \eta_d) (s_{T-1} + e) \}$$

s.t. (4) - (5)

As $\alpha > \pi^+ > \pi^-$, the optimal EV charging action is

$$v^* = \min\{\bar{v}, y_{T-1}\}$$

For $y_{T-1} > \overline{v}$, optimal consumption and storage operation are independent to y_{T-1} and $\frac{\partial}{\partial y_{T-1}}V_{T-1}^M = -\alpha$. For $y_{T-1} \le \overline{v}$, optimal consumtion and storage operation

For $y_{T-1} \leq \bar{v}$, optimal consumtion and storage operation is decided based on the remaining renewables $\tilde{r}_{T-1} = \max\{r_{T-1} - y_{T-1}, 0\}$. Consumption-storage co-optimization decision is based on myopic co-optimization in [3].

If household is net consuming, $\tilde{r}_{T-1} \leq l_{T-1}(\pi_{T-1}^+) - \underline{e}', \\ \frac{\partial}{\partial y_{T-1}}V_{T-1}^M = -\pi_{T-1}^+ < -\alpha. \\ \text{As } y_{T-1} \text{ reduces, } \tilde{r}_{T-1} \text{ increases, and household will be}$

As y_{T-1} reduces, \tilde{r}_{T-1} increases, and household will be in the net zero zone. In the net zero zone, $\frac{\partial}{\partial y_{T-1}}V_{T-1}^M$ is the negative of marginal utility of consumption, and as the consumption is increasing, marginal utility decreases. Therefore, $\frac{\partial}{\partial y_{T-1}}V_{T-1}^M$ increases. At last, household enters, net producing region, which means $\frac{\partial}{\partial y_{T-1}}V_{T-1}^M = -\pi_{T-1}^-$.

Hence, $\frac{\partial}{\partial y_{T-1}}V_{T-1}^M$ is a decreasing function of y_{T-1} , therefore, $V_{T-1}^M(x_{T-1})$ is a concave function of y_{T-1} .

Suppose, V_{t+1}^M is a concave function of y_{t+1} . Then, the objective function of bellman equation is concave and continuous in the compact feasible set, there exists a optimal solution for every y_t . Let's denote the optimal solution for $y_t = \kappa$, as $v_{\kappa}^*, e_{\kappa}^*, \mathbf{d}_{\kappa}^*$, and for $y_t = \zeta$ as $v_{\zeta}^*, e_{\zeta}^*, \mathbf{d}_{\zeta}^*$.

$$\begin{split} \lambda V_t^M(\kappa) &+ (1-\lambda) V_t^M(\zeta) \\ &= \lambda \{ g_t(\kappa, v_{\kappa}^*, e_{\kappa}^*, \mathbf{d}_{\kappa}^*) + \mathbb{E}[V_{t+1}^M(\kappa - v_{\kappa}^*)] \} \\ &+ (1-\lambda) \{ g_t(\zeta, v_{\zeta}^*, e_{\zeta}^*, \mathbf{d}_{\zeta}^*) + \mathbb{E}[V_{t+1}^M(\zeta - v_{\zeta}^*) \} \\ &\leq g_t(\xi, v_{\xi}, e_{\xi}, \mathbf{d}_{\xi}) + \mathbb{E}[V_{t+1}^M(\xi - v_{\xi})] \\ &\leq g_t(\xi, v_{\xi}^*, e_{\xi}^*, \mathbf{d}_{\xi}^*) + \mathbb{E}[V_{t+1}^M(\xi - v_{\xi}^*)] \\ &= V_t^M(\xi) = V_t^M(\lambda \kappa + (1-\lambda)\zeta) \end{split}$$

Here, $\xi = \lambda \kappa + (1 - \lambda)\zeta$, $v_{\xi} = \lambda v_{\kappa}^* + (1 - \lambda)v_{\xi}^*$, $e_{\xi} = \lambda e_{\kappa}^* + (1 - \lambda)e_{\xi}^*$, $\mathbf{d}_{\xi} = \lambda \mathbf{d}_{\kappa}^* + (1 - \lambda)\mathbf{d}_{\xi}^*$.

The first inequality holds, because of the concavity of the stage reward function of inductive hypothesis.

