

Controlling aggregates of deferrable loads for power system regulation

Federico Bliman, Andres Ferragut and Fernando Paganini
Universidad ORT Uruguay

Abstract—This paper considers the role of a demand aggregator that manages a large number of consumer loads, with the objective of participating in the frequency regulation market. The key feature to be exploited is load deferrability in time, which enables the aggregator to adapt the consumption profile and thus reduce its own consumption of regulation, and even be a provider of regulation services to others. Rather than a microscopic model that considers individual loads and their scheduling, we consider here a macroscopic viewpoint drawn from fluid models of queueing systems. Here the state variable is the quantity of currently dispatchable loads, and the control input dictates the fraction of those which are currently active. The dynamics is a simple nonlinear ODE that allows us to design a controller with a feedforward term to track an external regulation reference, and a feedback term to reduce the impact of random load oscillations, in an agnostic way to the microscopic scheduling. The performance of this controller is evaluated by simulation using practical regulation signals.

I. INTRODUCTION

Matching demand and supply in electric power systems is a complex task, given the well-known limitations in storage. Typical solutions involve markets at different time-scales (day ahead, hour ahead) in which quantities of power are traded to match demand forecasts, and ancillary services (reserves and regulation, see [9]) that are summoned close to real-time by the system operator (SO) to achieve the necessary balance. Of these services, frequency *regulation* occurs at the shorter time basis, with the objective of keeping frequency in its nominal value (50 or 60Hz). Providers of such regulation service receive a SO signal every few seconds, and must alter their power to track this reference.

This service has been traditionally provided by fast responding generators like hydro or gas turbines, which are already in the system providing some nominal power, and offer variations around it. For example, a seller of 500 MW of energy who commits 100 MW of regulation must be able to quickly vary its generation output in the range [400MW, 600MW], following a normalized SO command in the range $[-1, 1]$; when receiving for instance the signal value 0.5 it moves to deliver 550 MW of power.

The requirement for regulation, which traditionally arose from load uncertainty, is increasing due to the rise of renewable energy sources. These are non-dispatchable and face significant real-time variability; their impact on regulation requirements is expected to be significant [7]. A natural alternative to supply-side regulation is to exploit the response capabilities of the demand side, which are becoming available with the deployment of smart grid technologies [15].

Indeed, the intrinsic *deferrability* of some loads (heating, AC, EV battery charging, etc.) may help with regulation. A fully distributed proposal in this vein is [17], where appliances individually respond to measured frequency deviations.

Now, to have impact on frequency regulation similar to a medium sized hydro turbine it is necessary to control thousands of individual loads, so it makes sense to frame the problem at the level of a load *aggregator* (e.g. [2], [4]). By making load dispatch decisions for a cluster of loads, such aggregators could provide similar ancillary services as generators provide today.

Some recent work that considers this approach is now briefly surveyed. In [5], [6], [10] the focus is on Thermostatically Controlled Loads (TCLs) which maintain temperatures using a hysteretic ON-OFF control. Manipulating such cycle within the temperature constraints allows for flexibility in power consumption, exploited to provide ancillary services. In [5] a collection of TCLs is characterized by an equivalent battery model. In [6] the potential of this approach is demonstrated with practical data from California. More relevant to this paper is another line of work [11], [18] that exploits the time deferrability of generic loads, characterized by arrival times, deadlines, and power and energy requirements. In particular [18] investigates different options for *scheduling* such deferrable loads, comparing classical approaches from processor scheduling (earliest deadline first, least laxity first, see e.g. [8]) with a model predictive control proposal tailored to the power setting. In [11] the authors attempt to characterize the aggregate flexibility provided by such load arrival profile, again in terms of electricity storage.

