Characterizing capacity of flexible loads for providing grid support

Austin R. Coffman*, Zhong Guo*,†, and Prabir Barooah*

***Abstract*—Flexible loads are a resource for the Balancing Authority (BA) of the future to aid in the balance of supply and demand in the power grid. Consequently, it is of interest for a BA to know how much flexibility a collection of loads has, so to successfully incorporate flexible loads into grid level resource allocation. Loads' flexibility is limited by all their Quality of Service (QoS) requirements. In this work we present a characterization of capacity for a collection of flexible loads. This characterization is in terms of the Power Spectral Density (PSD) of the reference signal. Two advantages of our characterization are: (i) it easily allows for a BA to use the characterization for resource allocation of flexible loads and (ii) it allows for precise definitions of the power and energy capacity for a collection of flexible loads.**

I. Introduction

The inherent variability in renewable generation sources such as solar and wind is a challenge for the power grid operators to balance demand and generation. Ramp rate constraints prevent conventional generation from handling this mismatch between demand and generation completely, while grid level storage from batteries is expensive. Thus a new resource is being investigated to help fill the mismatch: flexible loads. Flexible loads have the ability to vary power consumption over their baseline demand without violating their Quality of Service (QoS). The baseline power consumption is the power consumed without interference from the grid. The amount of deviation from the baseline demand, requested from the grid, is the *reference signal*. The tracking of a zero-mean reference signal guises, in the eyes of the grid operator, flexible loads as batteries providing storage services. This battery-like behavior of flexible loads is often referred to as Virtual Energy Storage (VES) [1]. VES from flexible loads can be less expensive than energy storage from batteries [2]. Examples of flexible loads include residential air conditioners [3], water heaters [4], refrigerators, commercial HVAC systems [5], pumps for irrigation [6], pool cleaning [7] or heating [8].

If the grid operator expects the flexible loads to track the reference signal accurately, then the reference must not cause the loads to violate their QoS. From the viewpoint of the grid operator, flexible loads not tracking a reference makes them appear unreliable. From the viewpoint of the load, reference signals that continually require QoS violation provide incentive for loads to stop providing VES. In either case, avoidance of the above scenarios is paramount to the long term success of VES. That is, reference signals must be designed to respect the *capacity* of the collection of flexible loads.

Informally, the capacity of an ensemble represents limitations in ensemble behavior (e.g., the ability to track a reference signal) due to QoS requirements at the individual loads. Consequently, a key step in determining the capacity is relating the QoS requirements to requirements for the reference signal. Unfortunately, this is not a straight forward task and many varying approaches are present in the current literature [9]–[15]. A popular approach is to develop ensemble level necessary conditions [9], [11]; reference signals that satisfy these conditions ensure the ability of all loads in the collection to satisfy QoS while tracking the reference. Other approaches include geometry based characterizations [16], characterizations through distributed optimization [17], and characterizations that approximate the Minkowski sum of individual loads' "resource polytopes" [18]–[20].

A limitation of some of the currently available capacity characterization results is that a balancing authority (BA) requires a specific reference signal trajectory to utilize the characterization. This is because these approaches characterize the capacity in terms of a set of constraints on the ensembles time domain power consumption [9], [11], [18]–[20]. Using these characterizations for long term planning would then require the BA to predict its needs many months in advance. In the rare case that this is possible, the BA would have to generate a new predicted trajectory every time it wants to ensure flexible loads can meet its needs.

This paper is about capacity characterization for long term planning, which allows answering questions of the form "how many flexible loads are required to offset the volatility in its investment in a solar farm". That is, will the BA require 10000 or 5000 flexible loads?

It turns out that if one develops constraints on the statistics of the reference signal, then a notion of capacity that is also useful for long term planning can be developed. More specifically, consider Figure 1 where the Power Spectral Density (PSD) of a grid's net demand is allocated to resources (precise definition of PSD is provided later). This frequency based allocation does not require knowledge of a specific reference signal, but only of the statistics of the reference signal. This is because one single PSD represents infinitely many potential realizations of reference signal trajectories that share the same statistics. If a BA can describe its needs through statistical quantities, such as the PSD of the net

* University of Florida
† corresponding author, email: zhong.guo@ufl.edu.
ARC, ZG, and PB are with the Dept. of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32601, USA. The research reported here has been partially supported by the NSF through award 1646229 (CPS-ECCS).

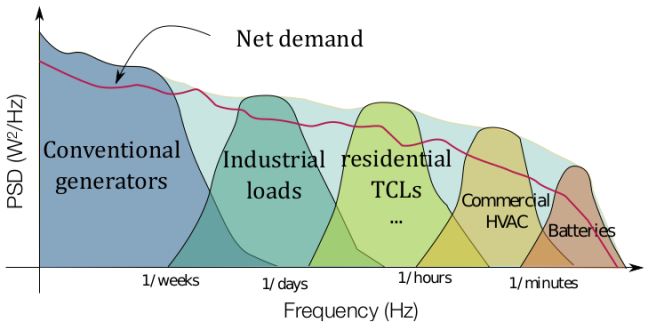


Fig. 1: An example spectral allocation of resources to meet the grids needs.

demand in Figure 1, then it can determine the required PSD from its resources. That is, the needs of the BA can be described without having to specify a specific reference signal. To avoid possible confusion with electrical power (in Watt) we use Spectral Density (SD) instead of Power Spectral Density (PSD) in the rest of the paper.

In order for the above argument to hold, it must be possible to express the flexibility of loads in terms of statistical properties, such as the SD. In fact, in [15], [21] it was shown that the QoS requirements of flexible loads can be characterized as constraints in the frequency domain. That is, while primarily an illustrative example, it is possible to quantitatively develop the regions shown in Figure 1.

In addition to [15], [21], there are many works that advocate for the frequency-domain specification of loads' abilities (e.g., [5], [22]) and for resource allocation (e.g., [23]). The results of real world VES experiments also suggest that specifying the spectral content of a reference signal is a feasible way to encapsulate the limitations in flexibility of a load [24]. A simplification of this concept is also widely used in today's power grid; ancillary services are classified by their response times and ramp rates [25].

Motivated by the advantages of working in the frequency domain we extend the work in [21], [26] and characterize the capacity through the SD of the reference signal, for an entire collection of flexible loads, based on the QoS requirements of each flexible load. Capacity characterization in terms of a SD, S, means that any signal whose SD is below S is *feasible*: the ensemble can track it without any load violating its local QoS. The QoS constraints considered here are quite general and encapsulate operating constraints for: (i) commercial HVAC systems, (ii) batteries, and (iii) a collection of thermostatically controlled loads (TCLs).

The contribution over past work is threefold. First, we characterize capacity through constraints on the statistics of the reference signal, rather than the reference signal itself. Second, our characterization of capacity allows for a BA to easily perform long term resource allocation. That is, as we will see, our capacity characterization will allow for the BA to answer questions such as: what is the benefit of investing in 10000 over 5000 flexible water heaters? Our third contribution is that our capacity characterization allows for precise definitions of the power (Watt) and energy (Watt-hour) capacity for a collection of flexible loads. The third contribution allows direct comparison between capacity of a real (e.g., electrochemical) energy storage device and a collection of flexible loads providing VES, using common terminology (MW and MWh).

We corroborate the advantages of our capacity characterization through numerical experiments. A convex optimization problem is posed that 'projects' the needs of the BA onto the set of feasible SD's. The needs of the BA is quantified as the SD of net-demand, as seen in Figure 1. To illustrate this process and our capacity characterization, as an example, we consider an ensemble of flexible HVAC systems in commercial buildings. In each simulation scenario, we determine the power and energy capacity of the ensemble of HVAC systems based on the obtained SD.

The capacity characterization presented is for continuously varying loads, that is loads that can vary power consumption freely within an interval. The capacity characterization can be extended to handle discrete loads (e.g., TCLs), by applying it to the work [11].

A preliminary version of this work is published in [26], which dealt with a single load. In this work we extend the results to a heterogeneous collection of loads.