Hence, V_t^M is a concave function of y_t for all $t \in \mathcal{T}$. \Box

C. Storage operation propositions

For the rest of the proofs, $d_t^* = \mathbf{1}^T \mathbf{d}_t^*$. The proof of the following propositions are shown in [3]

Proposition 3 (Storage salvage value under non-binding SoC assumption). Under assumption A1 and A2, for $\epsilon > 0$ that doesn't violate non-binding SoC assumption,

$$V_t(s_t + \epsilon, y_t, r_t) = V_t(s_t, y_t, r_t) + \beta \eta_c \epsilon$$
$$V_t(s_t - \epsilon, y_t, r_t) = V_t(s_t, y_t, r_t) - \beta / \eta_d \epsilon$$

Proposition 4 (Storage-total load complementarity condition). Under A1-A2, optimal storage operations and net consumption of the myopic co-optimization (??) satisfy

$$e_t^* z_t^* \le 0,$$

 $e_t^* (d_t^* + v_t^* - r_t) \le 0$

Proposition 5 (Optimal storage operation). Under A1-A2, for all $t \in T$, for the optimal consumption \mathbf{d}_t^* and EV charging v_t^* of (??), e_t^* satisfies

$$e_t^* = \begin{cases} \max\{r_t - d_t^* - v_t, -\underline{e}'\}, & r_t \le d_t^* + v_t^* \\ \min\{r_t - d_t^* - v_t, \overline{e}'\}, & r_t > d_t^* + v_t^* \end{cases}$$

and $e_t^* \leq r_t$.

D. Proof of Theorem 2

Proof. Let's consider three regions of r_t divided by $\Delta_t^{+'}$ and $\Delta_t^{-'}$.

1) $r_t < \Delta_t^{+'}(y_t)$: Define storage output augmented renewable $\tilde{r}_t := r_t - e_t^*$. Then, by Proposition 5, and $e_t^* \ge -\underline{e}'$,

$$0 < \tilde{r}_t < l_t(\pi_t^+) + h_{\tau_t}(y_t)$$

Hence, by Theorem 1,

$$v_t^* = h_{\tau_t}(y_t), \, d_{ti}^* = l_{ti}(\pi_t^+), \, \forall i = 1, \dots, K.$$

Then, $v_t^* + d_t^* > r_t$, and $e_t^* = -\underline{e}'$. Therefore,

$$z_t^* = v_t^* + d_t^* + e_t^* - r_t = \Delta_t^{+'}(y_t) - r_t$$

2) $r_t > \Delta_t^{-'}(y_t)$: For such r_t , storage output augmented renewable satisfies

$$\tilde{r}_t = r_t - e_t^* > l_t(\pi_t^-) + h_{\delta_t}(y_t)$$

Then, by Theorem 1,

$$v_t^* = h_{\delta_t}(y_t), \ d_{ti}^* = l_{ti}(\pi_t^-), \ \forall i = 1, \dots, K.$$

From Proposition 5, as $v_t^* + d_t^* < r_t$, $e_t^* = \bar{e}'$. Therefore,

$$z_t^* = v_t^* + d_t^* + e_t^* - r_t = \Delta_t^{-'}(y_t) - r_t$$

3) $r_t \in [\Delta_t^{+'}(y_t), \Delta_t^{-'}(y_t)]$: Note that $z_t^* = 0$ for $r_t = \Delta_t^{+'}(y_t)$ and $r_t = \Delta_t^{-'}(y_t)$. Hence, it's suffice to show that z_t^* is a monotone decreasing function of r_t in this region.

For $r_t = \Delta_t^{+'}(y_t)$, and $\epsilon > 0$ such that $r_t + \epsilon \in [\Delta_t^{+'}(y_t), \Delta_t^{-'}(y_t)]$, suppose optimal net consumption of $r_t + \epsilon$ is $\tilde{z}_t^* > 0$. Let's denote the optimal scheduling of each device as $\tilde{v}_t^*, \tilde{d}_t^*, \tilde{e}_t^*$. By Proposition 5, $\tilde{e}_t^* = -\underline{e'}$.

$$r_t + \epsilon - \tilde{e}_t^* = r_t + \epsilon + \underline{e}' \ge r_t - e_t^* = l_t(\pi_t^+) + h_{\delta_t}(y_t) = \Delta_t^+(y_t)$$

The first inequality holds because $e_t^* \ge -\underline{e}'$. The second equality comes from Theorem 1.

By Theorem 1, if $r_t + \epsilon - \tilde{e}_t^* \ge \Delta_t^+(y_t)$, $\tilde{v}_t^* + \tilde{\mathbf{d}}_t^* \le r_t + \epsilon - \tilde{e}_t^*$. Hence, $\tilde{z}_t^* \le 0$ which contradicts the assumption. Therefore, z_t^* is a monotone decreasing function with respect to r_t .