The present paper also employs an aggregate characterization, in this case with techniques from large-scale queueing systems¹. We employ a *fluid* ordinary differential equation model to track the population of loads under the potential of being serviced, and the macroscopic control decision of load deferral becomes a continuous control input. This allows standard methods from control theory to be brought to bear to analyze the regulation capabilities of such an aggregate. We start in Section II with a deterministic ODE model, and derive a suitable feedforward controller to achieve regulation tracking. Our model is refined in Section III by incorporating noise fluctuations, and a feedback control is designed to reduce them. In Section IV we combine noise reduction and reference tracking into an overall control design, whose performance is evaluated with real-life regulation signals from [14]. Conclusions are given in Section V.

Research supported in part by ANII under grant FSE_1_2014_1_102426. E-mail: {bliman, ferragut, paganini}@ort.edu.uy.

¹See also [1] for queueing analysis of aggregation, without deferrability.

II. REGULATION TRACKING THROUGH LOAD DEFERRING

Consider a demand aggregator that manages a large set of loads on the power system. This agent should have an estimation of the aggregate power demand profile for a time period (e.g. the following day), and will use the forecast to purchase this average power in the forward market. Regulation comes into play to deal with real-time variations around these predicted values.

Assume first that none of the loads are deferrable: in that case the aggregator has a randomly varying load profile that deviates from the forecast and so it becomes a *consumer* of regulation services, which must be obtained through the SO. Suppose instead that a portion of the loads is deferrable in time; this flexibility can be exploited to align as much as possible the consumed power to the forecast, reducing the regulation power requirement. When load deferrability is large the aggregator could eliminate the need to purchase regulation, as long as the total energy requirement has no bias. Furthermore, it could exploit the flexibility to become a *supplier* of regulation services to others in the network.

We begin in this section with the last scenario. We assume load demands arrive according to a predictable profile, and will use the option of deferring service to match the power consumption to an external reference. Suppose demands arrive at the aggregator at a rate λ , each requiring a mean amount of energy Q_0 and power output p_0 . In real life such quantities will be time-varying, following a time-of-day load profile; since regulation occurs at a fast time scale we will ignore this issue, working over a time horizon where these quantities can be considered constant.

Let $n(t)$ denote the number of loads at disposal in the system to be serviced. Each load requires an average service time $\tau = \frac{Q_0}{p_0}$ from the system, provided it is serviced at full power. To characterize the *deferrability* of loads we introduce a time window parameter h that represents the mean deadline for service. The loads are deferrable if this time exceeds the minimum service time, i.e.

$$h > \tau = \frac{Q_0}{p_0}.$$

A. Fluid model of load deferrals

The load aggregator may choose to serve the loads at a fraction of the power, or alternatively, serve only a fraction of the loads and defer the others. Our macroscopic model will be agnostic to these details. Let $u(t) \in [0, 1]$ be the service fraction. A simple model for the evolution of the system is the following first-order state-space system:

$$\dot{n}(t) = \lambda - \frac{1}{\tau}n(t)u(t), \quad (1a)$$

$$p(t) = p_0n(t)u(t). \quad (1b)$$

Here, the number of dispatchable loads grows as new demands arrive, whereas the second term accounts for service completions: $n(t)u(t)$ is the number of active loads, each completed at rate $1/\tau$. The instantaneous power consumption is then the average power times the number of *active* loads. A more complete justification of this model from a queueing

perspective is given in Section III, but we begin here by analyzing some of its properties.

Let us analyze first this system for a fixed value of $u(t) = u^*$. Imposing equilibrium in (1) we obtain:

$$n^* = \frac{\lambda\tau}{u^*}, \quad p^* = p_0\lambda\tau = \lambda Q_0. \quad (2)$$

In equilibrium, the amount of serviceable loads in the system is increased by the deferral action u^* . Note however that the average power output from the system is independent of u^* , and equal to the average energy per request times the frequency of requests. This amount of power is the *predictable* component of the demand and can be purchased in advance for the time-period considered.

