Real-time Flexibility Feedback for Closed-loop Aggregator and System Operator Coordination

Tongxin Li
CMS, California Institute of Technology
tongxin@caltech.edu

Steven H. Low
CMS, EE, California Institute of Technology
slow@caltech.edu

Adam Wierman
CMS, California Institute of Technology
adamw@caltech.edu

ABSTRACT

Consider a system operator that wishes to optimize its objectives over time subject to operational constraints as well as private constraints of controllable loads managed by an aggregator. In this paper, we design a *real-time* feedback signal for the aggregator to quantify and communicate its available flexibility to the system operator. The proposed feedback signal at each time is the conditional probability of future feasible trajectories that will be enabled by the operator's decision. We show that it is the unique distribution that maximizes a system capacity for flexibility. It allows the system operator to maintain feasibility and enhance future flexibility while optimizing its objectives. We illustrate how the design can be used by the system operator to perform online cost minimization and real-time capacity estimation, while provably satisfying the private constraints of the loads.

KEYWORDS

Aggregate flexibility, real-time closed-loop control, data-driven cost minimization, electric vehicle charging

1 INTRODUCTION

The need to manage the uncertainty and volatility caused by the growing penetration of renewable sources such as wind and solar power has created a desire to increase the ability of the system to provide flexibility via distributed energy resources (DERs) and aggregators have emerged as dominate players for coordinating these loads [3, 4]. The power of aggregators is that they are able to provide coordination among large pools of DERs and then give a single point of contact for independent system operators (ISOs) to call on for flexibility. This enables ISOs to minimize cost, respond to unexpected fluctuations of renewables, and even mitigate failures quickly and reliably.

To realize the potential benefits of aggregators, ISOs need to be able to call on the aggregator via a time-varying signal, e.g., a desired power profile, that satisfies the operational constraints and optimizes a system objective. The signal is then disaggregated by the aggregator in order to determine the behavior of the loads under its control. However, the loads have private constraints on their operation (e.g., satisfying energy demands of electric vehicles before their deadlines). These constraints limit the flexibility available to the aggregator and so the aggregator must also communicate with the ISO by providing a signal that quantifies its available flexibility. This signal is of crucial importance for the ISO when determining the signal it sends to the aggregator, and thus the aggregator and the ISO form a closed-loop control system.

This paper focuses on the design of this closed-loop system and, in particular, the design of the signal quantifying the available flexibility sent from the aggregator to the ISO. The question of how to design the signal providing information on aggregate flexibility of the aggregator to the operator, namely *the flexibility feedback signal*, is complex and has been the subject of significant research over the last decade, e.g., [2, 5, 8, 11, 13, 17, 21, 22, 28]. Any feedback design must balance between a variety of conflicting goals. In particular a design must be:

(1) **Concise.** Given the scale of aggregators and the complexity of the constraints of loads, it is impossible to communicate precise information about every load. Instead, aggregate flexibility feedback must be a concise summary of a system's constraints. Even if it was possible, providing exact information about the constraints of each load governed by the aggregator would not be desirable because the load constraints are typically private. Information conveyed to the ISO must limit the leakage about specific load constraints.

(2) **Informative.** The feedback sent by an aggregator needs to be informative enough that it allows the ISO to achieve operational objectives, e.g., minimize cost, and, most importantly, guarantee the feasibility of the whole system with respect to the private load constraints.

(3) **General.** Any design for a flexibility feedback signal must be general enough to be applicable for a wide variety of controllable loads, e.g., electric vehicles (EVs), heating, ventilation, and air conditioning (HVAC) systems, energy storage units, thermostatically controlled loads, residential loads, and pool pumps. It is impractical to imagine different feedback signals for each load, so the same design must work for all DERs.

The challenge and importance of the design of flexibility feedback signals has led to the emergence of a rich literature. In many cases, the literature focuses on specific classes of controllable loads, such as electric vehicles (EVs) [27], heating, ventilation, and air conditioning (HVAC) systems [12], energy storage units [8], thermostatically controlled loads [13] or residential loads and pool pumps [19, 22]. In the context of these applications, there have been a variety of approaches suggested, e.g., convex geometric approximations [5, 5, 8, 13, 28], scheduling based aggregation [20, 25, 26], and probability-based characterization [19, 22]. These approaches have all yielded some success, especially in terms of quantifying the aggregate flexibility available (we go into more detail about these approaches in the related work section below). However, to this point there are no real-time designs of the coordination between an aggregator and a system operator that achieve the goals laid out above. In particular, the goal of providing a *real-time* feedback signal that is concise and informative has seemed unapproachable

and so nearly all prior work has focused on slower-timescale estimations. In addition to having a flexibility feedback signal that is concise and informative, it is also desirable to have the feedback satisfy the following property:

(4) **Real-time.** The system is time-varying and non-stationary and so it is crucial that (nearly) real-time feedback can be defined and approximated if it is to be used in online feedback-based applications.

The need for real-time information requires that computation of the feedback signal be simple and efficient, which is in direct conflict with assuring generality across wide-ranging applications. In addition, it is highly desirable that the feedback signal be intuitive and interpretable, so that the ISO can use it at a policy level for planning purposes.

Contributions. In this paper we propose a novel design of a flexibility feedback signal that quantifies the flexibility available to an aggregator. We justify our design by proving several desirable properties of the flexibility feedback for real-time feedback-based applications. In particular our flexibility feedback allows the system operator to maintain feasibility and enhance flexibility in real time in an online setting. Finally we demonstrate our design through two example applications: online cost minimization and real time capacity estimation. Our design is conceptually simple, interpretable and we describe two approximations that are efficiently computable in real time. Finally, it is the unique design that attains a certain system capacity for flexibility in an offline setting.

In more detail, we introduce a model of the real-time closed-loop control system formed by a system opeartor and an aggregator. Within this model we define the "optimal" real-time flexibility feedback vector as the solution to an optimization problem that maximizes the entropy of the feedback vector. The use of entropy in this context is novel and we show that entropic maximization has a close relationship to maximization of the system capacity. Further, we justify axiomatically how entropic maximization is fundamentally necessary for providing informative and concise feedback from the aggregator to the operator.

To illustrate applicability of the optimal real-time flexibility feedback vector we propose using two applications: online cost minimization and system capacity estimation. We demonstrate the effectiveness of the flexibility feedback vector in these applications through using real EV charging data from Caltech's ACN-Data dataset [16]. In the case of online cost minimization, we use the flexibility feedback signal in the context of model predictive control and show that the signal is effective even when it is approximated via a data-driven approach based on reinforcement learning. In the case of system capacity estimation, we use the flexibility feedback signal in the context of Monte Carlo estimation and show that the signal is effective even when it is approximated via look-ahead estimation (rather than estimation based on historical data). In both cases we provide provable guarantees that when the aggregator communicates with the system operator via the optimal flexibility feedback signal the private constraints of the loads governed by the aggregator are respected despite the conciseness of the signal communicated to the operator. This work is the first to close the loop and both define a concise measure of aggregate flexibility and

show how it can be used by the system operator to optimize system objectives while respecting the constraints of loads.

Related literature. The growing importance of aggregators for the integration of controllable loads and the challenge of defining and quantifying the flexibility provided by aggregators means that a rich literature on the topic has emerged. Broadly, this work can be separated into three approaches.

Convex geometric approximation. The idea of representing the set of aggregate loads as a virtual battery model dates back to [11, 13]. In [28], flexibility of an aggregation of thermostatically controlled loads (TCLs) was defined as the Minkowski sum of individual polytopes, which is approximated by the homothets of a virtual battery model using linear programming. The recent paper [5] takes a different approach and defines the aggregate flexibility as upper and lower bounds so that each trajectory to be tracked between the bounds is disaggregatable and thus feasible. However, convex geometric approaches cannot be extended to generate real-time flexibility signals because the approximated sets cannot be decomposed along the time axis. In [2], a belief function of setpoints is introduced for real-time control. However, feasibility can only be guaranteed when each setpoint is in the belief set and this may not be the case for systems with memory.

Scheduling algorithm-driven analysis. Scheduling algorithms that enable the aggregation of loads have been studied in depth over the past decade. The authors of [9] introduced a decentralized algorithm with a real-time implementation for EV charging to track a given load profile. The authors of [25] considered the feasibility of matching a given power trajectory and show that causal optimal policies do not exist. In this work, aggregate flexibility was implicitly considered as the set of all feasible power trajectories. Three heuristic causal scheduling policies were compared and the results were extended to aggregation of deferrable loads and storage in [26]. Furthermore, decentralized participation of flexible demand from heat pumps and electric vehicles was addressed in [20]. Notably, the flexibility signals that have emerged from this literature are not general, i.e., the apply to specific policies and DERs.

Probability-based characterization. There is much less work on probabilistic methods. The aggregate flexibility of residential loads was defined based on positive and negative pattern variations by analyzing collective behaviour of aggregate users [22]. A randomized and decentralized control architecture for systems of deferrable loads was proposed in [19], with a linear time-invariant system approximation of the derived aggregate nonlinear model. Flexibility in this work was defined as an estimate of the proportion of loads that are operating. Our work falls into this category, but differs from previous papers in that entropy maximization for a closed-loop control system yield an interpretable signal that can be informative for operator objectives in real-time, as well as guarantee feasibility of the private constraints of loads.

Other approaches. Beyond the works described above, there are many other suggestions for metrics of aggregate flexibility, e.g., graphical-based measures [14] and data-driven approaches [14]. Most of these, and the approaches described above are evaluated at the aggregator level however, and much less attention has been paid to the question of real-time coordination between an ISO and an aggregator that controls decentralized loads.

The assessment and enhancement of aggregate flexibility are often considered independent of the operational objectives and constraints today. For instance, the notion of aggregated flexibility is reported to an ISO participating in a reserve market a day ahead and the scheduling is then conducted the next day after receiving the flexibility representation as defined in [5, 6, 11, 17], with notable exceptions, such as [27], which considered charging and discharging of EV fleets batteries for tracking a sequence of automatic generation control (AGC) signals. However, this approach has several limitations. First, in large-scale systems, knowing the exact states of each load is not realistic. Second, classical flexibility representations often rely on a precise state-transition model on the aggregator's side. Third, traditional ISO market designs, such as a day-ahead energy market, often make use of ex ante estimates of future system states. The forecasts of the future states can sometime be far from reality, because of either an inaccurate model is used, or an uncertain event occurs. In contrast, a real-time energy market [18, 23] provides more robust system control when facing uncertainty in the environment, e.g., from fast-changing renewable resources or human behavioral parameters. This further highlights the need for real-time flexibility feedback, and serves to differentiate the approach in our paper.