Now let's show that v_t^*, d_t^*, e_t^* are determined by Algorithm 1. From the proof of optimal net consumption, we proved optimal myopic decision for $r_t < \Delta_t^{+'}(y_t)$ and $r + t > \Delta_t^{-'}(y_t)$. Let's prove the myopic optimal schedule in the net-zero zone. In this region, myopic co-optimization becomes,

$$\max_{v,e,\mathbf{d}} \quad U_t(\mathbf{d}) + \left(\eta_c \mathbb{1}_{e_t \ge 0} - \mathbb{1}_{e_t \le 0}/\eta_d\right)\beta e_t \\ + \mathbb{E}[V_{t+1}^M(s_t, y_t - v, r_{t+1}, \pi_{t+1})] \\ \text{s.t.} \quad -\underline{e}' \le e_t \le \overline{e}' \qquad (\bar{\lambda}_e, \underline{\lambda}_e) \\ 0 \le \mathbf{d} \le \overline{\mathbf{d}} \qquad (\bar{\lambda}_d, \underline{\lambda}_d) \\ 0 \le v \le \overline{v} \qquad (\bar{\lambda}_v, \underline{\lambda}_v)$$

$$v + \mathbf{1}^T \mathbf{d} + e = r \qquad (\nu)$$

Here, $(\bar{\lambda}_e, \underline{\lambda}_e)$, $(\bar{\lambda}_d, \underline{\lambda}_d)$, $(\bar{\lambda}_v, \underline{\lambda}_v)$ are the Lagrange multipliers for the corresponding constraints.

Lagrangian of the optimization problem is

$$\mathcal{L} = U(\mathbf{d}) + \mathbb{E}[V_{t+1}^{M}(s_{t}, y_{t} - v, r_{t+1}, \pi_{t+1})] + (\eta_{c} \mathbb{1}_{e_{t} \geq 0} - \mathbb{1}_{e_{t} \leq 0}/\eta_{d})\beta e_{t} + \underline{\lambda}_{e}(e_{t} + \underline{e}') + \bar{\lambda}_{e}(\bar{e}' - e_{t}) + \underline{\lambda}_{d}^{T}\mathbf{d} + \bar{\lambda}_{d}^{T}(\bar{\mathbf{d}} - \mathbf{d}) + \underline{\lambda}_{v}v + \bar{\lambda}_{v}(\bar{v} - v) + \nu(v + \mathbf{1}^{T}\mathbf{d} + e - r)$$

As above optimization satisfies Slater's condition, KKT condition becomes necessary and sufficient condition. Let's

verify myopic optimal decisions in **Algorithm** 1 satisfies KKT condition.

1) $r_t \in [\Delta_t^{+'}(y_t), \Delta_{t,1}(y_t)]$: By complementary slackness, $\underline{\lambda}_e = 0$. Then, $\nu = \beta/\eta_d + \overline{\lambda}_e$.

 $\tilde{r}_t = r_t + \underline{e}' \in [\Delta_t^+(y_t), l_t(\beta/\eta_d) + h_{\sigma_t^+}(y_t))$. Hence, from Theorem 1, optimal EV charging and consumption decisions are

$$\psi_t^* = h_{w_{t+1}(\nu)}(y_t), \, d_{ti}^* = l_{ti}(\nu), \, \forall i = 1, \dots, K$$

where $\nu \in [\pi_t^+, \beta/\eta_d)$. Here, due to the monotonicity of L_t and $\partial(\bar{V}_t)$, v_t^* and d_t^* are monotone increasing with respect to r_t .

2) $r_t \in [\Delta_{t,1}(y_t), \Delta_{t,2}(y_t)]$: By complementary slackness, $\lambda_e = \underline{\lambda}_e = 0$. Hence, $\nu = \beta/\eta_d$. Then, by the stationarity condition, optimal EV charging and consumption decisions are

$$w_t^* = h_{\sigma_t^+}(y_t), \, d_{ti}^* = l_{ti}(\beta/\eta_d), \, \forall i = 1, \dots, K,$$

where $\sigma_t^+ = w_{t+1}(\beta/\eta_d)$. By monotonicity of L_t and $\partial(\bar{V}_t)$, optimal EV charging and consumption decisions are no less than the optimal EV charging and consumption decisions of previous case, respectively. Hence, monotonicity of myopic optimal policy still holds.