The paper proceeds as follows, Section II describes the QoS constraints of loads, Section III describes mathematical prerequisites and a characterization of individual load capacity, Section IV describe the spectral characterization of the constraints for the ensemble. Section V describes the method for determining how much of a BA's needs can be satisfied by flexible loads. Numerical results are provided in Section VI.

II. QoS Constraints of Individuals

Denote by $P(t)$ the power consumption of a flexible load at time t, and let $P^b(t)$ its baseline demand. The demand deviation is $\tilde{P}(t) := P(t) - P^b(t)$. The load provides VES service by controlling the deviation $\tilde{P}(t)$ to track a desired deviation signal, called a reference, while maintaining its own QoS. The first QoS constraint is simply an actuator constraint:

$$\text{QoS-1:} \quad \left|\tilde{P}(t)\right| \leq c_1, \quad \forall\ t, \tag{1}$$

where the constant c_1, the maximum possible deviation of power consumption, depends on the rated power and the baseline demand. Second, define the demand increment $\tilde{P}_\delta(t) := \tilde{P}(t) - \tilde{P}(t-\delta)$, where $\delta > 0$ is a predetermined (small) time interval. The second constraint is a ramping rate constraint:

$$\text{QoS-2:} \quad \left|\tilde{P}_\delta(t)\right| \leq c_2, \quad \forall\ t. \tag{2}$$

Third, define the additional energy use during any time interval of length T:

$$\tilde{E}(t) = \int_{t-T}^{t} \tilde{P}(\sigma) d\sigma. \tag{3}$$

The third QoS constraint is that

$$\text{QoS-3:} \quad \left|\tilde{E}(t)\right| \le c_3, \quad \forall\ t. \tag{4}$$

The parameter T in (3) can represent the length of a billing period. Ensuring (4) ensures that the energy consumed during a billing period close to the nominal energy consumed, although it is stronger than what is necessary. Eq. (3) has a representation in Laplace domain as $\tilde{E}(s) = G(s)\tilde{P}(s)$ where the BIBO stable transfer function G is given by $G(s) = \frac{1-e^{-sT}}{s}$.

To define the fourth and last QoS constraint, we associate with the VES system a *storage variable* $\tilde{\theta}(t)$ that is related to the demand deviation through a stable linear time invariant system $H(s)$ as follows:

$$\tilde{\theta}(s) = H(s)\tilde{P}(s), \tag{5}$$

and impose the constraint

$$\text{QoS-4:} \quad \left|\tilde{\theta}(t)\right| \le c_4, \quad \forall\ t. \tag{6}$$

The user-specified constants c_1, c_2, c_3, c_4 are called equipment bounds in the sequel.

1) Understanding QoS-4: To understand the storage variable, imagine a flexible HVAC system providing VES. A simple model of the indoor temperature θ_z of the building this HVAC system serves is:

$$C\dot{\theta}_z(t) = \frac{1}{R}\left(\theta_a(t) - \theta_z(t)\right) + \dot{q}_{\text{int}}(t) + \eta_{\text{COP}} P(t), \tag{7}$$

where C and R are thermal capacitance and resistance, $\theta_a(t)$ is the ambient temperature, η_{COP} is the coefficient of performance, and $\dot{q}_{\text{int}}(t)$ is an exogenous disturbance. The quantity $P(t)$ is the total electrical power consumption of the HVAC system.

In general, the baseline power consumption for a HVAC system is the value $P^b(t)$ that keeps the internal temperature of the load at a fixed value $\bar{\theta}$, which for (7) is

$$P^b(t) = -\frac{\left(\theta_a(t) - \bar{\theta}\right)}{\eta_{\text{COP}} R} - \frac{\dot{q}_{\text{int}}(t)}{\eta_{\text{COP}}}. \tag{8}$$

Since we are concerned with the flexibility in the load, we linearize (7) about the thermal setpoint $\bar{\theta}$ and the baseline power $P^b(t)$ yielding,

$$\dot{\tilde{\theta}}_z(t) = -\gamma\tilde{\theta}_z(t) + \beta\tilde{P}(t), \quad \gamma = \frac{1}{RC}, \quad \beta = \frac{\eta_{\text{COP}}}{C}, \tag{9}$$

where $\tilde{\theta}_z \triangleq \theta_z(t) - \bar{\theta}$ is the internal temperature deviation and $\tilde{P}$ is as defined at the beginning of this Section.

Taking the Laplace transform of (9), we get $\tilde{\theta}_z(s) = \frac{\beta}{s+\gamma}\tilde{P}(s)$. So, for the HVAC system example, the storage variable is simply the indoor temperature deviation from the baseline value, i.e., $\tilde{\theta}(t) := \tilde{\theta}_z(t)$ and $H(s) = \frac{\beta}{s+\gamma}$. A more complex model of HVAC dynamics would lead to a higher order $H(s)$. The model (7), while simplistic, has been shown to agree quite well with more realistic models for certain flexible loads [27].

Also, when the VES system is in fact a battery, $\tilde{\theta}(t)$ can be thought of as the amount of energy stored in the battery at time t, i.e., $\tilde{\theta}(t) = E_0 x_{\text{SoC}}(t)$, where E_0 is the nominal energy capacity in kWh and x_{SoC} is the state of charge. A simple dynamic model of this variable is $\dot{\tilde{\theta}}(t) = -\alpha\tilde{\theta}(t) + \tilde{P}(t)$, where $-\alpha\tilde{\theta}(t)$ accounts for the leakage, self degradation, and non-unity round trip efficiency of the battery. In this case $H(s) = \frac{1}{s+\alpha}$

The four constraints QoS 1-4, with parameters $Q^i \triangleq \left(\{c_i\}_{i=1}^4, T, \delta\right)^i$ specify the QoS set for the VES system. The question is, what is a feasible power deviation signal $\tilde{P}(t)$?

III. Individual Load Capacity

A. Preliminaries

To develop our capacity characterization we will require a few preliminaries, which we list in this section. First, we switch from a deterministic to a stochastic setting. In this setting we model the power deviation, $\tilde{P}$ as a continuous time stochastic process. The mean and autocorrelation function for $\tilde{P}$ are,

$$\mu_{\tilde{P}}(t) \triangleq \mathsf{E}[\tilde{P}(t)], \qquad \forall\ t, \tag{10}$$

$$R_{\tilde{P}}(\sigma, t) \triangleq \mathsf{E}[\tilde{P}(\sigma)\tilde{P}(t)], \quad \forall\ \sigma,\ t, \tag{11}$$

where $\mathsf{E}[\cdot]$ denotes expectation. We make the following assumption about the stochastic process $\tilde{P}$.

Assumption 1. *The stochastic process $\tilde{P}$ is wide sense stationary (WSS) with mean $\mu_{\tilde{P}}(t) = 0$ for all t.*

Since $\tilde{P}$ is the difference of the power consumption from a baseline value, it is intuitive that its mean is zero. Otherwise, loads are not providing storage services. Furthermore, WSS requires the variance and mean of the process $\tilde{P}$ to be time invariant and the autocorrelation function to be a function of $\tau = \sigma - t$. We denote the time invariant variance as $\sigma^2_{\tilde{P}}$.

In addition, for a continuous time WSS stochastic process $\{X(t)\}$ we have, through the Fourier transform, an alternative expression for the autocorrelation function [28],

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(\omega) e^{(j\omega\tau)} d\frac{\omega}{2\pi}, \text{ and} \tag{12}$$

$$S_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau) e^{(-j\omega\tau)} d\tau, \tag{13}$$

where $\omega \in \mathbb{R}$ is frequency, and j is the imaginary unit, and $S_X(\omega)$ is the (power) Spectral Density (SD) of X:

$$S_X(\omega) \triangleq \lim_{T\to\infty} \frac{1}{T}\mathsf{E}\left[\left|\int_0^T X(t) e^{-j\omega t} dt\right|^2\right]. \tag{14}$$

The equivalence of definitions (14) and (13) for a zero-mean WSS process is the Wiener-Khinchin theorem. Letting $\tau = 0$ in (12) results in,

$$R_X(0) = \int_{-\infty}^{\infty} S_X(\omega) d\frac{\omega}{2\pi}, \tag{15}$$

where, if the mean of X is zero, $R_X(0)$ is the variance of the process X. We introduce the following proposition from [28] that will be useful in the developments to follow.