Notation and Conventions. We use $\mathbb{P}(\cdot)$ and $\mathbb{E}(\cdot)$ to denote the probability distribution and expectation of random variables. The (discrete) entropy function is denoted by $\mathbb{H}(\cdot)$. To distinguish random variables and their realizations, we follow the convention to denote the former by capital letters (e.g., X) and the latter by lower case letters (e.g., x). Furthermore, we denote the length-t prefix of a vector x by $x_{\leq t} := (x_1, \ldots, x_t)$. Similarly, $x_{<t} := (x_1, \ldots, x_{t-1})$ and $x_{a \to b} := (x_a, \ldots, x_b)$. The concatenation of two vectors x and y is denoted by (x, y). Given two vectors $x, y \in \mathbb{R}^n$, we write $x \leq y$ if $x_i \leq y_i$ for all $i = 1, \ldots, n$. For $x \in \mathbb{R}$, denote $[x]_+ := \max\{0, x\}$.

2 PROBLEM FORMULATION

Consider a load aggregator and a system operator that interact over a discrete time horizon $[T] := \{1, \ldots, T\}$.

2.1 Load aggregator

Let ξ_t denote the *aggregator state* at time t that takes value in a certain set Ω. Let $\xi_{\leq t} := (\xi_s \in \Omega : s = 1, \ldots, t)$ denote the aggregator state trajectory up to time t and $\xi := \xi_{\leq T}$. The aggregator needs to accomplish a certain task over the horizon $[T]$, e.g., delivering energy to a set of electric vehicles (EVs) by their deadlines. To this end, it makes a decision ϕ_t at each time t according to a *disaggregation policy* ϕ. The decision ϕ_t changes the aggregator state ξ_t according to a state transition function which is not essential for our discussion. We hence omit its description and represent the dynamics of the aggregator simply by the state trajectory ξ. Besides accomplishing its task, the decision ϕ_t also produces a system input at time t that will affect a system cost, e.g., the aggregate EV charging rate increases load on the electricity grid. The aggregator has flexibility in its decisions ϕ_t for accomplishing its task and, we assume for this paper, is indifferent to these decisions as long as the task is accomplished by time T. At each time t the system operator sends a *signal* x_t to the aggregator to guide the aggregator's decision ϕ_t towards one that minimizes the system cost. The

signal x_t at time $t \in [T]$ takes value in a discrete set $\mathbb{X} \subseteq \mathbb{R}$.[1] Let $x_{\leq t} := (x_1, \ldots, x_t)$ denote the signal trajectory up to time t and $x := x_{\leq T}$. In general, the aggregator's decision $\phi_t := \phi_t(\xi_t, x_{\leq t})$ is a causal function of aggregator state ξ_t and signal trajectory $x_{\leq t}$ up to time t.[2] We use ϕ both to denote the disaggregation policy or the decision trajectory

$$\phi = (\phi_1(\xi_1, x_1), \ldots, \phi_T(\xi_T, x))$$

depending on the context. We often refer to a pair of disaggregation policy and aggregator state trajectories (ϕ, ξ) as an *aggregator trajectory*. That the aggregator must accomplish its task but is otherwise indifferent to its decisions ϕ can be modeled by the constraints:

$$g_i(x; \phi, \xi) \leq 0, \quad i = 1, \ldots, m, \tag{1}$$

where each g_i is an arbitrary function of ϕ, ξ and x.

The disaggregation policy ϕ can represent a variety of control strategies, such as a scheduling algorithm for EV charging, energy disaggregation, or price signals. We illustrate our model of an aggregator using an EV charging application.

Example 2.1 (Aggregator: EV charging). *Consider an aggregator that is an EV charging facility with n users. Each user j has a private vector $(a(j), d(j), e(j), r(j)) \in \mathbb{R}^4$ where $a(j)$ denotes its arrival (connecting) time; $d(j)$ denotes its departure (disconnecting) time, normalized according to the time indices in $[T]$; $e(j)$ denotes the total energy to be delivered, and $r(j)$ is its peak charging rate. Fix a set of n users with their private vectors $(a(j), d(j), e(j), r(j))$, the aggregator state ξ_t at time $t \in [T]$ is a collection of length-3 vectors $(d(j), e_t(j), r(j) : a(j) \leq t \leq d(j))$ for each EV that has arrived and has not departed by time t. Here $e_t(j)$ is the remaining energy demand of user j at time t. The decision $\phi_t(j)$ is the energy delivered to each user j at time t. A policy ϕ can be well-known scheduling policies such as earliest-deadline-first, least-laxity-first, etc. The aggregator decision $\phi_t(j) \in \mathbb{R}_+$ at each time t updates the state, in particular $e_t(j)$, even though do not explicitly represent the state transition function. The decision also produces an aggregate charging energy $\sum_{j:a(j) \leq t \leq d(j)} \phi_t(j)$ that affects the load on the power grid and operational cost.*

Suppose, in the context of demand response, the system operator (a local utility company, or a building management) sends a signal x_t that is the aggregate energy that can be allocated to EV charging. The aggregator makes charging decisions $\phi_t(j)$ to track the signal x_t received from the system operator as long as they will meet the energy demands of all users before their deadlines. Then the constraints in

---

[1]We assume that the set $\mathbb{X}$ is discrete only for simplicity of presentation. Our results, for example, the definition of optimal flexibility feedback (Definition 3.1), Theorem 1 can be extended to continuous space using a density function as the flexibility feedback, changing the summations to integrals, replacing the discrete entropy functions by differential entropy functions, and redefining the system capacity $\log |\mathcal{S}(\phi, \xi)|$ as the volume of the space consisting of feasible signal trajectories.

[2]The main results can be easily extended to allow for non-causal policies.

(1) *include the following constraints on the charging decisions:*

$$\phi_t(j) = 0 \, , \ t < a(j), j = 1, \ldots, n,$$

$$\phi_t(j) = 0 \, , \ t > d(j), j = 1, \ldots, n,$$

$$\sum_{j=1}^{n} \phi_t(j) = x_t, t = 1, \ldots, T, \tag{2a}$$

$$\sum_{t=1}^{T} \phi_t(j) = e(j), j = 1, \ldots, n, \tag{2b}$$

$$0 \le \phi_t(j) \le r(j), t = 1, \ldots, T \tag{2c}$$

where constraint (2a) *ensures that the aggregator decision ϕ_t tracks the signal x_t at each time $t \in [T]$, the constraint (2b) guarantees that EV j's energy demand is satisfied, and the other constraints say that the aggregator cannot charge an EV before its arrival, after its departure, or at a rate that exceeds its limit.*

2.2 System operator

As Example 2.1 illustrates, the aggregator decisions ϕ_t produce a system input that affects system operation. The goal of the system operator is to compute a signal x_t at time $t \in [T]$ to guide the aggregator's decisions ϕ_t so as to minimize the system cost given by a *cost function* $f : \mathbb{X}^T \to \mathbb{R}$. The signal trajectory x must satisfy certain operational constraints, parameterized by an *environmental parameter* ζ:

$$h_i(x; \zeta) \le 0, \ i = 1, \ldots, k. \tag{3}$$

Example 2.2 (Operational constraints). *Suppose x_t represents the total load of an EV charging facility, or on a power system. If the operator performs peak shaving, then* (3) *may be:*

$$x_t \le \overline{\gamma}, \ t = 1, \ldots, T. \tag{4}$$

If the operator limits ramp rates, then (3) *may be:*

$$|x_{t+1} - x_t| \le \varepsilon, \ t = 1, \ldots, T. \tag{5}$$

Example 2.3 (Cost function). *Suppose the electricity cost at each time $t \in [T]$ is a function $f_t : \mathbb{X} \to \mathbb{R}_+$. Then the total electricity cost is $f(x) := \sum_{t=1}^{T} f_t(x_t)$.*

The goal of the system operator is to choose the signal x so as to solve:

$$\min_x f(x) \tag{6a}$$

$$\text{subject to } g_i(x; \phi, \xi) \le 0, \ i = 1, \ldots, m, \tag{6b}$$

$$h_i(x; \zeta) \le 0, \ i = 1, \ldots, k. \tag{6c}$$

i.e., the operator wishes to minimize its cost f subject to its operational constraints (6c) while the load aggregator needs to fulfill its obligations in the form of constraints (6b). This is an offline problem that involves global information at all times $t \in [T]$. The challenge is that the constraints (6b) are private to the aggregator. It is impractical for the aggregator to communicate the constraint functions $g_i(x; \phi, \xi)$ to the operator because of privacy concerns or computational effort, and because in an online setting, even the aggregator will not know all the constraints at each time t that involve future information, e.g., future EV arrivals in Example 2.1.

Remark 1. For simplicity, we describe our model in an offline setting where the cost and the constraints (e.g., see (2b) in Example 2.1) in the optimizatoin problem (6) are expressed in terms of the entire trajectories $(x; \phi, \xi)$. All functions defined in this paper, however, are *causal* in that they depend *only on local information available at time t*. Hence these functions are designed for solving an online version of the offline problem (6).

2.3 Online feedback-based solution

We explore a solution where the system operator and the aggregator jointly solve an online version of (6) in a closed loop in real time, as illustrated in Figure 1. Our approach does not require the aggregator to know the system operator's optimization problem (6), but only the signal x_t at each time t from the operator. It does not require the system operator to know the aggregator constraints (6b), but only a feedback signal p_t (to be designed) from the aggregator. The system operator generates its signal x_t using a causal function π_t and the aggregator generates its feedback p_t using a causal function ψ_t. By an "online feedback" solution, we mean that these functions (π_t, ψ_t) use only information available locally at time t.

Specifically, our approach proceeds as follows. At each time t, the aggregator computes a length-$|\mathbb{X}|$ vector

$$p_t(\cdot | x_{<t}; \xi_t) \ = \ \psi_t(x_{<t}; \xi_t) \ =: \ \psi_t(x_{<t}) \tag{7a}$$

based on its current state ξ_t and previously received signal trajectory $x_{<t} = (x_1, \ldots, x_{t-1})$, and sends it to the system operator. We will omit ξ_t in the notation when it is not essential to our discussion and simplify the probability vector as p_t.[3] The system operator then computes a (possibly random) signal

$$x_t \ = \ \pi_t(p_t; \zeta) \ =: \ \pi_t(p_t) \tag{7b}$$

based on the aggregator feedback p_t and sends it to the aggregator. We will omit ζ in the notation when it is not essential to our discussion. The aggregator makes its decision $\phi_t(\xi_t, x_{\le t})$. It then computes the next feedback p_{t+1} and the cycle repeats.