A second conclusion of (2) is that, if only a fraction u^* of the loads are active at a given time, the average time spent by each request in the system is τ/u^* . We would like this to be below the deadline h , which imposes a first condition on the choice of u^* :

$$u^* > \frac{\tau}{h} = \frac{Q_0}{p_0h} =: \eta.$$

Here $\eta \in [0, 1]$ is a measure of deferrability of the loads (more deferrability for smaller η).

We would like to analyze this system with input $u(t)$ and output $p(t)$, in order to understand which class of signals can be tracked by using the power fraction as a control input, while keeping with the deadline constraint. We do so by linearizing the system around the equilibrium point n^*, p^*, u^* . Denoting by $\delta n, \delta p$ and δu the deviation of variables from equilibrium, the linearized dynamics are:

$$\dot{\delta n} = -\frac{1}{\tau}u^*\delta n - \frac{1}{\tau}n^*\delta u, \quad (3a)$$

$$\delta p = p_0(u^*\delta n + n^*\delta u). \quad (3b)$$

The transfer functions associated with the above system in the Laplace domain can be readily computed to yield:

$$G_{un}(s) := \frac{\hat{\delta n}}{\hat{\delta u}} = \frac{-\frac{n^*}{\tau}}{s + \frac{u^*}{\tau}},$$

$$G_{up}(s) := \frac{\hat{\delta p}}{\hat{\delta u}} = \frac{p_0n^*s}{s + \frac{u^*}{\tau}};$$

and will be useful in the subsequent analysis.

B. Tracking a reference signal

The control objective is for the power output to track a *reference signal* $r(t)$, which is of the form $r(t) = r^* + \delta r(t)$. Here r^* is the nominal power consumption negotiated with the grid in advance, thus $r^* = p^* = \lambda Q_0$ assuming the forecast for the mean load is accurate. The variation to track will have the form

$$\delta r(t) = \theta r^* \rho(t), \quad (4)$$

where $\theta \in [0, 1]$ is the fraction of the nominal consumption power that is offered for regulation, and $\rho(t) \in [-1, 1]$ is the *regulation signal* sent by the SO, as discussed in Section I.

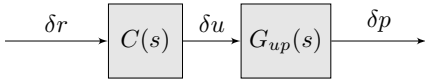


Fig. 1. Controller design for tracking the regulation signal

A first proposal to achieve the desired tracking is through a feedforward controller, as depicted in Fig. 1. Observe that by choosing $C(s)$ as:

$$C(s) = \frac{1}{G_{up}(s)} = \frac{s + \frac{u^*}{\tau}}{p_0 n^* s} = \frac{1}{p_0 n^*} \left(1 + \frac{u^*}{\tau s} \right), \quad (5)$$

the linearized system should be able to match deviations in the regulation signal δr .

Noting that $p_0 n^* = p_0 \lambda \tau / u^* = r^* / u^*$, the above proportional-integral law can be expressed in the time domain as follows:

$$\delta u(t) = \frac{u^*}{r^*} \left[\delta r(t) + \frac{u^*}{\tau} \int_0^t \delta r(w) dw \right].$$

Now replacing with (4) and noting $\tau = \eta h$ leads to

$$\delta u(t) = u^* \theta \left[\rho(t) + \frac{u^*}{\eta h} \int_0^t \rho(w) dw \right]. \quad (6)$$

Of course, the control input is subject to the saturation constraint $u(t) \in [0, 1]$ (or even $u(t) > \eta$ for all t to be covered for deadlines). Consequently, the above expression indirectly constrains the class of regulation signals $\rho(t)$ that our system can track; in particular $\rho(t)$ must have mean-value zero, otherwise the integral term will necessarily lead to saturation; in fact $\rho(t)$ should not have a persistent sign for too long, in relation to the mean deadline h .