Proposition 1. *[28] Let $X(t)$ be zero-mean WSS stochastic process and input to the linear time invariant BIBO stable system $M(s)$ with output $Y(t)$. Then Y is WSS, X and Y are jointly WSS, and*

$$\begin{aligned} &(i) \quad \mathsf{E}[Y] = M(j\omega)\Big|_{\omega=0} \mathsf{E}[X], \\ &(ii) \quad S_Y(\omega) = |M(j\omega)|^2 S_X(\omega), \end{aligned}$$

where S_X, S_Y are the SD's of X and Y.

Furthermore, the Chebyshev inequality for a random variable X, will be useful:

$$\mathsf{P}\big(|X - \mathsf{E}[X]| \geq k\big) \leq \frac{\sigma_X^2}{k^2}, \quad \forall\ k > 0, \tag{16}$$

where $\mathsf{P}(\cdot)$ denotes probability.

B. Inequality Constraints: Spectral Characterization

The QoS constraints QoS 1-4 are characterized probabilistically in the following way. The inequalities in QoS 1-4 turn into probabilistic inequalities; the probability of the QoS constraint *not* being met is required to be small:

$$\mathsf{P}\left(\left|\tilde{P}(t)\right| \geq c_1\right) \leq \epsilon_1, \qquad \forall\ t, \tag{17}$$
$$\mathsf{P}\left(\left|\tilde{P}_\delta(t)\right| \geq c_2\right) \leq \epsilon_2, \qquad \forall\ t, \tag{18}$$
$$\mathsf{P}\left(\left|\tilde{E}(t)\right| \geq c_3\right) \leq \epsilon_3, \qquad \forall\ t, \tag{19}$$
$$\mathsf{P}\left(\left|\tilde{\theta}(t)\right| \geq c_4\right) \leq \epsilon_4, \qquad \forall\ t, \tag{20}$$

The quantities $\{\epsilon_i\}_{i=1}^4$ set the tolerance level for satisfying the respective constraint and are chosen to be small.

In order to pose the inequality constraints (17)-(20) in terms of $S_{\tilde{P}}$, two steps are taken. The first step is to utilize the Chebyshev inequality (16) to bound the probabilities in (17)-(20) as a function of the variance of the given random variable. The second step is to then use the Wiener-Khinchin theorem (12) to express the variance as the integral of $S_{\tilde{P}}$.

Lemma 1. *Let $\tilde{P}$ satisfy Assumption 1, then for all t,*

$$\mathsf{E}[\tilde{E}(t)] = 0,\ \mathsf{E}[\tilde{P}_\delta(t)] = 0,\ \textit{and}\ \mathsf{E}[\tilde{\theta}(t)] = 0.$$

Proof. Apply the result of Proposition 1-(i) for $\mathsf{E}[\tilde{E}(t)]$ and $\mathsf{E}[\tilde{\theta}(t)]$. The linearity of expectation suffices for $\mathsf{E}[\tilde{P}_\delta(t)]$. □

With the result in Lemma 1 and Chebyshev's inequality (16) we formulate sufficient conditions for the inequality constraints (17)-(20) as follows,

$$\sigma_{\tilde{P}}^2 \leq c_1^2\epsilon_1, \qquad \sigma_{\tilde{P}_\delta}^2 \leq c_2^2\epsilon_2, \tag{21}$$
$$\sigma_{\tilde{E}}^2 \leq c_3^2\epsilon_3, \qquad \sigma_{\tilde{\theta}}^2 \leq c_4^2\epsilon_4, \tag{22}$$

so that the probability of exceeding the inequality constraints (17)-(20) will be less than the respective specified amount, $\{\epsilon_i\}_{i=1}^4$. The variance $\sigma_{\tilde{P}_\delta}^2$ is equivalently,

$$\sigma_{\tilde{P}_\delta}^2 = \mathsf{E}\left[\left(\tilde{P}_\delta(t)\right)^2\right] = 2\left(R_{\tilde{P}}(0) - R_{\tilde{P}}(\delta)\right). \tag{23}$$

Now applying the Wiener-Khinchin theorem (12) to the LHS of each inequality in (21)-(22) we have,

$$\int_0^\infty S_{\tilde{P}}(\omega) d\frac{\omega}{\pi} \leq c_1^2\epsilon_1, \tag{24}$$
$$\int_0^\infty S_{\tilde{P}}(\omega)\left(2 - 2\cos(\omega\delta)\right) d\frac{\omega}{\pi} \leq \epsilon_2 c_2^2, \tag{25}$$
$$\int_0^\infty S_{\tilde{E}}(\omega) d\omega = \int_0^\infty |G(j\omega)|^2 S_{\tilde{P}}(\omega) d\frac{\omega}{\pi} \leq c_4^2\epsilon_4, \tag{26}$$
$$\int_0^\infty S_{\tilde{\theta}}(\omega) d\frac{\omega}{\pi} = \int_0^\infty |H(j\omega)|^2 S_{\tilde{P}}(\omega) d\frac{\omega}{\pi} \leq c_3^2\epsilon_3. \tag{27}$$

Definition 1. *(Individual load constraint set) Let $\epsilon = [\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4]$, the set of feasible SDs is then*

$$\mathcal{S}_\epsilon = \{S_{\tilde{P}} \ :\ S_{\tilde{P}} \geq 0,\ \textit{and}\ (24) - (27)\}.$$

That is, if $\tilde{P}$ has SD in the set $\mathcal{S}_\epsilon$, then the load's power deviation can match $\tilde{P}$ while satisfying the constraints (17)-(20). The set $\mathcal{S}_\epsilon$ is given explicit dependence on ϵ to reflect the dependence of constraints (24)-(27) on ϵ.

C. Power and energy capacity: the capacity of a SD

The past subsection culminated in the definition of the constraint set for an individual load in Definition 1. The next natural thing to define then, is the energy and power capacity for each SD that is in this constraint set. We now do this. Utilizing the concept of SD we can define not only the power and energy capacity, but also the temperature and rate capacity. All of these definitions will be in terms of the SD of the power deviation.

Definition 2. *Let $\mathcal{S}_\epsilon$ be as defined in Definition 1, then for a given SD $S \in \mathcal{S}_\epsilon$ denote the following,*

$$\textbf{\textit{Pow}}(S) = \sqrt{\frac{1}{\pi\epsilon_1}\int_0^\infty S(\omega)d\omega} \quad (W),$$
$$\textbf{\textit{Rate}}(S) = \sqrt{\frac{1}{\pi\epsilon_2}\int_0^\infty \left(2 - 2\cos(\omega\delta)\right) S(\omega)d\omega} \quad (W),$$
$$\textbf{\textit{Eng}}(S) = \sqrt{\frac{1}{\pi\epsilon_3}\int_0^\infty |G(j\omega)|^2 S(\omega)d\omega} \quad (Wh),$$
$$\textbf{\textit{ThEng}}(S) = \sqrt{\frac{1}{\pi\epsilon_4}\int_0^\infty |H(j\omega)|^2 S(\omega)d\omega} \quad (^\circ C),$$

which are the power, rate, energy, and temperature capacity for the given SD $S \in \mathcal{S}_\epsilon$.