The operator chooses its signal x_t in order to solve the time-t problem in an online version of (6), so the function π_t denotes the mapping from the aggregator feedback p_t to an optimal solution of the time-t problem. See Section 4 for an example.

The focus of this paper is to propose an aggregator feedback ψ_t in (7a) that quantifies its future flexibility that will be enabled by an operator decision x_t. The feedback p_t therefore is a surrogate for the aggregator constraints (6b) to guide the operator's decision. Specifically, define the set of all *feasible signal trajectories* for the aggregator as:

$$\mathcal{S}(\phi, \xi) := \left\{ x \in \mathbb{X}^T : x \text{ satisfies (6b)} \right\}.$$

Throughout, we assume that $\mathcal{S}(\phi, \xi)$ is non-empty. We propose that the aggregator function $\psi_t(x_{<t}; \xi_t)$ computes the conditional probabilities of future signal trajectories $x_{>t} := (x_{t+1}, \ldots, x_T)$ that satisfy the aggregator constraints (6b), as a function of the operator's signal choice x_t, conditioned on the signal trajectory

---

[3]Note that in (9b) we slightly abuse the notation and use p_t to denote a conditional distribution. This is only for computational purposes and the information sent from an aggregator to an operator at time $t \in [T]$ is still a length-$|\mathbb{X}|$ probability vector, conditioned on fixed $x_{<t}$.
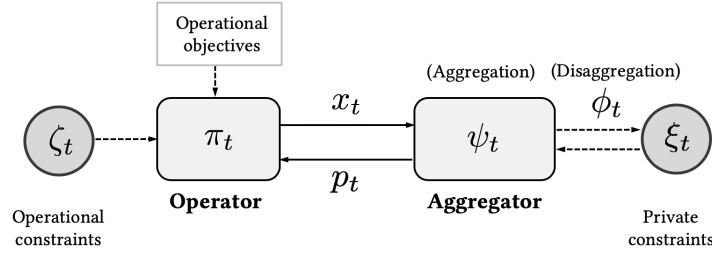
Figure 1: A feedback control approach for solving an online version of (6).

$x_{<t} := (x_1, \ldots, x_{t-1})$ up to time $t-1$. Formally, let $\mathcal{P}$ denote the probability simplex:

$$\mathcal{P} := \left\{ p \in \mathbb{R}^{|\mathbb{X}|} : p(x) \geq 0, x \in \mathbb{X}; \sum_{x \in \mathbb{X}} p(x) = 1 \right\}.$$

Fix any aggregator trajectory (ϕ, ξ). Then the aggregator function $\psi_t : \mathbb{X}^{t-1} \times \Omega \to \mathcal{P}$ at each time t is: $p_t = \psi_t(x_{<t}; \xi_t)$ such that for each $x_t \in \mathbb{X}$,

$$\psi_t(x_{<t}; \xi_t) := p(\cdot|(x_{<t})) \in \mathcal{P}. \tag{8}$$

We refer to p_t as *flexibility feedback* sent at time $t \in [T]$ from the aggregator to the system operator. Given current aggregator state ξ_t and signal trajectory $x_{<t}$, the conditional probability $\psi_t(x_{<t}; \xi_t)$ depends not just on the operator decision x_t, but also on the future evolution of the aggregator state ξ_t. In this paper, we do not fully specify the details of the dynamical process ξ_t. For different applications, ξ_t may evolve according to different state transition functions, possibly with stochastic inputs. These details will determine the value of the flexibility feedback $p_t = \psi_t(x_{<t}; \xi_t)$ defined in (8).

In this sense, (8) does not specify a specific aggregator function ψ_t, but a class of possible functions ψ_t. Every function in this collection is *causal* in that it depends only on information available to the aggregator at time t. In contrast to most aggregate flexibility notions in the literature [5, 8, 11, 13, 17, 21, 22, 28], the flexibility feedback here is specifically designed for an online feedback control setting.

3 OPTIMAL FLEXIBILITY FEEDBACK

In this section we propose a specific function ψ_t in the class defined by (8) for computing aggregator feedback to quantify its future flexibility. We will justify our proposal by showing that the proposed ψ_t has several desirable properties for solving an online version of (6) using the real-time feedback-based approach (7).

3.1 Definition

The intuition behind our proposal is that the conditional probability $p_t(x_t) := p_t(x_t|x_{<t})$ measures the resulting future flexibility of the aggregator if the system operator chooses x_t as the signal at time t, given the signal trajectory up to time $t-1$. The sum of the conditional entropy of p_t thus is a measure of how informative p_t is. This suggests choosing a conditional distribution p_t that maximizes its conditional entropy. Fix any aggregator trajectory

(ϕ, ξ). Consider the optimization problem:

$$F(\phi, \xi) := \max_{p_1, \ldots, p_T} \sum_{t=1}^{T} \mathbb{H}(X_t|X_{<t}) \text{ subject to } X \in \mathcal{S}(\phi, \xi) \quad (9a)$$

where the variables are conditional distributions:

$$p_t := p_t(\cdot|\cdot) := \mathbb{P}_{X_t|X_{<t}}(\cdot|\cdot), \qquad t \in [T] \tag{9b}$$

$X \in \mathbb{X}^T$ is a random variable distributed according to the joint distribution $\prod_{t=1}^{T} p_t$ and $\mathbb{H}(X_t|X_{<t})$ is the conditional entropy of p_t defined as:

$$\mathbb{H}(X_t|X_{<t}) := \sum_{x_1, \ldots, x_t \in \mathbb{X}} \left(-\prod_{\ell=1}^{t} p_\ell(x_\ell|x_{<\ell}) \right) \log p_t(x_t|x_{<t}). \tag{9c}$$

By definition, a quantity conditioned on "$x_{<1}$" means an unconditional quantity, so in the above, $\mathbb{H}(X_1|X_{<1}) := \mathbb{H}(X_1) := \mathbb{H}(p_1)$. The chain rule shows that $\sum_{t=1}^{T} \mathbb{H}(X_t|X_{<t}) = \mathbb{H}(X)$. Hence (9) can be interpreted as maximizing the entropy $\mathbb{H}(X)$ of a random trajectory X sampled according to the joint distribution $\prod_{t=1}^{T} p_t$, conditioned on X satisfying (1), where the maximization is over the collection of conditional distributions $(p_1, \ldots, p_T)$. We provide in Section 3.3 an axiomatic justification of maximizing the entropy $\mathbb{H}(X)$ of the signal trajectory X in (9a).

DEFINITION 3.1 (OPTIMAL FLEXIBILITY FEEDBACK). *Fix any aggregator trajectory* (ϕ, ξ). *The flexibility feedback* $p_t^* = \psi_t^*(x_{<t}; \xi_t)$ *for* $t \in [T]$ *is called the* **optimal flexibility feedback** *if* $(p_1^*, \ldots, p_T^*)$ *is the unique optimal solution of* (9).

Remark 2. Even though the optimization problem (9) involves variables p_t for the entire time horizon $[T]$, the individual variables p_t in (9b) are conditional probabilities that depend only on information available to the aggregator at times t. Therefore the optimal flexibility feedback ψ_t^* in Definition 3.1 is indeed causal and in the class of functions ψ_t^* defined in (8). The existence and uniqueness of p_t^* is guaranteed by Theorem 1 below, which also implies that ψ_t^* is unique. $\square$

We demonstrate Definition 3.1 using a toy example.

Example 3.1 (Optimal flexibility feedback p^*). Consider the following instance of Example 2.1. Suppose the number of charging time slots is $T = 3$ and there is one customer, whose private vector is $(1, 3, 1, 1)$ and possible energy levels are 0 (kWh) and 1 (kWh), i.e., $\mathbb{X} \equiv \{0, 1\}$. Since there is only one EV, the scheduling algorithm ϕ (disaggregation policy) assigns all power to this single EV. For

this particular choices of ξ and ϕ, the set of feasible trajectories is $S(\phi, \xi) = \{(0,0,1), (0,1,0), (1,0,0)\}$, shown in Figure 2 with the corresponding optimal conditional distributions given by (9).

$$p_1^*(x_1) = \begin{cases} \frac{2}{3} & \text{if } x_1 = 0 \\ \frac{1}{3} & \text{if } x_1 = 1 \end{cases}$$

$$p_2^*(x_2|x_1) = \begin{cases} \frac{1}{2}, & \text{if } (x_1, x_2) \in \{(0,0),(0,1)\} \\ 1, & \text{if } (x_1, x_2) = (1,0) \\ 0, & \text{otherwise} \end{cases}$$

$$p_3^*(x_3|x_1, x_2) = \begin{cases} 1 & \text{if } (x_1, x_2, x_3) \in \{(0,0,1),(0,1,0),(1,0,0)\} \\ 0 & \text{otherwise} \end{cases}$$
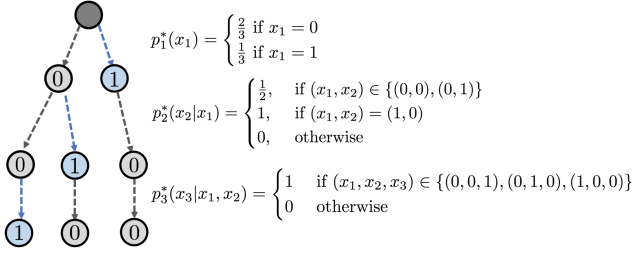
Figure 2: Feasible trajectories of power signals and the computed optimal flexibility feedback in Example 3.1.

3.2 Properties of p_t^*

We now show that the proposed optimal flexibility feedback p_t^* has several desirable properties. We start by computing p_t^* explicitly. Fix any aggregator trajectory (ϕ, ξ). Given any signal trajectory $x_{\leq t}$, define the set of *subsequent* feasible trajectories as:

$$S(\phi, \xi | x_{\leq t}) := \left\{ x_{>t} \in \mathbb{X}^{T-t} : g_i(\phi, \xi, x) \leq 0, \ \forall i = 1, \ldots, m \right\}. \tag{10}$$

where $x := (x_{\leq t}, x_{>t})$. The size $|S(\phi, \xi | x_{\leq t})|$ of the set of subsequent feasible trajectories is a measure of future flexibility, conditioned on $x_{\leq t}$. Our first result justifies our calling p_t^* the optimal flexibility feedback: p_t^* is a measure of the future flexibility that will be enabled by the operator's signal x_t and it attains a measure of system capacity for flexibility (see Remark 3 below). By definition, $S(\phi, \xi | x_{<1}) := S(\phi, \xi)$ and $p_1^*(x_1 | x_{<1}) := p_1^*(x_1)$.