To illustrate the behavior of the proposed controller, we simulated the system driven by a real-life regulation signal $\rho(t)$ taken from [14]. We considered a random profile of loads arriving at the aggregator as a Poisson process, of rate $\lambda = 4$ jobs per minute, with (exponentially distributed) energy request of mean $Q_0 = 2$ kWh, and power $p_0 = 1$ kW when serviced. The fraction of serviced loads $u(t)$ is driven by (6), the output of the linear controller around a fixed equilibrium value of $u^* = 0.5$ and includes the effect of the saturation constraint $u(t) \in [0, 1]$. Of the possible scheduling policies based on $u(t)$ (described in more detail in Section III) we chose here a *random* selection for the fraction of served loads; however this choice has minimal impact.

Simulation results are shown in Fig. 2, corresponding to $\theta = 0.5$; we see that the aggregator output closely matches the regulation request. Thus the aggregator can offer 50% regulation around the average power $p^* = 240$ kW.

We do notice, however, some tracking errors which are attributed to the randomness in the system. Depending on their entity, such errors may result in practice in penalties for not following the correct profile [12]. It is thus important to understand these fluctuations, which are not captured by our ODE model (1). In the next section we tackle this issue with tools from queueing theory analysis.

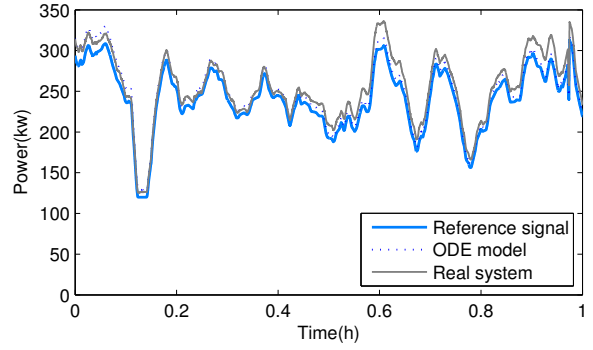


Fig. 2. Tracking of a power regulation signal via service deferrals.

III. IMPACT OF UNCERTAINTY ON REGULATION

To incorporate randomness we will consider a more descriptive stochastic model. Assume that loads arrive at the aggregator as a Poisson process of intensity λ , each requesting a random amount of energy following an exponential distribution of mean Q_0 . If the system feeds loads at a fraction u^* of the maximum power, the number of loads present in the system $N(t)$ behaves as an $M/M/\infty$ queue [16] with arrival rate λ and service completion rate:

$$\mu = \frac{p_0 u^*}{Q_0} = \frac{u^*}{\tau}.$$

If the arrival rate λ is large, the random process $N(t)$ can be well approximated by a deterministic trajectory following the ODE (1); perturbations around the equilibrium value can be approximated by a random noise input to (1).

In mathematical terms, let L a scaling factor, and $N_L(t)$ represent the random process with arrival rate $L\lambda$. Provided $N_L(0)/L \rightarrow n(0)$, the rescaled random process satisfies:

$$\bar{N}_L(t) = \frac{N_L(t)}{L} \rightarrow_{L \rightarrow \infty} n(t)$$

uniformly over compact sets, where $n(t)$ is the solution of (1) with fixed $u = u^*$ and initial condition $n(0)$ [16].

As for variability around equilibrium, a diffusion approximation can be performed. If the initial condition satisfies $N_L(0)/L = n^*$, then $n(t) = n^* \forall t$ and the random process

$$\delta N_L(t) = \frac{N_L(t) - Ln^*}{\sqrt{L}}$$

converges in distribution [16] to the solution of the following stochastic differential equation:

$$\dot{\delta n} = -\mu \delta n + \sqrt{2\lambda} v_0(t),$$

where $v_0(t)$ is stationary white noise of unit power spectral density². The factor 2λ in the net noise power spectrum comes from the two sources of variability, arrivals and departures, each contributing a term λ , associated with the Poisson arrival and departures. This type of variability analysis has been used before by the authors to track population profiles

²A more formal version is $d(\delta n) = -\mu \delta n dt + \sqrt{2\lambda} dW$, where $W(t)$ is standard Brownian motion. Such δn constitutes an Ornstein-Uhlenbeck process. For second order analysis, however, the above description suffices.

in P2P network applications [3], [13], which have the same type of fluid model as our present system.