The reason for these definitions is that for a signal $\tilde{P}(t)$ with SD S, the following inequalities follow:

$$\mathsf{P}\left(\left|\tilde{P}(t)\right| \geq \textbf{Pow}(S)\right) \leq \epsilon_1, \quad \forall\ t, \tag{28}$$
$$\mathsf{P}\left(\left|\tilde{E}(t)\right| \geq \textbf{Eng}(S)\right) \leq \epsilon_4, \quad \forall\ t, \tag{29}$$

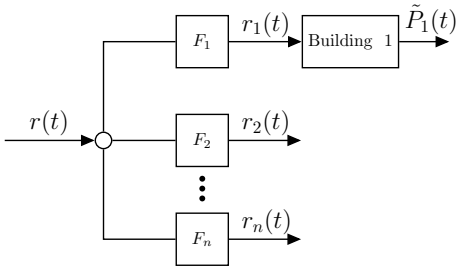


Fig. 2: An example control architecture for flexible loads to assist the grid.

where $\tilde{E}(t)$ is related to $\tilde{P}(t)$ through (26). This is an application of the Chebyshev inequality (16) with $k = \mathbf{Pow}(S)$ (respectively, $\mathbf{Eng}(S)$). The definitions $\mathbf{ThEng}(S)$ and $\mathbf{Rate}(S)$ have the same interpretation. Practically, Definition 2 through the Chebyshev inequality yields a connection between the capacity of the SD and upper/lower bounds in the time domain in equations (28)-(29). As an example, if a signal has SD S, then for a given ϵ_1 the probability of the signal being above $\mathbf{Pow}(S)$ is less than ϵ_1. This suggests that the quantities in Definition 2 will serve as tight probabilistic bounds for time series with the specified SD. Section VI-D provides further investigation of our capacity definition (Definition 2) through numerical experiments.

IV. Ensemble Capacity

Section III constructed constraints on the SD $S_{\bar{P}}$ so to respect the capacity requirements for an individual load. In this section we do the same for n heterogeneous loads by defining an "ensemble SD" and utilizing the constraints on the individuals to develop constraints for the ensemble SD. First, the ensemble power deviation is defined as

$$\bar{P}(t) \triangleq \sum_{i=1}^{n} \tilde{P}^i(t). \tag{30}$$

The ensemble SD is then the SD of $\bar{P}(t)$, that is,

$$S_{\bar{P}}(\omega) = \int_{-\infty}^{\infty} R_{\bar{P}}(\tau) e^{-j\omega\tau} d\tau, \tag{31}$$

where $R_{\bar{P}}(\tau)$ is the autocorrelation function of $\bar{P}$. With the requirement that each $\tilde{P}^i(t)$ is jointly WSS, the ensemble power deviation (30) is also WSS and the definition (31) is valid.

The SD of the sum, being the Fourier transform of the autocorrelation of the sum, depends on how the component signals $\tilde{P}^i$ are correlated to one another. In the limiting case when they are uncorrelated with one another, the autocorrelation of the sum is the sum of the autocorrelations since the cross correlations are 0. In that case the SD of the sum is also the sum of the SDs. Similarly, in the limiting cases of being perfectly anti-correlated or correlated, one can come up with precise relationship. However, the most interesting case from a practical point of view is when they signals are positively correlated but without having a perfect correlation of 1. To understand the reason, let us imagine how these signals $\tilde{P}^i$ will be generated. It is reasonable to assume a control architecture shows in Figure 2: a reference signal from the balancing authority, $r(t)$, is split into n distinct components, $r_i(t)$, $i = 1, \ldots, n$, so that $r_i(t)$ is within the capacity of load i, and some control system is used to ensure that $\tilde{P}^i$ tracks r_i. For instance, r_i can be computed by band-pass filtering r with a filter $F_i(j\omega)$. In a well designed system, these filters will ensure that the signals r_i have sufficient positive correlation. Otherwise, each load - assuming they perfectly track their references - will be working against each other instead of working collaboratively to supply the total r. Although the design of the coordination architecture for the loads is beyond the scope of this work, we make the following assumption:

Assumption 2. *For any pair of loads ℓ, m, $\mathsf{E}[\tilde{P}_\ell(t)\tilde{P}_m(t+\tau)] =: R_{\ell,m}(\tau) \geq 0$ for every τ.*

Lemma 2. *The ensemble SD of n loads, under Assumption 2, satisfies*

$$\Sigma_{\bar{P}} \leq S_{\bar{P}} \leq n\Sigma_{\bar{P}}, \tag{32}$$

where $\Sigma_{\bar{P}} := \sum_{i=1}^{n} S_{\tilde{P}^i}$.

Proof. Since the signals in the sum are jointly WSS, we have

$$R_{\bar{P}(t),\bar{P}(t)}(\tau) = \sum_{\ell=1}^{n}\sum_{m=1}^{n} R_{\ell,m}(\tau). \tag{33}$$

The cross correlation $R_{\ell,m}(\tau)$ when $\ell \neq m$ is non negative by Assumption 2, $R_{\bar{P}(t),\bar{P}(t)}(\tau) \geq \sum_{\ell=1}^{n} R_{\ell,\ell}(\tau)$, from which the lower bound follows due to linearity of the Fourier transform. The upper bound is proven by considering the definition (14) for $S_{\bar{P}}(\omega)$. We have for all $\omega \in \mathbb{R}$: $S_{\bar{P}}(\omega) =$

$$\lim_{T\to\infty} \frac{1}{T}\mathsf{E}\left[\left|\sum_{i=1}^{n}\int_0^T \tilde{P}^i(t)e^{-j\omega t}dt\right|^2\right] \tag{34}$$

$$\leq \lim_{T\to\infty} \frac{n}{T}\sum_{i=1}^{n}\mathsf{E}\left[\left|\int_0^T \tilde{P}^i(t)e^{-j\omega t}dt\right|^2\right] = n\Sigma_{\bar{P}}(\omega), \tag{35}$$

where the bound is from Jensen's inequality: $\left|\sum_{i=1}^{n} x_i\right|^2 \leq n\sum_{i=1}^{n}|x_i|^2$ since $|\cdot|^2$ is convex. □

Corollary 1. *For a homogeneous collection of n loads in which $\tilde{P}_\ell(t) = \tilde{P}_m(t)$ for every $\ell, m = 1, \ldots, n$, the ensemble SD is*

$$S_{\bar{P}} = n\Sigma_{\bar{P}}.$$

The proof follows from the same argument used in establishing the upper bound case of Lemma 2.

In addition, for a heterogeneous ensemble the lower bound in Lemma 2 will be loose as n increases. A better bound can be obtained by 'binning' the heterogeneous collection into several homogeneous ensembles and utilizing the result of Corollary 1.

Corollary 2. *For a heterogeneous collection of n loads with n_{bin} bins indexed by ℓ with n_ℓ loads in the ℓ^{th} bin, the ensemble SD is bounded as,*

$$\sum_{\ell=1}^{n_{bin}} n_\ell \Sigma_{\bar{P}^\ell} \leq S_{\bar{P}} \leq n_{bin} \sum_{\ell=1}^{n_{bin}} n_\ell \Sigma_{\bar{P}^\ell},$$

with

$$\Sigma_{\bar{P}^\ell} \triangleq \sum_{(i \in bin\ \ell)} S_{\tilde{P}^i}.$$

Proof. Apply directly the same procedure as in Lemma 2, except for n_{bin} loads with the ℓ^{th} SD being $n\Sigma_{\bar{P}^\ell}$ (Corollary 1). □

A. Ensemble constraint set

We now develop a constraint set on $\Sigma_{\bar{P}^\ell}$ based on the set of constraints for the individual SDs that are in the definition of $\Sigma_{\bar{P}^\ell}$. In light of the results of Corollary 2, we assume a homogeneous collection of loads. The idea is to sum the individual constraints over n_ℓ so that the definition of $\Sigma_{\bar{P}^\ell}$ can be inserted. We do this for the rate constraint below,

$$\sum_{(i \in \text{bin } \ell)} \int_0^\infty (1-\cos(\omega\delta))\, S_{\tilde{P}^i}(\omega) d\omega \leq n_\ell \frac{\pi\epsilon_2 c_2^2}{2},$$
$$\iff \int_0^\infty (1-\cos(\omega\delta))\, \Sigma_{\bar{P}^\ell}(\omega) d\omega \leq n_\ell \frac{\pi\epsilon_2 c_2^2}{2}.$$