THEOREM 1. *The optimal flexibility feedback p_t^* is given by*

$$p_t^*(x_t | x_{<t}) = \frac{|S(\phi, \xi | (x_{<t}, x_t))|}{|S(\phi, \xi | x_{<t})|}, \quad \forall (x_{<t}, x_t) \in \mathbb{X}^t. \tag{11}$$

for $t \in [T]$. Moreover, the optimal value $F(\phi, \xi)$ of (9) is equal to $\log |S(\phi, \xi)|$.

PROOF. We prove the statement by induction on T. It is straightforward to verify the results when $T = 1$. We suppose the theorem is true when $T = m$. Suppose $T = m + 1$. Let

$$F(\phi, \xi | x_1) := \max_{p_2, \ldots, p_T} \sum_{t=2}^T \mathbb{H}(X_t | X_{<t})$$

denote the optimal value corresponding to the time horizon $t \in [T] \backslash \{1\}$, conditioning on x_1. We have

$$F(\phi, \xi) = \max_{p_1} \sum_{x_1 \in \mathbb{X}} p_1(x_1) F(\phi, \xi | x_1) + \mathbb{H}(p_1).$$

By the induction hypothesis, $F(\phi, \xi | x_1) = \log |S(\phi, \xi | x_1)|$. Therefore,

$$F(\phi, \xi) = \max_{p_1} \sum_{x_1 \in \mathbb{X}} p_1(x_1) \log |S(\phi, \xi | x_1)| + \mathbb{H}(p_1)$$

$$= \max_{p_1} \sum_{x_1 \in \mathbb{X}} p_1(x_1) \log \left(\frac{|S(\phi, \xi | x_1)|}{p_1(x_1)} \right)$$

whose optimizer p_1^* satisfies (11) and we get $F(\phi, \xi) = \log |S(\phi, \xi)|$. The theorem follows by finding the optimal conditional distributions $p_2^*, \ldots, p_T^*$ inductively. □

Given the unique optimal flexibility feedback $(p_1^*, \ldots, p_T^*)$ guaranteed by Theorem 1, let $q^*(x) = \prod_{t=1}^T p_t^*(x_t | x_{<t})$ denote the joint distribution of the signal trajectory x. Then (11) implies that the joint distribution q^* is the uniform distribution over the set $S(\phi, \xi)$ of all feasible trajectories:

$$q^*(x) := \begin{cases} 1/|S(\phi, \xi)| & \text{if } x \in S(\phi, \xi) \\ 0 & \text{otherwise} \end{cases}. \tag{12}$$

Remark 3 (System capacity $F(\phi, \xi)$). Fix any aggregator trajectory (ϕ, ξ). The size $|S(\phi, \xi)|$ is a measure of flexibility inherent in the aggregator. We will hence call $\log |S(\phi, \xi)|$ the *system capacity*. Theorem 1 then says that the optimal value of (9) is the system capacity, $F(\phi, \xi) = \log |S(\phi, \xi)|$. Moreover the optimal flexibility feedback $(p_1^*, \ldots, p_T^*)$ is the unique collection of conditional distributions that attains the system capacity in (9). This is intuitive since the entropy of a random trajectory x in $S(\phi, \xi)$ is maximized by the uniform distribution q^* in (12) induced by the conditional distributions $(p_1^*, \ldots, p_T^*)$. □

Theorem 1 directly implies the following important properties of the optimal flexibility feedback.

COROLLARY 2 (FEASIBILITY AND FLEXIBILITY). *Let $p_t^* = p_t^*(\cdot | x_{<t})$ be the optimal flexibility feedback at each time $t \in [T]$.*

(1) *For any signal trajectory $x = (x_1, \ldots, x_T)$, if*

$$p_t^*(x_t | x_{<t}) > 0 \quad \text{for all } t \in [T]$$

then $x \in S(\phi, \xi)$.

(2) *For all $x_t, x_t' \in \mathbb{X}$ at each time t, if*

$$p_t^*(x_t | x_{<t}) \geq p_t^*(x_t' | x_{<t})$$

then $|S(\phi, \xi | (x_{<t}, x_t))| \geq |S(\phi, \xi | (x_{<t}, x_t'))|$.

We elaborate on the implication of Corollary 2 on our online feedback-based solution approach.

Remark 4 (Feasibility and flexibility). Corollary 2 says that the proposed optimal flexibility feedback p_t^* provides the right information for the system operator to choose its signal x_t at time t. Specifically, the first statement of the corollary says that if the operator always chooses a signal x_t with positive conditional probability $p_t^*(x_t) > 0$ for each time t, then the resulting signal trajectory is guaranteed to be feasible, $x \in S(\phi, \xi)$, i.e., the system will remain feasible at *every* time t along the way.

Moreover, according to the second statement of the corollary, if the system operator chooses a signal x_t with a larger $p_t^*(x_t)$ value at time t, then the system will be more flexible going forward than if it had chosen another signal x_t' with a smaller $p_t^*(x_t')$ value, in the

sense that there are more feasible trajectories in $\mathcal{S}(\phi, \xi | (x_{<t}, x_t))$ going forward. □

As noted in Remark 1, despite characterizations that involve the whole trajectory (x, ϕ, ξ), such as $x \in \mathcal{S}(\phi, \xi)$, these are *online* properties. This guarantees the feasibility of the online closed-loop control system depicted in Figure 1, and confirms the suitability of p_t^* for online applications.

3.3 Axiomatic justification of (9)

As explained in Remark 4, the optimal flexibility feedback p_t^* quantifies succinctly for the system operator the future flexibility of the aggregator that will be enabled by the operator's choice of next signal x_t. Intuitively, the system has "more flexibility" at time t if the distribution $p_t(\cdot | x_{<t})$ is "more uniform". This view suggests using an entropic measure to quantify flexibility, such as the cost function of the optimization problem (9) that underlies our proposed flexibility feedback. In this subsection we justify this intuition using an axiomatic argument.

Fix any aggregator trajectory (ϕ, ξ). Consider a flexibility metric as a function of any flexibility feedback $p \in \{p_1, \ldots, p_T\}$. Recall that p is a conditional distribution. For any p, let $F(p)$ represent a candidate metric for quantifying aggregate flexibility. Consider any time slots $\tau \in [T]$, the metric should also be able to provide a value, given the marginal distributions $\overline{p}_t := \sum_{\mathbf{x}_{<t}} p_t(\cdot | x_{<t}) \prod_{\tau < t} p_\tau(x_\tau | x_{<\tau})$.

We require the metric F to satisfy several conditions (axioms):

(1) *Continuity*: $F(\overline{p}_t)$ is a continuous function of $\overline{p}_t$, $t \in [T]$.
(2) *(Strong) additivity*: $F(q) = \sum_{t=1}^T F(p_t)$ if $q := \prod_{t=1}^T p_t$.
(3) *Subadditivity*: $F(q_{t,t'}) \leq F(\overline{p}_t) + F(\overline{p}_{t'})$ where $\overline{p}_t, \overline{p}_{t'}$ are marginal distributions corresponding to time slots t and t' and $q_{t,t'}$ is their joint distribution.
(4) *Symmetry*: $F(q_{t,t'}) = F(q_{t',t})$ where $q_{t,t'}$ and $q_{t',t}$ are joint distributions of time slots t and t'.
(5) *Expansibility*: $F(\overline{p}_t') = F(\overline{p}_t')$ for all $\overline{p}_t$, $t \in [T]$ where $\overline{p}_t' = (\overline{p}_t, 0)$, i.e., concatenate a zero entry to $\overline{p}_t$.

Additivity is useful because the tracking of a random signal trajectory $x := (x_1, \ldots, x_T)$ can then be decomposed using the chain rule into sub-problems of tracking each signal x_t at time t, conditioned on previous signal trajectory $x_{<t}$. Subadditivity is motivated by the property that fixing a signal x_t may restrict the choice of feasible signals $x_{t'}$ since the signals $x_1, \ldots, x_T$ may be correlated. This means that measuring the joint distribution of $(x_t, x_{t'})$ gives lower flexibility than measuring the coordinates x_t and $x_{t'}$ independently. For symmetry, the permutation of components in the distribution $\overline{p}_t$ does not change $F(\overline{p}_t)$ since the switch of positions does not affect the underlying distribution. Expansibility is natural since adding a new component that equals to zero means x_t can never choose a certain power level. So the aggregate flexibility will not change.

These five conditions imply that the flexibility metric $F(\overline{p}_t)$ (for all $t \in [T]$) must be an entropy function:

$$\mathbb{H}(\overline{p}_t) := \sum_{x \in \mathbb{X}} \sum_{\mathbf{x}_{<t}} p_t(x | \mathbf{x}_{<t}) \prod_{\tau < t} p_\tau(x_\tau | x_{<\tau})$$
$$\cdot \log \left(\frac{1}{\sum_{\mathbf{x}_{<t}} p_t(x | \mathbf{x}_{<t}) \prod_{\tau < t} p_\tau(x_\tau | x_{<\tau})} \right)$$

Data: Sequential cost functions $f_1, \ldots, f_T$ and states $\xi_1, \ldots, \xi_T$
Result: Total cost $\sum_{t=1}^T f_t(x_t)$
for $t \in [T]$ **do**
   Generate flexibility feedback:
$$\overline{p}_t = \psi_t^{\text{SAC}}(x_{<t}; \xi_t)$$
   Generate control signal and compute cost:
$$x_t = \pi_t^{\text{RHC}}(\overline{p}_t)$$
$$\text{cost} = \text{cost} + f_t(x_t)$$
   Update state:
   $\xi_{t+1}(x_t; \xi_t) \longleftarrow \xi_t$
end
return cost

Algorithm 1: The RHC scheme for online cost minimization.

up to multiplicative factors and $F(p_t)$ is the conditional entropy of p_t. This is a classical result about entropy; see [1, 7].