3) $r_t \in [\Delta_{t,2}(y_t), \Delta_{t,3}(y_t)]$: As $e_t^* = 0$, by complementary slackness, $\bar{\lambda}_e = \underline{\lambda}_e = 0$. From Theorem 1, optimal EV charging and consumption decisions are

$$w_t^* = h_{w_{t+1}}(y_t), \, d_{ti}^* = l_{ti}(\nu), \, \forall i = 1, \dots, K$$

where $\nu \in [\beta \eta_c, \beta/\eta_d]$ and $v_t^* + d_t^* = r_t$. Similarly, monotonicity of v_t^* and d_t^* holds due to monotonicity of L_{ti} and $\partial(\bar{V}_t)$.

4) $r_t \in [\Delta_{t,3}(y_t), \Delta_{t,4}(y_t)]$: By complementary slackness, $\lambda_e = \underline{\lambda}_e = 0$ and $\nu = \eta_c \beta$. By stationarity condition, optimal EV charging and consumption decisions are

$$v_t^* = h_{\sigma_t^-}(y_t), \, d_{ti}^* = l_{ti}(\beta/\eta_d), \, \forall i = 1, \dots, K,$$

where $\sigma_t^- = w_{t+1}(\beta \eta_c)$. Montonicity of the optimal decisions still holds.

5) $r_t \in [\Delta_{t,4}(y_t), \Delta_t(y_t)^{-'})$: By complementary slackness, $\lambda_e = 0$. Then, $\nu = \beta \eta_c + \underline{\lambda}_e$. For $\tilde{r}_t = r_t - \bar{e}' \in [l_t(\beta \eta_c) + h_{\sigma_t^-}(y_t), \Delta_t^-(y_t))$, by Theorem 1, optimal EV charging and consumption decisions are

$$v_t^* = h_{w_{t+1}}(y_t), \ d_{ti}^* = l_{ti}(\nu), \ \forall i = 1, \dots, K,$$

where $\nu \in [\beta \eta_c, \pi_t^-]$.

Overall, monotonicity of v_t^* , d_t^* holds due to the monotonicity of l_{ti} and $\partial(\bar{V}_t)$.

E. Proof of Theorem 3

Proof. It's suffice to show that for s_t that satisfies A2 and $\gamma = (\epsilon, 0, 0, 0)$ ($\epsilon > 0$), optimal value function that satisfies (12), satisfies

$$V_t(x_t + \gamma) = V_t(x_t) + \beta \epsilon \tag{14}$$

If V_t satisfies above equations, V_t reduces to V_t^M . Let's show above equations by backward induction. For t = T - 1,

$$V_{T-1}(x_{T-1} + \boldsymbol{\gamma}) = \max_{v,e,\mathbf{d}} \{g_{T-1}(x_{T-1} + \boldsymbol{\gamma}, v, e, \mathbf{d}) \\ +\beta(s_{T-1} + \epsilon + \eta_c \mathbb{1}_{e_t \ge 0} e - \mathbb{1}_{e_t < 0}/\eta_d)e) + \alpha(y_{T-1} - \alpha v)\} \\ = \max_{v,e,\mathbf{d}} \{g_{T-1}(x_{T-1}, v, e, \mathbf{d}) + \beta(s_{T-1} + \eta_c \mathbb{1}_{e_t \ge 0} e - \mathbb{1}_{e_t < 0}/\eta_d)e) \\ + \alpha(y_{T-1} - \alpha v)\} + \beta\epsilon = V_{T-1}(x_{T-1}) + \beta\epsilon.$$

Suppose V_{t+1} satisfies (14), then, from (12),

$$V_t(x_t + \boldsymbol{\gamma}) = \max_{v,e,\mathbf{d}} \{g_t(x_t + \boldsymbol{\gamma}, v, e, \mathbf{d}) + \mathbb{E}[V_{t+1}(x_{t+1} + \boldsymbol{\gamma})]\}$$

= $\max_{v,e,\mathbf{d}} \{g_t(x_t, v, e, \mathbf{d}) + \mathbb{E}[V_{t+1}(x_{t+1})]\} + \beta \epsilon$
= $V_t(x_t) + \beta \epsilon.$

F. Proof of Proposition 2

Proof. Procrastination thresholds τ_t and δ_t follows from the Theorem 2 in [4]. Also, characterization of σ_t^+ and σ_t^- comes from the proof of Theorem 2.

Also, recursive relation of τ_t and δ_t comes from the proof of Proposition 1 in [4].