The other constraints are obtained in a similar fashion, and the full constraint set for $\Sigma_{\bar{P}^\ell}$ is: $\mathcal{S}(n_{\text{bin}}, n_\ell, \ell) \triangleq$

$$\Bigg\{ \Sigma_{\bar{P}^\ell} \ :\ \Sigma_{\bar{P}^\ell} \geq 0,$$
$$\int_0^\infty \Sigma_{\bar{P}^\ell}(\omega) d\omega \leq n_{\text{bin}} n_\ell \left(\pi\epsilon_1 c_1^2\right),$$
$$\int_0^\infty (1-\cos(\omega\delta)) \Sigma_{\bar{P}^\ell}(\omega) d\omega \leq n_{\text{bin}} n_\ell \left(\frac{\pi\epsilon_2 c_2^2}{2}\right),$$
$$\int_0^\infty |G(j\omega)|^2 \Sigma_{\bar{P}^\ell}(\omega) d\omega \leq n_{\text{bin}} n_\ell \left(\pi\epsilon_3 c_3^2\right),$$
$$\int_0^\infty |H(j\omega)|^2 \Sigma_{\bar{P}^\ell}(\omega) d\omega \leq n_{\text{bin}} n_\ell \left(\pi\epsilon_4 c_4^2\right) \Bigg\}.$$

We use $\mathcal{S}(1, n_\ell, \ell)$ and $\mathcal{S}(n_{\text{bin}}, n_\ell, \ell)$ to denote the constraint sets for the ℓ^{th} SD in the lower and upper bounds in Corollary 2, respectively. Additionally, $\mathcal{S}(1, n, 1)$ exactly represents the constraint set for the single ensemble SD in Corollary 1 of a homogeneous ensemble with n loads.

V. Resource Allocation for Flexible Loads

We illustrate here how the ensemble constraint set in Section IV can be used by a Balancing Authority (BA) for resource allocation. Conceptually, our proposed method "projects" the needs of the BA onto the ensemble constraint set. We first describe how a BA can incorporate the ensemble constraint set into an optimization problem for resource allocation and then how a BA can determine its needs spectrally.

A. Allocation through projection

We denote by S^{BA} a SD that represents the BA's needs. Computation of this SD will be discussed in the next section. Resource allocation is performed by projecting the SD S^{BA} - what the grid needs - onto the Cartesian product $\times_{\ell=1}^{n_{\text{bin}}} \mathcal{S}(n_{\text{bin}}, n_\ell, \ell)$ - what the ensemble of loads can provide. The projection problem for a given value of n_{bin} is,

$$\min_{\{\Sigma_{\bar{P}^\ell}\}_1^{n_{\text{bin}}}} \quad \int_0^\infty \left(\Sigma^{\text{agg}}(\omega) - S^{\text{BA}}(\omega)\right)^2 d\omega \tag{36}$$
$$\text{s.t.} \quad \forall\ \ell \in \{1, \ldots, n_{\text{bin}}\}, \quad \Sigma_{\bar{P}^\ell} \in \mathcal{S}(n_{\text{bin}}, n_\ell, \ell),$$
$$\Sigma^{\text{agg}} = \sum_{\ell=1}^{n_{\text{bin}}} \Sigma_{\bar{P}^\ell} \tag{37}$$

Doing so allocates the needs of the grid across all loads; the loads will cover as much of the needs of the BA as they can while maintaining their QoS. In other words, the projection computes the regions shown in Figure 1 corresponding to each class of flexible loads.

Comment 1. *The problem (36) encapsulates resource allocation for both heterogeneous and homogeneous ensembles. If the ensemble of loads is heterogeneous, the BA can solve (36) twice. If the BA uses $\hat{n}$ bins, the first time the BA will set $n_{bin} = \hat{n}$ (upper bound) and the second time using $n_{bin} = 1$ (lower bound). Consequently, for a purely homogeneous collection the BA only needs to solve (36) once with $n_{bin} = 1$, and $n_1 = n$ (see Corollary 1). The LHS of (37) evaluated at the optimal solution of (36) then represents a bound on $S_{\bar{P}}$ for a heterogeneous ensemble, or is exactly $S_{\bar{P}}$ for a homogeneous ensemble.*

B. Computation

In order to solve the problem (36), the integrals appearing will have to be approximated. The approximation is obtained easily by going from continuous time t to discrete time index $k = 0, 1, \ldots$, so that functions such as $\tilde{P}(t)$ are replaced by their sampled-data analog $\tilde{P}_k := \tilde{P}(k\Delta t)$, with a constant sampling period Δt. In that case, the corresponding Fourier transforms are replaced by the DTFT (discrete time Fourier transform), and integrals of the form $\int_0^\infty S_{\tilde{P}}(\omega) d\omega$ are replaced by $\int_0^\pi S_{\tilde{P}}^D(\Omega) d\Omega$, where $S_{\tilde{P}}^D(\Omega)$ (D superscript for discrete/sampled time) is the SD of the sampled version of $\tilde{P}$. The resulting integral on $[0, \pi]$ can then be approximated through any numerical integration technique. For example, consider the N points indexed by j, $\Omega_j = 0,\ \frac{\pi}{N}, \ldots, \frac{N-1}{N}\pi$ then the decision variable for the finite dimensional problem is the SD $S_{\tilde{P}}^D(\Omega)$ evaluated at these N points. The resulting finite dimensional problem will be a strictly convex optimization problem, and hence can be easily solved using readily available software.

C. Spectral Needs of the BA

In the following we provide an example procedure for a BA to spectrally determine its needs, as illustrated in Figure 1. The BA first estimates the SD, Φ^{ND}, of its net

demand, i.e., demand minus renewable generation. It can estimate this quantity from time series data of demand and renewable generation, or through a modeling effort, or a combination thereof. The next step for the BA is to fit a parameterized model to Φ^{ND}, which is termed S^{ND}. All controllable resources, including generators, flywheels, batteries, and flexible loads, together have to supply Φ^{ND} (or its parameterized model S^{ND}). The third step is obtain the portion of S^{ND} that flexible loads have to provide (similar to what is shown in Figure 1) by "filtering" S^{ND}. Letting $F(j\omega)$ be an appropriate filter, then the SD of the signal the grid authority would like flexible loads to contribute is:

$$S^{\text{BA}}(\omega) = |F(j\omega)|^2 S^{\text{ND}}(\omega). \tag{38}$$

The quantity S^{BA} is the frequency domain analog of the reference signal $r(t)$ that will be asked from the loads, and will be referred to as the *reference SD* in the sequel. In the numerical example in this paper, we empirically estimate Φ^{ND} from time series data obtained from Bonneville Power Administration (BPA), a balancing authority in the Pacific Northwest, and then obtain S^{ND} by fitting an ARMA(p,q) model to Φ^{ND}. An example of S^{ND} is shown as the orange line in Figure 1, with S^{BA} being any of the shaded regions in Figure 1.

The procedure described above - for the BA to determine its needs in the spectral density domain - is *completely* independent from our characterization of capacity presented. The next step is to use the results of the procedure, the BA's spectral needs, to find the closest SD of the loads to the grid's need.

VI. Numerical Examples

An example of determining the capacity of a collection of HVAC systems in commercial buildings is illustrated in this section. For the numerical experiments we construct homogeneous and heterogeneous ensembles from two types of HVAC systems. The parameters of each type are displayed in Table I, and are chosen so that the HVAC systems are representative of those in small and large commercial buildings (hence the large and small superscripts in Table I).

To aid exposition of the results, we define the following power and energy capacity indices,

$$\zeta^{\text{P}} = \frac{\textbf{Pow}(\Sigma^{\text{agg}})}{\textbf{Pow}(S^{\text{BA}})} \times 100\%, \tag{39}$$

$$\zeta^{\text{E}} = \frac{\textbf{Eng}(\Sigma^{\text{agg}})}{\textbf{Eng}(S^{\text{BA}})} \times 100\%, \tag{40}$$

so as to show the percentage of power and energy capacity required by the BA that can be covered by the loads. In the above, the numerator SD abstractly represents the LHS of (37) at the optimal solution of problem (36). In all scenarios we solve the discrete time finite dimensional version of the convex optimization problem (36) with CVX [29]. All relevant simulation parameters, if not specified otherwise, can be found in Table I.