The results in this section justify the design of using the unique optimal solution of (9) as our flexibility feedback p_t^*. The design attains the system capacity $F(\phi, \xi)$. Moreover it characterizes the aggregate flexibility in real-time and allows a decomposition (see Section 5.1 for details) of aggregate flexibility over t via

$$\sum_{t=1}^T \mathbb{H}(p_t^*) = F(\phi, \xi).$$

We use this decomposition in Section 4 for online cost minimization where p_t^* is used as a penalty in a RHC-based online algorithm. We also use it in Section 5 for estimating the system capacity $F(\phi, \xi)$ empirically using a Monte Carlo method. Finally, computing the optimal flexibility feedback is demanding. We provide two approximations for p_t^*, one for the case where sufficient historical data is available and the other when it is not. The first is a data-driven approach using reinforcement learning (Section 4.2) and the second is a look-ahead approximation (Section 5.2).

4 ONLINE COST MINIMIZATION

Consider the cost minimization problem introduced in Example 2.3. In this setting, the operator seeks to minimize the cost in an online manner, i.e., at time t the operator only knows the objective functions $f_1, \ldots, f_t$ and the flexibility feedback $p_1, \ldots, p_t$.

We first describe a receding horizon control scheme for the operator that, given the flexibility feedback and the objective functions, allows the operator to compute the signals $x_1, \ldots, x_T$. Then, we introduce a deep reinforcement learning-based approach for the aggregator to compute an approximation of the optimal flexibility feedback. Finally, we illustrate our method with simulations.

4.1 Operator: Receding horizon control

The task of the operator is to, given the optimal flexibility feedback, generate signals $x_1, \ldots, x_T$ that are always feasible with respect to both the sets of private and operational constraints *and* that minimize cost. For the objective of cost minimization, we propose an approach that uses receding horizon control (RHC) to achieve this in an adaptive, online manner – see Algorithm 1.

We focus on a specific class of constraints and assume the operational constraints $h_i(\zeta, x) \leq 0$, $i = 1, \ldots, k$ can be decoupled as ($t \in [T]$):

$$h_i^{(t)}(\zeta, x_t) \leq 0, \; i = 1, \ldots, k_t. \tag{13}$$

First, we consider the following equivalent offline optimization of (6). Recall q^* from (12).

$$\min \sum_{t=1}^{T} f_t(x_t) - \beta \log q^*(x) \tag{14}$$

$$\text{subject to } h_i(\zeta, x) \leq 0, \; i = 1, \ldots, k,$$

where $\beta > 0$ is a tuning parameter. Decomposing the joint distribution $q^*(x) = \prod_{t=1}^{T} p_t^*(x_t | x_{<t})$ by the optimal conditional distributions given by (9), the objective function (14) becomes

$$\sum_{t=1}^{T} f_t(x_t) - \beta \log \left(\prod_{t=1}^{T} p_t^*(x_t | x_{<t}) \right)$$

$$= \sum_{t=1}^{T} \left(f_t(x_t) - \beta \log p_t^*(x_t | x_{<t}) \right). \tag{15}$$

Eq. (15) motivates the following RHC-based operator function at time $t \in [T]$, which includes the flexibility feedback $p_t(\cdot | x_{<t})$ as a penalty term in a greedy minimization:

$$\pi_t^{\text{RHC}} := \underset{x \in \mathbb{X}}{\arg\min} \; f_t(x) - \beta \log p_t(x | x_{<t}) \tag{16}$$

$$\text{subject to (13).} \tag{17}$$

Crucially, the following shows that feasibility is guaranteed when the flexibility feedback is optimal.

COROLLARY 3. *Suppose at each time $t \in [T]$, the optimal flexibility feedback $p_t^*(\cdot | x_{<t})$ is sent to an operator constrained by (13). Then, the trajectory $x = (x_1, \ldots, x_T)$ generated by the RHC-based operator function π_t^{RHC} in (16) is always feasible, i.e., $x \in \mathcal{S}(\phi, \xi)$.*

PROOF. Applying Theorem 2, it suffices to show that the signal $x_t = \pi_t^{\text{RHC}}(p_t^*(\cdot | x_{<t}))$ generated at time t satisfies $p_t^*(x_t | x_{<t}) > 0$. Suppose not, then there is a control signal $x^* = \pi_t^{\text{RHC}}(p_t^*(\cdot | x_{<t}))$ such that $p_t^*(x | x_{<t}) = 0$ for some t implies the objective in (16) becomes positive infinity. Our assumption $\mathcal{S}(\phi, \xi) \neq \emptyset$ implies that $p_t^*(\cdot | x_{<t})$ is not an all-zero vector. Therefore, x is not the optimal solution of (16), yielding a contradiction. □

4.2 Aggregator: Data-driven approximation of the optimal flexibility feedback

As we have already noted, computing the optimal flexibility feedback is computationally intensive. Thus, instead of computing it precisely, it is desirable to approximate it. For the case of online cost minimization, it is possible to take a data-driven approach. In particular, we propose the use of reinforcement learning to learn a function $\psi_t : \Omega \to \mathcal{P}$ that outputs the estimated flexibility feedback $\bar{p}_t$ given the current system state ξ_t. Note that we do not directly learn the disaggregation of x_t, which would have too large an action space. Instead, we fix a specific scheduling algorithm and learn the feedback vectors directly. This is another benefit of the concise representation of the feedback vectors.

More specifically, we train an agent function ψ_t using soft actor-critic (SAC) [10], with the following generic reward function $r : \Omega^t \times \mathbb{X}^t \times \mathcal{P} \to \mathbb{R}$:

$$r(\xi_{\leq t}, \mathbf{x}_{\leq t}, \mathbf{p}_t) = \mathbb{H}(\mathbf{p}_t) - \sum_{i=1}^{m} c_i \min_{\mathbf{x}_{>t}} [g_i(\mathbf{x}; \phi, \xi_{\leq t})]_+. \tag{18}$$

The first term maximizes the entropy of the flexibility feedback vector, as a heuristic for the objective in (9a). The second term penalizes the choice of x_t that leads to an infeasible trajectory. Note that the reward function is independent of the price functions. We provide more details in Appendix A. We next demonstrate in simulations that feeding back to the operator the approximate optimal flexibility obtained from reinforcement learning is sufficient for achieving the desirable properties proven in Section 3.

4.3 Experiments

In the following, we show our experimental results for online EV charging, using real EV charging data ACN-Data [16], which is a dataset collected from adaptive EV charging networks (ACNs) at Caltech and JPL. The detailed choices of SAC parameters and the design of the reward function for the SAC approach are presented in Appendix A.
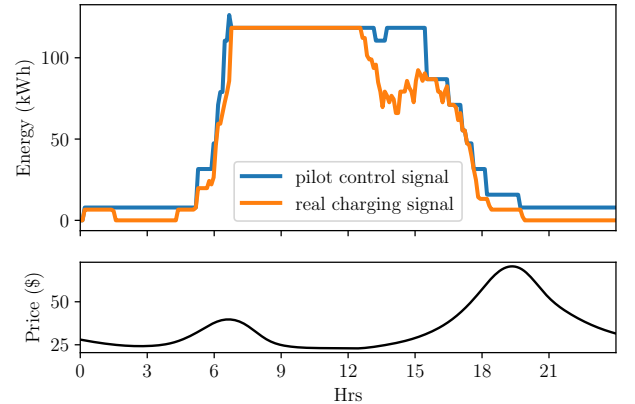

Figure 3: Pilot control signals and real energy allocated to EVs generated by Algorithm 1.

Charging curves. In Figure 3, pilot control and real energy signals are shown. The agent is trained on data collected at Caltech from Nov. 1, 2018 to Dec. 1, 2019 with linear price functions $f_t = 1 - t/24$, where $t \in [0, 24]$ (unit: Hrs) is the time index and tested on Dec. 18, 2019 for JPL with average LMPs on the CAISO (California Independent System Operator) day-ahead market in 2016, shown on the bottom. The scheduling policy is fixed to be LLF (see Appendix D for more details). The set of power levels $\mathbb{X}$ is a discrete set that contains 60 distinct power levels from 0 kWh to 360 kWh. We use tuning parameter $\beta = 4000$. The pilot control signals are optimal solutions of (16), which are always bounded from below by the real charging signals, representing the aggregate charging rates $\sum_i \phi_i(t)$ for $t \in [0, 24]$. The figure highlights that, with a suitable choice of tuning parameter, the operator is able to schedule charging at

time slots where prices are lower and avoid charging at the peak of prices, as desired. Note that the operational constraints used in this experiments is $x_t \leq 150$ (kWh) for every $t \in [T]$ and the learned flexibility feedback is able to automatically flatten the charging curve within this range, without explicitly knowing it.

5 SYSTEM CAPACITY $F(\phi, \xi)$ ESTIMATION

In addition to minimizing cost, another important goal of the operator is to quantify the amount of flexibility available at each time. This is crucial for purposes of ensuring the ability to respond to failures and planning for capacity investment. However, given that the private constraints of loads are not visible to the operator, such estimation is challenging. Further, measuring the exact size of $\mathcal{S}(\phi, \xi)$ is intractable even if such constraints were visible, since the subset in $\mathbb{R}^T$ specified by inequalities (1) can be non-convex and even computing the volume of a convex body can be a hard problem [24]. Furthermore, since a system's states are time-variant, the flexibility of the system also changes over time.

In this section, we illustrate how the optimal flexibility feedback can be used to estimate the system capacity $F(\phi, \xi)$. To this end, we propose an empirical estimation of the system capacity that uses an approximation of the optimal flexibility feedback and demonstrate our method using a case study of EV charging.

5.1 Operator: Monte Carlo estimation

The task of the operator is to, given the optimal flexibility feedback, estimate the system capacity $F(\phi, \xi)$ while also generating signals $x_1, \ldots, x_T$ that are always feasible with respect to both the operational constraints (6c) and the private aggregator constraints (6b). The approach we propose is an empirical estimation of the system capacity using Monte Carlo estimation. In particular, we consider

$$\mu_N(\phi, \xi) := \frac{1}{N} \sum_{\ell=1}^{N} \sum_{t=1}^{T} \mathbb{H}\left(p_t(\cdot|x_{<t}(\ell))\right), \tag{19}$$

where the summation is over T discrete time slots and N trajectories. For each, the corresponding entropy function computes the entropy of the flexibility feedback vector p_t conditioned on the generated signals $x_{<t}(\ell)$ at each time $t \in [T]$:

$$\mathbb{H}\left(p_t(\cdot|x_{<t}(\ell))\right) := -\sum_{x \in \mathbb{X}} p_t\left(x|x_{<t}(\ell)\right) \log p_t\left(x|x_{<t}(\ell)\right).$$

The goal of this approach is that, with suitable choices of operator functions π_t, when the number N of sampled trajectories becomes large, the approximation converges to the system capacity $F(\phi, \xi)$. To see why, suppose at each time $t \in [T]$, the operation π_t^{OPT} is a stochastic function that samples a signal X_t according to the optimal flexibility feedback p_t^*, i.e., for all $t \in [T]$ and $x_t \in \mathbb{X}$,

$$\mathbb{P}\left(\pi_t^{\text{OPT}}\left(p_t^*(\cdot|x_{<t}) = x_t\right)\right) = p_t^*(x_t|x_{<t}).$$

In this context, the theorem below shows that we obtain an estimate of the system capacity $F(\phi, \xi)$ using Monte Carlo estimation.