A. The impact of uncertainty on regulated power

The presence of noise at the input of our system means that the output power will deviate from its intended value. To evaluate this impact we begin with the situation of a *fixed* deferral policy $u \equiv u^*$, which in the absence of noise would produce a constant consumption of power $p \equiv p^*$. In the presence of noise, the output variations in power are characterized by incorporating the noise into (3):

$$\delta \dot{n} = -\frac{1}{\tau} u^* \delta n + v, \quad (7a)$$

$$\delta p = p_0 u^* \delta n; \quad (7b)$$

where $v(t)$ is white noise of power spectral density $S_v(\omega) \equiv 2\lambda$. The transfer function from the noise input to the output δp is given by

$$G_{vp}(s) = \frac{p_0 u^*}{s + \frac{u^*}{\tau}}. \quad (8)$$

The noise variance at the output of this stable filter can be found (see e.g., [19]) from the corresponding \mathcal{H}_2 norm:

$$\begin{aligned} E[(\delta p)^2] &= \int_{-\infty}^{\infty} |G_{vp}(j\omega)|^2 S_v(\omega) \frac{d\omega}{2\pi} \\ &= \|G_{vp}(s)\|_{\mathcal{H}_2}^2 2\lambda \\ &= (p_0 u^*)^2 \frac{\tau}{2u^*} 2\lambda = p^* p_0 u^*. \end{aligned} \quad (9)$$

A first conclusion is that choosing $u^* < 1$ can reduce the variability of the instantaneous power consumption $p(t)$, with respect to the case of non-deferrable loads. Here we see the favorable impact of the flexibility of deferring loads in smoothing out the power profile, even if this deferral is chosen in a fixed, uncontrolled way. It appears one should work with u^* as small as possible, but of course this runs against the deadline constraint expressed in mean value by $u^* > \eta$; indeed, as $u^* \rightarrow \eta$ there will be increased probability of loads missing their deadlines.

To further optimize the system to minimize this possibility, the exact scheduling of the loads must be taken into account. Proposals such as earliest deadline first (EDF) or least laxity first (LLF) [8] should be incorporated to cope with the deadlines. Here we analyze three possibilities:

- *Equal sharing*: The load aggregator chooses to serve all present jobs with power $p_0 u^*$. While this may be an infeasible policy in practice, it serves as a reference point for analysis. It corresponds to the Processor Sharing discipline of queueing theory.
- *Random*: The load aggregator chooses a fraction u^* of the available jobs at random. This policy is very easy to implement in a decentralized environment, by distributing the value of u^* and the loads choose whether to become active or not based on a local random variable.
- *Least-Laxity-First (LLF)*: Here, the load aggregator chooses a fraction u^* of the loads ordered by decreasing *laxity*, i.e. the remaining amount of time before the job needs to become active in order to meet its deadline [8].

We simulated the system using these scheduling policies and different values of u^* . In Fig. 3 we show confidence intervals (obtained through multiple runs) for the standard deviation of the measured noise power, and compare them with the theoretical value. We observe that power variability is oblivious to the exact scheduling performed, and correctly captured by the analysis. In other words, the main knob a load aggregator has to reduce variability in power consumption is reducing the fraction of serviced loads u^* .