TABLE I: Simulation parameters

Par.	Unit	Value	Par.	Unit	Value
c_1^{Small}	kW	4	γ^{Small}	$1/\text{hour}$	2.78
c_2^{Small}	kW	0.8	β^{Small}	$^\circ C/\text{kWh}$	0.3597
$c_4^{\text{Small}}, c_4^{\text{Large}}$	$^\circ$C	1.11	γ^{Large}	$1/\text{hour}$	177.6
c_3^{Small}	kWh	0.5	β^{Large}	$^\circ C/\text{kWh}$	0.0450
c_1^{Large}	kW	40	T	Day	1
c_2^{Large}	kW	8	$\{\epsilon_i\}_{i=1}^4$	N/A	0.05
c_3^{Large}	kWh	5	δ	Sec	10

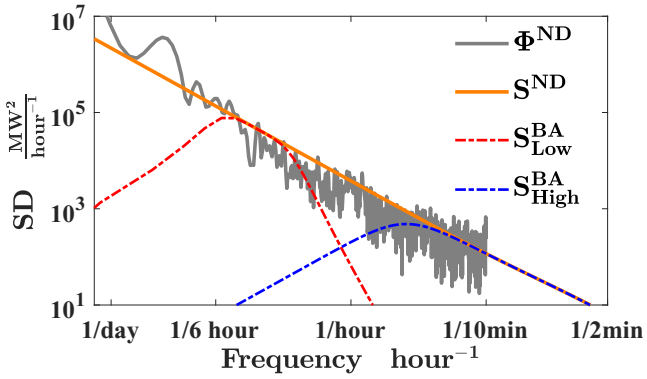


Fig. 3: Empirical net demand SD, modeled SD for BPA's net demand, and the two reference SD's for the high and low frequency passband.

A. BA's spectral needs

The net demand data is collected from BPA (a BA in the pacific northwest United States). The SD of the net demand is determined using the method described in Section V-C. The empirical net demand SD Φ^{ND} is estimated by the *pwelch* command in MATLAB. We choose an ARMA(2,1) model to fit to the empirically estimated SD. Since the estimate Φ^{ND} will cap out at the Nyquist frequency $1/10$min, we extrapolate the net demand SD to the higher frequencies. The empirical SD (denoted Φ^{ND}) and the extrapolated net demand SD (denoted S^{ND}) are shown in Figure 3.

We then choose two passbands to filter S^{ND}: (i) a low passband $[1/6,1/2]$ (1/hour) and (ii) a high passband $[1/30,1]$ (1/min). The results of "filtering" S^{ND} (see eq. (38)) are also shown in Figure 3. The low passband SD is termed $S^{\text{BA}}_{\text{Low}}$ and roughly corresponds to the region for TCLs in Figure 1. The high passband SD is termed $S^{\text{BA}}_{\text{High}}$ and roughly corresponds to the region for HVAC systems in Figure 1.

B. Meeting the BA's needs with large buildings

In this scenario we consider a homogeneous collection of large commercial buildings. The idea is to illustrate how many of these large commercial buildings would be required to meet the grids needs as a function of frequency. To do this, the two reference SDs obtained from the previous section are projected (by solving (36)) onto the *same* ensemble constraint set with a varying number of large commercial buildings.

TABLE II: Equivalent Battery Capacity

Num. of Loads	High Passband	Low Passband
3000	120 MW & 12.6 MWh	19 MW & 15 MWh
6000	120 MW & 12.6 MWh	37 MW & 30 MWh
9000	120 MW & 12.6 MWh	55 MW & 45 MWh
12000	120 MW & 12.6 MWh	73 MW & 60 MWh
15000	120 MW & 12.6 MWh	90 MW & 75 MWh

First, we show the results for $n = 15000$ large commercial buildings in Figure 3. For the reference SD with high passband, the loads are able to meet the requirements of the grid, with an aggregate power and energy capacity index of $\zeta^P \approx 100\%$ and $\zeta^E \approx 100\%$. However, for the reference SD with low passband the loads are unable to meet the grids needs, with an aggregate power and energy capacity index of $\zeta^P = 43\%$ and $\zeta^E = 26\%$.

For a varying number of loads the power and energy capacity index are recorded and plotted as a function of n in Figure 5. These results indicate that the BA would require $\approx$ 3000 large commercial buildings in the higher passband and $\approx$ 32000 large commercial buildings for the low passband, to fulfill its needs. We also see that the BA's energy capacity requirement is met with fewer loads (compared to its power capacity requirement) at the higher frequency range. In the lower frequency range, this conclusion is reversed.

Additionally, we compute the power and energy capacity values in Definition 2 for both the low and high passband over a subset of the varying number of loads considered in the preceding experiment. These values are displayed in Table II.

In summary, this numerical experiment suggests that the commercial HVAC systems considered here are more suitable to assist the grid by tracking signals with SD in the higher passband. If a BA wanted the commercial HVAC systems to track reference signals in the lower frequency range, than significantly more large commercial HVAC systems would be required.

C. Effects of heterogeneity

We present a numerical experiment to illustrate the bounds in Corollary 2. We consider $n_{\text{bin}} = 2$ bins with, $n_1 = n_S = 900$ small and $n_2 = n_L = 2100$ large commercial buildings. We solve the problem (36) twice, with the appropriate parameters detailed in Comment 1 for both times, so to obtain the upper and lower bounds in Corollary 2. The SD that represents the grid's needs in this experiment is the same SD with high passband for the experiments in section VI-B.

The results for this scenario are shown in Figure 6, where 'upper bound' and 'lower bound' represent the upper and lower bounds in Corollary 2, respectively. Additionally, the shaded region in Figure 6 represents the region where the true ensemble SD lives. For this example, the results indicate that the power capacity index is bounded as $17\% \leq \zeta^P \leq 69\%$ and the energy capacity index is bounded as $30\% \leq \zeta^E \leq 86\%$.

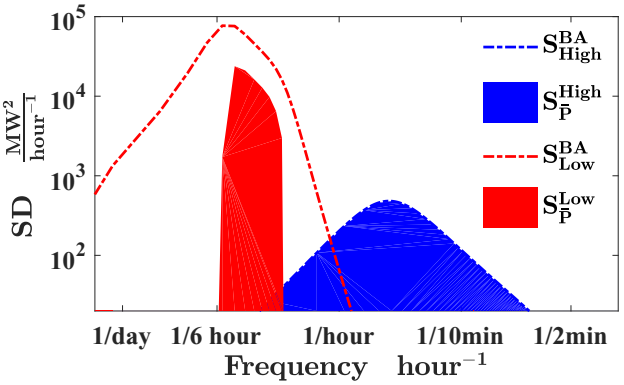


Fig. 4: The two reference SDs and the ensembles SD (boundary of the shaded regions) obtained by solving (36) for a homogeneous collection of $n = 15000$ loads.

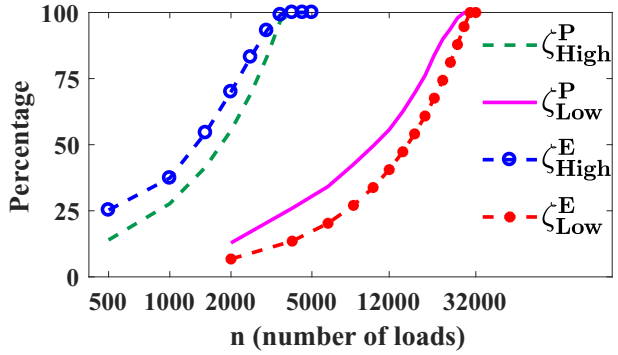


Fig. 5: The aggregate power and energy capacity index for the two different reference SDs with a homogeneous collection of loads plotted against n.

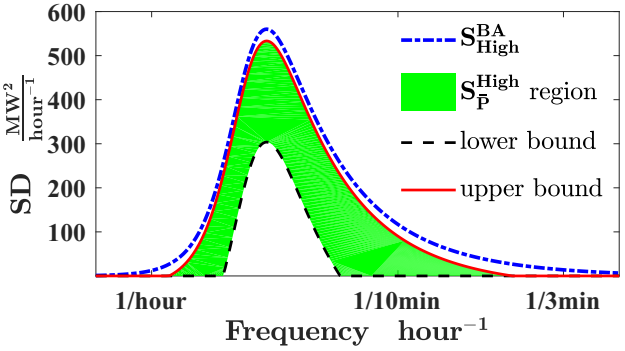


Fig. 6: The upper and lower bounds on the ensemble SD for a heterogeneous ensemble.