THEOREM 4. *If the N trajectories* $\{(x_1(\ell), \ldots, x_T(\ell))\}_{\ell=1}^{N}$ *are generated i.i.d. by* $\{\pi_1^{\text{OPT}}, \ldots, \pi_T^{\text{OPT}}\}$, *then the empirical estimate in* (19) *converges to the system capacity almost surely, i.e.,*

$$\mu_N(\phi, \xi) \xrightarrow{a.s.} F(\phi, \xi) \text{ as } N \to \infty.$$

Note that, in addition to providing a method for estimating the system capacity, the theorem also validates that the entropy of the flexibility feedback sent each time reflects the system's current flexibility. This indicates that, for instance, if the feedback vector is a uniform distribution on $\mathbb{X}$, then the system has maximal flexibility.

PROOF OF THEOREM 4. Suppose N trajectories $\{x(1), \ldots, x(N)\}$ are sampled i.i.d. according to the optimal flexibility feedback. Equivalently, for all $\ell = 1, \ldots, N$, the entropy of the optimal flexibility feedback $p_t^*(x_{<t}(\ell))$ can be written as the following conditional entropy $\mathbb{H}(p_t(\cdot|x_{<t}(\ell))) = \mathbb{H}(X_t|X_{<t} = x_{<t})$, where each $X_t \in \mathbb{X}$ is a random signal drawn according to $p_t^*(x_{<t}(\ell))$. We claim that, if the random power signal X_t is sampled according to $p_t^*(\cdot|x_{<t})$ conditioned on previous power signals $x_{<t} = (x_1, \ldots, x_{t-1})$ for all $t \in [T]$, then the accumulated flexibility over $t \in [T]$ is equal to the system capacity $F(\phi, \xi)$ in expectation,

$$\mathbb{E}_Y\left[\sum_{t=1}^{T} \mathbb{H}(X_t|X_{<t} = Y_{<t})\right] = F(\phi, \xi) \tag{20}$$

where the expectation is taken over the randomness of the signal trajectory Y that has the same distribution as X. The equality in (20) follows by noticing that the left hand side equals to the objective function in (9a), with the flexibility feedback there at each time $t \in [T]$ being optimal. Noting that the expectation in (20) equals to $F(\phi, \xi)$, the law of large numbers implies the theorem. □

5.2 Aggregator: Look-ahead approximation of the optimal flexibility feedback

As we have discussed, computing the exact optimal flexibility feedback vectors $p_1^*, \ldots, p_T^*$ is computationally intensive and so approximations are desirable. In Section 4.2 we have presented a data-driven approach for estimation via reinforcement learning. Here, we take a different approach based on looking ahead rather than referring to historical data. This approach is preferable in highly non-stationary situations. The approximation is presented in Appendix B. Notably, one may wonder if sending approximately optimal flexibility feedback to the operator is sufficient for achieving the desirable properties discussed in Section 3. In fact, it is and the results can be extended to hold for approximately optimal flexibility feedback computed as described above. Perhaps the most important of these properties is feasibility, and so we provide a detailed discussion of the extension for feasibility in Appendix C.

5.3 Experiments

In our experiments, we apply Monte Carlo estimation and look-ahead approximation to the ACN-Data [16].

System capacity $F(\phi, \xi)$ estimation. Figure 4 shows the estimated 365-day (average) system capacities $\mu_N(\phi, \xi)$ in (19) calculated by Monte Carlo estimation using the look-ahead approximation with $N = 5$, $k = 1$ and $T = 240$ from Sep. 1, 2018 to Aug. 31, 2019. We use parameters that match the setup of the garage. The total number of charging stations is fixed as 54, with peak power rate 6.6 kWh. The set of power levels $\mathbb{X}$ is a discrete set that contains 60 distinct power levels from 0 kWh to 360 kWh (for the definition of the parameters, see Appendix B). Note that, corresponding to this setting, in the case that every power trajectory in

the length-240 time horizon is feasible, the maximal system capacity is $240 \times \log 60 \approx 983$.

An interesting observation from this figure is that, although there are fewer users after Nov. 1, 2018 (because of switching from free-charging to paid-charging), there is no significant decrease of system capacity. Additionally, notice that there is a decline of users during the holidays, and therefore total flexibility drops during the Christmas season.
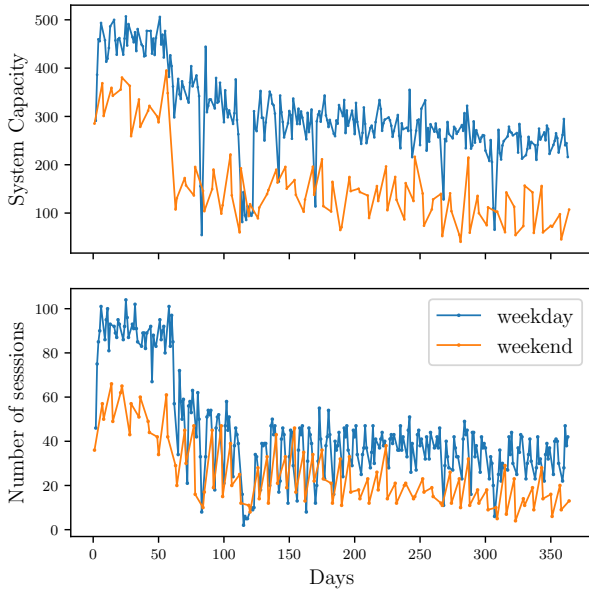

Figure 4: Estimated system capacities at Caltech with LLF compared with the number of charging sessions over a year, from Sep. 1, 2018 to Aug. 31, 2019. Weekends and weekdays are separated.

Real-time flexibility feedback. Now, let us study the quality of the real-time flexibility feedback. Eq. (20) gives the desired decomposition of system capacity, which enables us to characterize the spectrum of flexibility fluctuations. We show experimental results for real-time flexibility by considering a charging system within a single day. We use the same setting of parameters as described in Fig 4.

We consider the case when operational constraints present, and the operator seeks to perform peak shaving. In Figure 5, we vary the peak power limit defined in Example 2.2 and it shows that the smaller the limit is set to be, the lower real-time aggregate flexibility the system has. Note that summing the real-time aggregate flexibility over time estimates the system capacity. Therefore a sharper limit induces a lower system capacity. Supplementary experimental results can be found in Appendix E.

6   CONCLUDING REMARKS

This paper formalizes and studies the closed-loop control framework created by the interaction between a system operator and an
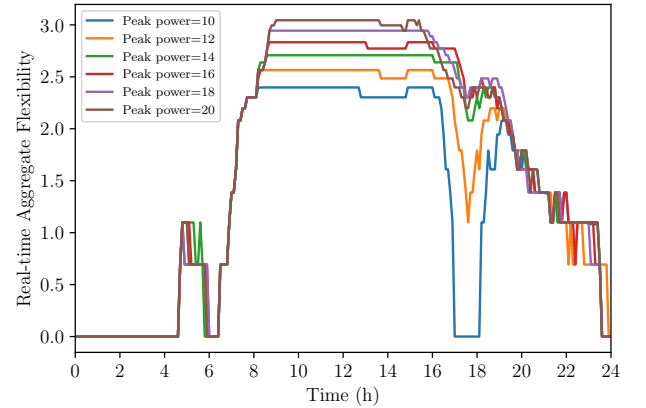

Figure 5: Impact of peak shaving constraints on real-time flexibility. We vary the peak power limit $\overline{\gamma}$(kWh) on Sep. 4, 2018.

aggregator. Our focus is on the feedback signal provided by the aggregator to the operator that summarizes the real-time availability of flexibility among the loads controlled by the aggregator. We present the design of an optimal flexibility feedback signal based on entropic maximization. We prove a close connection between the optimal flexibility feedback signal and the system capacity, and show that when the signal is used the system operator can perform online cost minimization and system capacity estimation while provably respecting the private constraints of the loads controlled by the aggregator. Further, we illustrate the effectiveness of these designs using simulation experiments of an EV charging facility.

There is much left to explore about this optimal flexibility feedback signal presented in this work. In particular, computing it is computationally intensive and we have presented two approaches for estimation. Improving these and developing other approximations is of particular interest. Further, exploring the use of flexibility feedback for operational objectives beyond cost minimization and capacity estimation is an important goal. Finally, exploring the application of flexibility feedback in other settings, such as frequency regulation and real-time pricing, is exciting.

ACKNOWLEDGMENTS

REFERENCES

[1] János Aczél, Bruno Forte, and Che Tat Ng. 1974. Why the Shannon and Hartley entropies are âĂŸnaturalâĂŹ. *Advances in applied probability* 6, 1 (1974), 131–146.

[2] Andrey Bernstein, Jean-Yves Le Boudec, Mario Paolone, Lorenzo Reyes-Chamorro, and Wajeb Saab. 2016. Aggregation of power capabilities of heterogeneous resources for real-time control of power grids. In *2016 Power Systems Computation Conference (PSCC)*. IEEE, 1–7.

[3] Scott Burger, Jose Pablo Chaves-Ávila, Carlos Batlle, and Ignacio J Pérez-Arriaga. 2017. A review of the value of aggregators in electricity systems. *Renewable and Sustainable Energy Reviews* 77 (2017), 395–405.

[4] Duncan S Callaway and Ian A Hiskens. 2010. Achieving controllability of electric loads. *Proc. IEEE* 99, 1 (2010), 184–199.

[5] Tianyi Chen, Na Li, and Georgios B Giannakis. 2018. Aggregating flexibility of heterogeneous energy resources in distribution networks. In *2018 Annual American Control Conference (ACC)*. IEEE, 4604–4609.