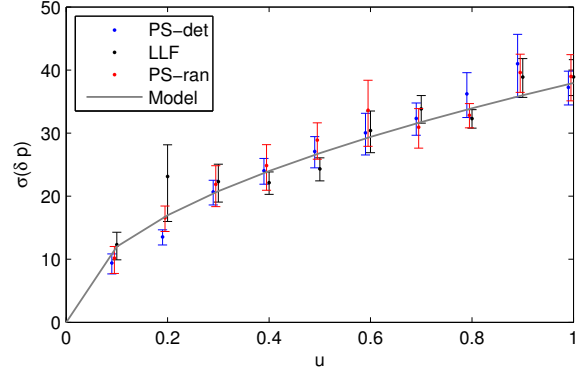


Fig. 3. Variability of power output as a function of u^* , for different scheduling policies.

Scheduling *does* have an impact, however, in meeting the load deadline requirements. In Fig. 4, we plot the fraction of loads that finish with expired deadlines for the different scheduling policies. The equal sharing and random policies behave in the same way, with a smooth decrease of expired jobs as a function of u^* . In the case of LLF, which takes deadlines explicitly into account, there is a sharp decrease in expired jobs after $u^* > \eta$. This means that, provided the system can implement a suitable scheduling policy, the value of u^* can be reduced towards the minimum η , thereby reducing regulation requirements.

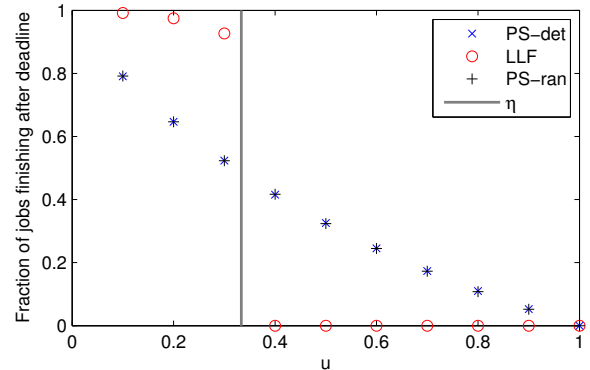


Fig. 4. Fraction of jobs completed after the deadline with varying u^* , and different scheduling policies.

B. Optimizing regulation requirements through feedback

In our analysis of system noise so far we only considered a fixed, static choice of the load deferral fraction, captured by the parameter u^* . However, further improvements could

be sought by controlling the variable $u(t)$ in feedback, in this case using as natural measurement the state $n(t)$.

We now analyze such a scenario. Consider again the linearized system from (3) restoring the input δu and adding the noise v , with output δp :

$$\dot{\delta n} = -\frac{1}{\tau}u^*\delta n - \frac{1}{\tau}n^*\delta u + v, \quad (10a)$$

$$\delta p = p_0(u^*\delta n + n^*\delta u); \quad (10b)$$

where again $v(t)$ is white noise of power spectrum 2λ .

Since we are working with stochastic noise and signal variances for performance, a natural feedback design strategy is \mathcal{H}_2 -optimal control, seeking to minimize for instance

$$J := E[(\delta p)^2 + \beta(p^*)^2(\delta u)^2], \quad (11)$$

weighted sum of the regulation error variance with a penalty on control effort. The latter penalization is natural to induce the control input to stay within its saturation limits³.

Noting that we are in a state-feedback situation, the optimal \mathcal{H}_2 control [19] will have the form of a static state feedback $\delta u = -K\delta n$; in this scalar case we can work directly with the gain K , more conveniently written as

$$K = \frac{u^*}{n^*}a \quad (12)$$

with the parameter $0 \leq a < 1$.