D. Time domain results

To illustrate how our definition of capacity given in Definition 2 is useful in the time domain, we estimate the probabilities (28)-(29) for a given ϵ and SD $S \in \mathcal{S}_\epsilon$. To do this, we use the SD $S^{\text{High}}_{\bar{P}}$ that was obtained from the

numerical experiment in section VI-B for $n = 4200$ large commercial HVAC systems. From the SD we independently generate $M = 1 \times 10^4$ time series that are indexed by i and each denoted as $\{\tilde{P}_i(t_k)\}_k$. We use the notation $\tilde{P}$, since we will show the results for the time series (and consequently the bounds) normalized by n. From each of these time series, the following estimate of the probability (28) is obtained:

$$\hat{\mathsf{P}}_i\left(\left|\tilde{P}\right| \geq \mathbf{Pow}(S)\right) = \frac{1}{N}\sum_{k=1}^{N} \mathbf{I}\left(\left|\tilde{P}_i(t_k)\right| \geq \mathbf{Pow}(S)\right), \tag{41}$$

where $\mathbf{I}(\cdot)$ is the indicator function. The same computation is done for (29) as well, where the time series $\{\tilde{E}_i(t_k)\}_k$ is obtained by using $\{\tilde{P}_i(t_k)\}_k$ and the dynamics (3). Our overall estimate is then the average over all of the $\hat{\mathsf{P}}_i(\cdot)$, and for the computation (41) this is

$$\hat{\mathsf{P}}\left(\left|\tilde{P}\right| \geq \mathbf{Pow}(S)\right) = \frac{1}{M}\sum_{i=1}^{M} \hat{\mathsf{P}}_i\left(\left|\tilde{P}\right| \geq \mathbf{Pow}(S)\right).$$

Similarly, the same computation is done for (29) as well. The estimates $\hat{\mathsf{P}}_i(\cdot)$ are each shown in Figures 7-8 as a histogram. The average estimates are displayed in each figure and both are well below their respective thresholds. In all scenarios, the quantities $\hat{\mathsf{P}}_i(\cdot)$ are never above the respective threshold.

In addition, we also plot two sample paths of the time series, used to compute the probabilities, for all four of the QoS metrics. The result of this is shown in Figure 9, where we also plot the capacity bounds from Definition 2 and the equipment bounds c_1, c_2, c_3, c_4 from (17)-(20). Similarly, the time series for the rate and temperature quantities are obtained by using the time series $\{\tilde{P}_i(t_k)\}_k$ and the respective linear system. Since the SD $S_{\tilde{P}}^{\text{High}}$ is for n homogeneous loads, we normalize the time series and capacity bounds by n so that the plots in Figure 9 can be interpreted for a single load. We show the sample paths for a subset of the total horizon for clarity. Long run information is captured in the histograms, since each data point in the histogram is over the entire horizon.

For all four QoS, the two time series are typically well within capacity as defined in Definition 2 (the black horizontal lines in the figures) and any violations satisfy the desired probabilistic bound. In fact, in all scenarios Definition 2 is tighter than the equipment bounds, which for some of the QoS are quite loose. The equipment bounds play a large part in the capacity characterization of some past works, e.g., [9]. The numerical experiment here illustrates the dangers in quoting directly equipment bounds as capacity: in this simple experiment the actual energy capacity is $\approx$ 3 KWh for a single load, and not 5 KWh as the equipment bound indicates.

VII. Summary and Conclusion

Our characterization of virtual energy storage capacity of flexible loads allows a balancing authority to quantify which loads can contribute how much to mitigating supply demand mismatch. In contrast to past work, our method can be used for long term planning. The key insight is to characterize capacity through statistics of the reference signal rather than on specific instances of the signal. This framework also has another benefit: it enables us to define power and energy capacity of virtual energy storage in language of real energy storage: kW's and kWh's.

An open problem is to extend the framework to non-stationary statistics of the grid's demand-supply mismatch and/or loads' demand deviation. Another open problem is the extension to the case of on/off loads. A key challenge for on/off loads will be to characterize their cycling constraint.

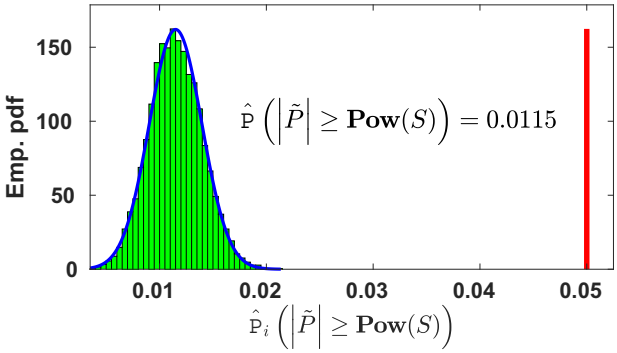


Fig. 7: Empirically estimated probability bounds for the power capacity with $\epsilon_1 = 0.05$.

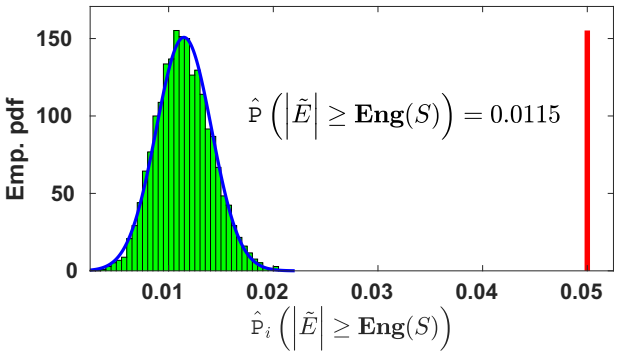


Fig. 8: Empirically estimated probability bounds for the energy capacity with $\epsilon_4 = 0.05$.

References

[1] P. Barooah, *Smart Grid Control: An Overview and Research Opportunities*. Springer Verlag, 2019, ch. Virtual energy storage from flexible loads: distributed control with QoS constraints, pp. 99–115.

[2] N. J. Cammardella, R. W. Moye, Y. Chen, and S. P. Meyn, "An energy storage cost comparison: Li-ion batteries vs Distributed load control," in *2018 Clemson University Power Systems Conference (PSC)*, Sep. 2018, pp. 1–6.

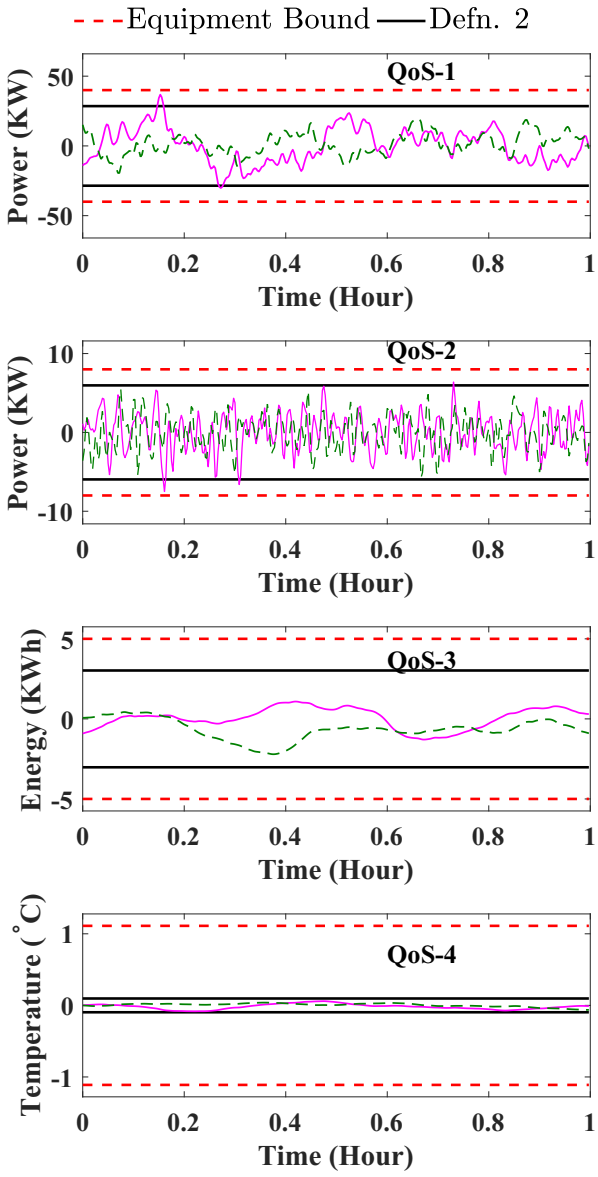


Fig. 9: Two sample paths of time domain trajectories for the SD $S_{\bar{P}}^{\text{High}}$ plotted with the capacity bounds from Definition 2 and the equipment bounds, all normalized by the number of loads n.