[6] Xin Chen, Emiliano Dall'Anese, Changhong Zhao, and Na Li. 2018. Aggregate Power Flexibility in Unbalanced Distribution Systems. *arXiv preprint arXiv:1812.05990* (2018).

[7] Imre Csiszár. 2008. Axiomatic characterizations of information measures. *Entropy* 10, 3 (2008), 261–273.

[8] Michael P Evans, Simon H Tindemans, and David Angeli. 2019. A Graphical Measure of Aggregate Flexibility for Energy-Constrained Distributed Resources. *IEEE Transactions on Smart Grid* (2019).

[9] Lingwen Gan, Ufuk Topcu, and Steven H Low. 2012. Optimal decentralized protocol for electric vehicle charging. *IEEE Transactions on Power Systems* 28, 2 (2012), 940–951.

[10] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. 2018. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *arXiv preprint arXiv:1801.01290* (2018).

[11] He Hao and Wei Chen. 2014. Characterizing flexibility of an aggregation of deferrable loads. In *53rd IEEE Conference on Decision and Control*. IEEE, 4059–4064.

[12] He Hao, Yashen Lin, Anupama S Kowli, Prabir Barooah, and Sean Meyn. 2014. Ancillary service to the grid through control of fans in commercial building HVAC systems. *IEEE Transactions on smart grid* 5, 4 (2014), 2066–2074.

[13] He Hao, Borhan M Sanandaji, Kameshwar Poolla, and Tyrone L Vincent. 2014. Aggregate flexibility of thermostatically controlled loads. *IEEE Transactions on Power Systems* 30, 1 (2014), 189–198.

[14] Emre C Kara, Jason S Macdonald, Douglas Black, Mario Bérges, Gabriela Hug, and Sila Kiliccote. 2015. Estimating the benefits of electric vehicle smart charging at non-residential locations: A data-driven approach. *Applied Energy* 155 (2015), 515–525.

[15] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[16] Zachary J Lee, Tongxin Li, and Steven H Low. 2019. ACN-Data: Analysis and Applications of an Open EV Charging Dataset. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. ACM, 139–149.

[17] Daria Madjidian, Mardavij Roozbehani, and Munther A Dahleh. 2018. Energy storage from aggregate deferrable demand: Fundamental trade-offs and scheduling policies. *IEEE Transactions on Power Systems* 33, 4 (2018), 3573–3586.

[18] Mousa Marzband, Andreas Sumper, José Luis Domínguez-García, and Ramon Gumara-Ferret. 2013. Experimental validation of a real time energy management system for microgrids in islanded mode using a local day-ahead electricity market and MINLP. *Energy Conversion and Management* 76 (2013), 314–322.

[19] Sean P Meyn, Prabir Barooah, Ana Bušić, Yue Chen, and Jordan Ehren. 2015. Ancillary service to the grid using intelligent deferrable loads. *IEEE Trans. Automat. Control* 60, 11 (2015), 2847–2862.

[20] Dimitrios Papadaskalopoulos, Goran Strbac, Pierluigi Mancarella, Marko Aunedi, and Vladimir Stanojevic. 2013. Decentralized participation of flexible demand in electricity marketsâĂŤPart II: Application with electric vehicles and heat pump systems. *IEEE Transactions on Power Systems* 28, 4 (2013), 3667–3674.

[21] Nasrin Sadeghianpourhamami, Nazir Refa, Matthias Strobbe, and Chris Develder. 2018. Quantitive analysis of electric vehicle flexibility: A data-driven approach. *International Journal of Electrical Power & Energy Systems* 95 (2018), 451–462.

[22] Intisar Ali Sajjad, Gianfranco Chicco, and Roberto Napoli. 2016. Definitions of demand flexibility for aggregate residential loads. *IEEE Transactions on Smart Grid* 7, 6 (2016), 2633–2643.

[23] Pierluigi Siano and Debora Sarno. 2016. Assessing the benefits of residential demand response in a real time distribution energy market. *Applied Energy* 161 (2016), 533–551.

[24] Miklós Simonovits. 2003. How to compute the volume in high dimension? *Mathematical programming* 97, 1-2 (2003), 337–374.

[25] Anand Subramanian, Manuel Garcia, A Dominguez-Garcia, D Callaway, Kameshwar Poolla, and Pravin Varaiya. 2012. Real-time scheduling of deferrable electric loads. In *2012 American Control Conference (ACC)*. IEEE, 3643–3650.

[26] Anand Subramanian, Manuel J Garcia, Duncan S Callaway, Kameshwar Poolla, and Pravin Varaiya. 2013. Real-time scheduling of distributed resources. *IEEE Transactions on Smart Grid* 4, 4 (2013), 2122–2130.

[27] George Wenzel, Matias Negrete-Pincetic, Daniel E Olivares, Jason MacDonald, and Duncan S Callaway. 2017. Real-time charging strategies for an electric vehicle aggregator to provide ancillary services. *IEEE Transactions on Smart Grid* 9, 5 (2017), 5141–5151.

[28] Lin Zhao, Wei Zhang, He Hao, and Karanjit Kalsi. 2017. A geometric approach to aggregate flexibility modeling of thermostatically controlled loads. *IEEE Transactions on Power Systems* 32, 6 (2017), 4721–4731.

A LEARNING FLEXIBILITY FEEDBACK

Soft actor-critic (SAC) [10] is an off-policy maximum entropy deep reinforcement learning algorithm, which in many complicated learning scenarios (such as control of humanoid robotics) outperforms deep deterministic policy gradient (DDPG) approaches, especially when the action space is a continuous and high-dimensional. The policy in our experiments is fixed to be a parameterized family of Gaussian distributions.

A.1 Approximate agent

We train an agent $\psi_t^{\text{SAC}} : \overline{\Omega} \to \mathcal{P}$ using SAC whose input at time $t \in [T]$ is a state parameter ξ_t that encodes the remaining energy to be delivered and the remaining charging time for the EV being charged at each station $i \in [W]$ and time $t \in [T]$, denoted by $\psi_t^{\text{SAC}}(i)$. Knowing the states $\xi_{<t}$, scheduling algorithm ϕ and the signals $x_{<t}$ gives the state ξ_t.

A.2 Parameters in our experiments

In the experiments, the state space is $\mathbb{R}_+^{2 \times W}$ where W is the total number of charging stations and a state vector for each charging station is $(e_t, [d(j) - t]^+)$, i.e., the remaining energy to be charged and the remaining charging time if it is being used; otherwise the vector is an all-zero vector. The action space is $\mathbb{R}_+^{|\mathbb{X}|}$. Moreover, the outputs of the neural networks are normalized into the probability simplex $\mathcal{P}$ afterwards. Hyper-parameters in our experiments are shown in Table 1.

Soft actor-critic	
Parameter	Value
optimizer	Adam [15]
learning rate	$3 \cdot 10^{-4}$
discount (γ)	0.5
relay buffer size	10^6
number of hidden layers	2
number of hidden units per layer	256
number of samples per minibatch	256
non-linearity	ReLU
temperature parameter (α)	0.5
Markov decision process	
power levels ($\mathbb{X}$)	$\{1, 2, \ldots, 20\}$ (in kWh)
number of stations (W)	Caltech (54) / JPL (52)
state space	$\mathbb{R}_+^{2 \times W}$
action space	$[0, 1]^{|\mathbb{X}|}$
reward $r_{\text{EV}}(\xi_t, p_t)$	$\sigma_1 = 0.1, \sigma_2 = 0.2, \sigma_3 = 2$
time interval (Δ)	6 minutes
operational constraints	$x_t \le 150$ (kWh), $\forall\ t \in [T]$

Table 1: Hyper-parameters in the experiments.

A.3 Reward function in training

For the deferrable loads in Example 2.1, once constraints are violated, they can on longer be satisfied by future decisions. Therefore, the minimization can be removed and we have the following specific reward function for EV charging scenario:

$$r_{\text{EV}}(\xi_{\le t}, p_t) = \mathbb{H}(p_t) + \sigma_1 \left\| \phi_t(\xi_{\le t}, x_{\le t}) \right\|_2$$

$$- \sigma_2 \left[e(j) - \sum_{t=1}^{T} \phi_t(j) \right]_+ - \sigma_3 \left| \pi_t^{\text{RHC}}(p_t) - \sum_{j=1}^{n} \phi_t(j) \right| \quad (21)$$

where $\sigma_0, \sigma_1, \sigma_2$ and σ_3 are positive constants. The second term is to enhance charging performance and the last two terms are realizations of the last term in (18) for constraints (2a) and (2b). The other constraints in Example 2.1 can automatically be satisfied by enforcing the constraints in the fixed scheduling algorithm ϕ.

With the settings described above, in Figure ?? we show a typical training curve of the reward function in (21). The constants in (21) are $\sigma_1 = 0.1$, $\sigma_2 = 0.2$ and $\sigma_3 = 2$.

B LOOK-AHEAD APPROXIMATION

Before presenting the design, we first introduce some generalized notation for feasible power levels that, that extends (10) to the case of k-step look-ahead.

$$\mathcal{S}_k(\phi, \xi | x_{<t}) := \Big\{ x_{t \to t+k-1} \in \mathbb{X}^k : \exists x_{t+k \to T} \text{ s.t.}$$

$$g_i(\phi, \xi, x) \le 0, \forall i = 1, \ldots, m \Big\}.$$

Further, the closed-form expression in Theorem 1 motivates us to consider the following approximation of the optimal flexibility feedback for all $x \in \mathbb{X}$ and $x_{<t} \in \mathbb{X}^{t-1}$:

$$\widehat{p}_{t,k}(x | x_{<t}) = \frac{|\mathcal{S}_k(\phi, \xi | (x_{<t}, x))|}{|\mathcal{S}_k(\phi, \xi | x_{<t})|}. \quad (22)$$

This, in turn, leads to the following recursive formula for all $2 \le k \le T - t + 1$:

$$|\mathcal{S}_k(\phi, \xi | x_{<t})| = \sum_{x \in \mathcal{S}_1(\phi, \xi | x_{<t})} |\mathcal{S}_{k-1}(\phi, \xi | (x_{<t}, x))|. \quad (23)$$

Here, we use k as the *look-ahead depth* and note that, when $k = T$, the approximation in (22) becomes exact.

Now, using this notation, in order to estimate the system capacity $F(\phi, \xi)$, we need to estimate the size of $\mathcal{S}_k(\phi, \xi | x_{<t})$ and $\mathcal{S}_k(\phi, \xi | x_{\le t})$ using the recursive formula in (23). Accomplishing this depends on calculating the feasible set for selecting x_t, given fixed $x_{<t}$, i.e., characterizing the set $\mathcal{S}_k(\phi, \xi | x_{<t})$ with look-ahead depth $k = 1$.

To provide a characterization of the first-order approximation $\mathcal{S}_1(\phi, \xi | x_{<t})$, we make a monotonicity assumption on the disaggregation policy ϕ, defined as follows.

DEFINITION B.1 (MONOTONICITY). *A (causal) disaggregation policy ϕ is **monotone** if for any $t \in [T]$, $x_t \le y_t$ implies that for all $x_{<t} \in \mathcal{S}_t(\phi, \xi)$,*

$$\phi_t(\xi_{<t}, (x_{<t}, x)) \preceq \phi_t(\xi_{<t}, (x_{<t}, y)). \quad (24)$$

Assuming that a scheduling algorithm ϕ is monotone, the feasible set of x_t conditioning on $x_{<t}$ can be characterized by a closed interval, as stated in the following theorem.

THEOREM 5. *Consider a system of deferrable loads with constraints specified by Example 2.1 and ξ being the associated states. For $t \in [T]$, for any ξ and monotone scheduling algorithm ϕ, the set $\mathcal{S}_1(\phi, \xi | x_{<t})$ can be written as the intersection of the set of power signals $\mathbb{X}$ and*

a closed interval: $S_1(\phi, \xi | x_{<t}) = \mathbb{X} \bigcup I_t(\phi, \xi)$, where $I_t(\phi, \xi) :=$ $[\alpha_t, \beta_t]$ is a closed interval in $\mathbb{R}_{\geq 0}$.

PROOF OF THEOREM 5. To prove the theorem, it is equivalent to show that if $x \in \mathbb{X}$ and $y \in \mathbb{X}$ with $x < y$ are two feasible power levels in $S_1(\phi, \xi | x_{<t})$, then any $z \in \mathbb{X}$ with $x \leq z \leq y$ is also in $S_1(\phi, \xi | x_{<t})$. Since the disaggregation policy ϕ is monotonically causal, the inequality $z \leq y$ guarantees that the power scheduled with $x_t = z$ to each load is always larger or equal to the case when $x_t = x$. Therefore, considering that $x_t = y$ is a feasible choice, since z does not violate any constraint for satisfying the demands of deferrable loads, it must also be a feasible power signal. □

It is typically straightforward to verify that the classical scheduling policies such as the least-laxity-first (LLF) scheduling and the earliest-deadline-first (EDF) scheduling are monotone, and thus the theory above applies. We demonstrate this for two classical policies, LLF and EDF, and one new policy termed feasibility interval maximization (FIM) in Appendix D. FIM is a new policy motivated by the feasibility analysis in this paper. Provided with a power signal $x_t > \alpha_t$, FIM assigns power to the loads with negative laxity, proportionally to $-\rho_{\text{Lax}}(j, t)$. Note that our purpose in discussing FIM is to demonstrate a contrast with LLF and EDF in our experimental results, not to present an "optimal" policy.

C FEASIBILITY VIA APPROXIMATE FLEXIBILITY FEEDBACK

Although the approximation of flexibility feedback in (22) is not precise, in this section we show that it is accurate enough to ensure feasibility under certain conditions. Specifically, consider a system of deferrable loads with constraints specified by Example 2.1. The following lemma states if none of the loads demands "excessive" energy upon arrival and the system has enough capacity for charging every load at their peak rates, then the system is always feasible by choosing the power signal according to the approximate flexibility feedback.

LEMMA 6. Consider a system of deferrable loads with constraints specified by Example 2.1. Suppose the following conditions hold:

(1) There is no operational constraints and there exists $x \in \mathbb{X}$ such that $x \geq \sum_{j=1}^{n} r(j)$.

(2) At each time $t \in [T]$, the selected $x_t \in \mathbb{X}$ satisfies (with look-ahead depth $k \geq 1$)

$$\widehat{p}_{t,k}(x | x_{<t}) > 0.$$

(3) The aggregator state ξ satisfies that

$$\rho_{\text{Lax}}(j, t) \geq 0, \text{ for } t \in [a(j), a(j) + 1]$$

and $j \in [n]$ where $\rho_{\text{Lax}}(j, t)$ is the laxity of the load j defined in (25).

It is guaranteed that $S_1(\phi, \xi | x_{<t}) \neq \emptyset$ for all $t \in [T]$ and any monotone disaggregation policy ϕ.

PROOF. Assuming that at time $t - 1 \in [T]$, the chosen power signal x_{t-1} satisfies $\widehat{p}_{t-1,k}(x_{t-1} | x_{<t-1}) > 0$, it remains to validate that there always exists some $x_t \in \mathbb{X}$ such that the approximated flexibility feedback $\widehat{p}_{t,k}(x | x_{<t}) > 0$, conditioning on the previously selected power levels $x_1, \ldots, x_{t-1}$. To see this, note that condition

(1) and (2) ensures that if the system is feasible at the previous time step $t - 1$ (i.e., $S_1(\phi, \xi | x_{<t-1}) \neq \emptyset$), then there is always a feasible power level in $\mathbb{X}$ for x_t, assuming there is no new loads arrive at the current time t. Condition (3) further guarantees that the demands of the new loads can also be satisfied, as long as the disaggregation policy ϕ is monotone. Therefore $\widehat{p}_{t,k}(x | x_{<t}) > 0$. By induction over $t \in [T]$ the proof is completed. □

D MONOTONICITY OF COMMON POLICIES

In this section we show that LLF, EDF, and FIM are monotone policies. Throughout, we fix the environment parameter ξ for deferrable loads and denote by $(a(j), d(j), e_t(j), r(j))$ the charging states of the j-th load at time $t \in [T]$ where $a(j), d(j)$ and $r(j)$ are defined in Example 2.1 and $e_t(j)$ is the remaining energy to be delivered at time $t \in [T]$. Additionally, let $\delta_t(j) := [d(j) - t]^+$ be the remaining charging duration (excluding the current time slot).

Least-laxity-first (LLF) scheduling. The laxity of the load $j \in [n]$ at time $t \in [T]$ is defined as

$$\rho_{\text{Lax}}(j, t) := \begin{cases} \delta_t(j) - e_t(j)/r(j), & t \geq a(j) \\ +\infty, & t < a(j) \end{cases}. \qquad (25)$$

If the laxity is negative, the car will never be fully charged and the the system becomes infeasible. Therefore, for any monotone disaggregation policy ϕ, the corresponding bounds must satisfy

$$\alpha_t \geq \sum_{j \in \underline{N}(t)} \min\left\{r(j), -\rho_{\text{Lax}}(j, t) r(j)\right\}, \qquad (26a)$$

$$\beta_t \leq \sum_{j \in \underline{N}(t)} \min\left\{r(j), e_t(j)\right\}, \qquad (26b)$$

where $\underline{N}(t) := \{j \in [n] : \rho_{\text{Lax}}(j, t) \leq 0\}$. It is immediate to see that the equalities (26a) and (26b) hold for LLF.

Earliest-deadline-first (EDF) scheduling. Under EDF the summation in (26a) for α_t needs to be replaced by a summation over $\underline{N}(t) \bigcup N_{\text{EDF}}(t)$ where a load j is in $N_{\text{EDF}}(t)$ if there exists $i \in \underline{N}(t)$ such that $d_j(t) \leq d_i(t)$:

$$N_{\text{EDF}}(t) := \left\{j \in [n] : \exists i \in \underline{N}(t) \text{ s.t. } d_j(t) \leq d_i(t)\right\}.$$

Feasibility-interval-maximization (FIM) scheduling. Recall that, when provided with a power signal $x_t > \alpha_t(\phi, \xi)$, FIM assigns power to the loads with negative laxity, proportionally to $-\rho_{\text{Lax}}(j, t)$. To understand the motivation behind FIM, observe that increasing the laxity of the loads decreases the lower bound $\alpha_t(\phi, \xi)$. Therefore, intuitively, it is desirable to ensure that as many loads as possible have non-negative laxity. Clearly, FIM is monotone, since the higher x_t is, the larger amount of energy is assigned to the EVs. The upper and lower bounds of the interval can be computed the same as in (26a) and (26b).

E SUPPLEMENTARY SIMULATION RESULTS

In Table 2, we summarize the quality of the capacity estimation, undelivered energy percentage, and tracking error (see (28) and (27) for definitions) for three scheduling policies, EDF, LLF and FIM, on both the Caltech and JPL garages. We measure performance using

the mean squared error (MSE) as the tracking error:

$$\text{MSE}(\phi, x) := \sum_{k=1}^{N} \sum_{t=1}^{T} \left| \sum_{j=1}^{n} \phi_t^{(k)}(j) - x_t^{(k)} \right|^2 / (N \times T), \quad (27)$$

where $x_t^{(k)}$ is the t-th power signal for the k-th test and $\phi_t^{(k)}(j)$ is the energy scheduled to the j-th load at time t for the k-th test. Additionally, define the mean percentage error with respect to the undelivered energy as

$$\text{MPE}(\phi, \xi) := \sum_{k=1}^{N} \sum_{t=1}^{T} \sum_{j=1}^{n} \phi_t^{(k)}(j) / \left((N \times T) \cdot \sum_{j=1}^{n} e_j \right), \quad (28)$$

where e_j is the energy request for each load j.

The results are averaged over the days from Sep. 1, 2018 to Aug. 31, 2019. The results show that FIM achieves the highest (estimated) system capacity, and lowest tracking error. However, as FIM always

maximizes the feasible charging interval, as a trade-off, its average percentage of undelivered energy is always the largest.

ϕ	System Capacity	Undelivered (%)	Tracking Error (kWh)
EDF	262.6726	**5.6044**	11.1570
LLF	271.7024	10.1406	3.1356
FIM	**271.9571**	10.3148	**2.7661**
EDF	221.5389	**6.6837**	15.5949
LLF	242.4101	11.8726	7.0418
FIM	**242.8318**	12.1583	**6.6034**

Table 2: Estimated system capacity $F(\phi, \xi)$, undelivered energy percentage $\text{MPE}(\phi, \xi)$, and tracking error $\text{MSE}(\phi, x)$ for Caltech (top) and JPL (bottom) comparing EDF, LLF, and FIM.