[3] A. Coffman, A. Bušić, and P. Barooah, "A study of virtual energy storage from thermostatically controlled loads under time-varying weather conditions," in *5th International Conference on High Performance Buildings*, July 2018, pp. 1–10.

[4] M. Liu, S. Peeters, D. S. Callaway, and B. J. Claessens, "Trajectory tracking with an aggregation of domestic hot water heaters: Combining model-based and model-free control in a commercial deployment," *IEEE Transactions on Smart Grid*, 2019.

[5] H. Hao, A. Kowli, Y. Lin, P. Barooah, and S. Meyn, "Ancillary service for the grid via control of commercial building HVAC systems," in *American Control Conference*, June 2013, pp. 467–472.

[6] A. Aghajanzadeh and P. Therkelsen, "Agricultural demand response for decarbonizing the electricity grid," *Journal of Cleaner Production*, vol. 220, pp. 827 – 835, 2019.

[7] Y. Chen, A. Bušić, and S. Meyn, "State estimation for the individual and the population in mean field control with application to demand dispatch," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1138–1149, March 2017.

[8] Z. E. Lee, Q. Sun, Z. Ma, J. Wang, J. S. MacDonald, and K. Max Zhang, "Providing Grid Services With Heat Pumps: A Review," *ASME Journal of Engineering for Sustainable Buildings and Cities*, vol. 1, no. 1, 01 2020, 011007.

[9] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent, "Aggregate flexibility of thermostatically controlled loads," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 189–198, Jan 2015.

[10] H. Hao, D. Wu, J. Lian, and T. Yang, "Optimal coordination of building loads and energy storage for power grid and end user services," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2017.

[11] A. Coffman, N. Cammardella, P. Barooah, and S. Meyn, "Aggregate capacity of TCLs with cycling constraints," *arXiv preprint arXiv:1909.11497*, 2019.

[12] R. Yin, E. C. Kara, Y. Li, N. DeForest, K. Wang, T. Yong, and M. Stadler, "Quantifying flexibility of commercial and residential loads for demand response using setpoint changes," *Applied Energy*, vol. 177, pp. 149 – 164, 2016.

[13] H. Hao, J. Lian, K. Kalsi, and J. Stoustrup, "Distributed flexibility characterization and resource allocation for multi-zone commercial buildings in the smart grid," in *2015 54th IEEE Conference on Decision and Control (CDC)*, Dec 2015, pp. 3161–3168.

[14] J. T. Hughes, A. D. Domnguez-Garca, and K. Poolla, "Virtual battery models for load flexibility from commercial buildings," in *2015 48th Hawaii International Conference on System Sciences*, Jan 2015, pp. 2627–2635.

[15] D. Wang, K. Ma, P. Wang, J. Lian, and D. J. Hammerstrom, "Frequency-domain flexibility characterization of heterogeneous end-use loads for grid services," in *2020 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2020, pp. 1–5.

[16] S. Kundu, K. Kalsi, and S. Backhaus, "Approximating flexibility in distributed energy resources: A geometric approach," in *2018 Power Systems Computation Conference (PSCC)*, June 2018, pp. 1–7.

[17] F. Lin and V. Adetola, "Flexibility characterization of multi-zone buildings via distributed optimization," in *2018 Annual American Control Conference (ACC)*, June 2018, pp. 5412–5417.

[18] F. L. Müller, O. Sundström, J. Szabó, and J. Lygeros, "Aggregation of energetic flexibility using zonotopes," in *2015 54th IEEE Conference on Decision and Control (CDC)*, 2015, pp. 6564–6569.

[19] S. Barot and J. A. Taylor, "A concise, approximate representation of a collection of loads described by polytopes," *International Journal of Electrical Power & Energy Systems*, vol. 84, pp. 55 – 63, 2017.

[20] M. S. Nazir, I. A. Hiskens, A. Bernstein, and E. Dall'Anese, "Inner approximation of minkowski sums: A union-based approach and applications to aggregated energy resources," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 5708–5715.

[21] P. Barooah, A. Bušić, and S. Meyn, "Spectral decomposition of demand side flexibility for reliable ancillary service in a smart grid," in *48th Hawaii International Conference on Systems Science (HICSS)*, January 2015, pp. 2700–2709, invited paper.

[22] E. M. Krieger, "Effects of variability and rate on battery charge storage and lifespan," Ph.D. dissertation, Princeton University, 2013.

[23] J. Apt, "The spectrum of power from wind turbines," *Journal of Power Sources*, vol. 169, no. 2, pp. 369 – 374, 2007.

[24] Y. Lin, P. Barooah, S. Meyn, and T. Middelkoop, "Experimental evaluation of frequency regulation from commercial building HVAC systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 776 – 783, 2015.

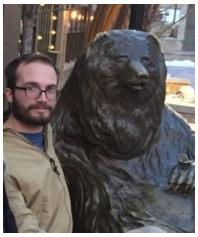
[25] B. Kirby, "Ancillary services: Technical and commercial insights," Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, Tech. Rep., 2007, prepared for Wärtsilä North America Inc.

[26] A. R. Coffman, Z. Guo, and P. Barooah, "Capacity of flexible loads for grid support: statistical characterization for long term planning," in *2020 American Control Conference (ACC)*, July 2020, pp. 533–538.

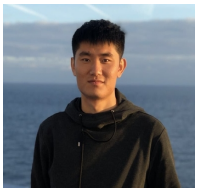
[27] S. Huang and D. Wu, "Validation on aggregate flexibility from residential air conditioning systems for building-to-grid integration," *Energy and Buildings*, vol. 200, pp. 58 – 67, 2019.

[28] B. Hajek, *Random processes for engineers*. Cambridge university press, 2015.

[29] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," http://cvxr.com/cvx, Feb. 2011.



Austin R. Coffman received both his B.S. (Aerospace Engineering) and M.S. (Mechanical Engineering) from the University of Florida in 2016 and 2017, respectively. He began his PhD studies at the University of Florida in May 2017. After graduation he hopes to work in industry solving complex engineering problems. He is the recipient of the Iva and Norman Tuckett FINS fellowship in 2017, and a Chateaubriand fellow for the 2019-2020 academic year. His academic interests range from optimization and reinforcement learning to Markov chains



Zhong Guo received his B.E. degree in Mechanical Design and Automation from Henan University of Science and Technology in 2017 and M.S. degree in Mechanical Engineering from University of Florida in 2019. He is currently a PhD student in Mechanical Engineering at University of Florida. His research interests include system identification, optimization, and reinforcement learning.

Prabir Barooah received his Ph.D. in 2007 from the University of California, Santa Barbara. From 1999 to 2002 he was a research engineer at United Technologies Research Center, East Hartford, CT. He received the M. S. degree in Mechanical Engineering from the University of Delaware in 1999 and the B.Tech degree in Mechanical Engineering from the Indian Institute of Technology, Kanpur, in 1996. He was born in Jorhat, Assam. Dr. Barooah is the winner of the ASEE SE Sections Outstanding Researcher Award (2013), NSF CAREER award (2010), General Chairs' Recognition Award for Interactive papers at the 48th IEEE Conference on Decision and Control (2009), best paper award at the 2nd Int. Conf. on Intelligent Sensing and Information Processing (2005), and NASA group achievement award (2003).