Substituting the feedback law in the linearized state-space model (10), we arrive at the closed loop:

$$\dot{\delta n} = -\frac{u^*}{\tau}(1-a)\delta n + v, \quad (13a)$$

$$\delta p = p_0u^*(1-a)\delta n, \quad (13b)$$

$$\delta u = -\frac{u^*}{n^*}a\delta n. \quad (13c)$$

The closed loop transfer function from noise to state is

$$G_{vn}^a(s) = \frac{1}{s + \frac{u^*}{\tau}(1-a)}, \quad (14)$$

from where we compute the stationary state variance

$$E[(\delta n)^2] = \|G_{vn}^a(s)\|_{\mathcal{H}_2}^2 2\lambda = \frac{\tau\lambda}{u^*(1-a)} = \frac{n^*}{1-a}.$$

Expressions for the variances in (11) follow from (13):

$$E[(\delta p)^2] = (p_0u^*(1-a))^2 \frac{n^*}{1-a} = \frac{(p^*)^2}{n^*}(1-a), \quad (15)$$

$$E[(\delta u)^2] = \frac{(u^*)^2 a^2}{n^*(1-a)}. \quad (16)$$

Therefore our cost from (11) becomes

$$J = \frac{(p^*)^2}{n^*} \left[(1-a) + \beta(u^*)^2 \frac{a^2}{1-a} \right].$$

The above expression clearly expresses the tradeoff between regulation and control effort as a function of the gain parameter $a \in [0, 1)$. If $a = 0$ there is no feedback control and we are back in the situation of Section III-A, with the

³Other control designs that explicitly incorporate L_∞ bounds on u would be more precise, we choose this version for simplicity. β is dimensionless.

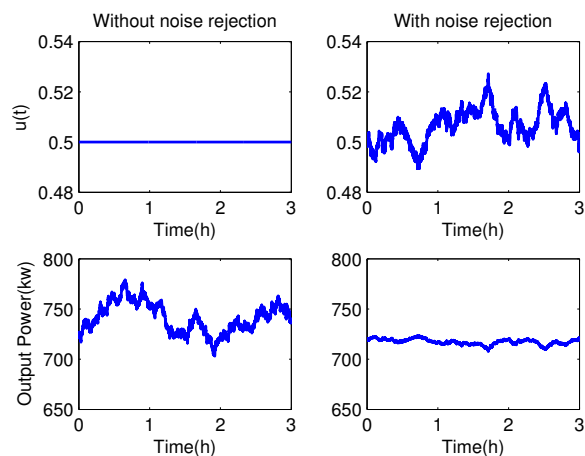


Fig. 5. Noise rejection via state feedback.

same performance. Setting $a \rightarrow 1$ would eliminate noise from the regulated power output, but make the control signal variance explode beyond its constraints. Intermediate values could potentially reduce the regulation variance while still keeping control within its admitted bound.

We now simulate the system with this fixed value of $u^* = 0.5$ against a system that continuously updates u to the deviations in δn following equation (13c). For this simulation we choose $a = 0.8$, which is a compromise between noise reduction and deviations in the control signal that may move the system far away from the nominal values.

In Fig. 5 we plot the results, showing the value of the control signal $u(t)$ (above) and the output power $p(t)$ (below). We can see that the state feedback is able to achieve an important reduction in the power variability, while the control signal $u(t)$ stays near the nominal value u^* .

As an additional remark, simulations with different scheduling policies show that, again in this case, the results are agnostic to the exact job scheduling policy.

IV. REGULATION TRACKING UNDER UNCERTAINTY

In the previous two sections we analyzed separately two aspects of the regulation control problem based on the deferral of aggregates of loads: tracking a reference signal and reducing noise. The natural conclusion is to now integrate these separate efforts into a control strategy that seeks reference tracking under noise.

We pursue this using again linearized models around the nominal operating point. The open loop model coincides with (10), but now we must consider in addition an external reference $\delta r(t)$ which the power output must track, so the performance specification will involve the *tracking error* $e(t) := \delta p(t) - \delta r(t)$. The integrated controller should have access to measurements of the state variable $\delta n(t)$, and also to the external reference $\delta r(t)$, producing an action $\delta u(t)$ on the plant (10) so as to minimize the error variance, while keeping a check on control effort. This could be framed as a joint \mathcal{H}_2 -optimal control design with cost

$$J' := E[e^2 + \beta(p^*)^2(\delta u)^2] \quad (17)$$

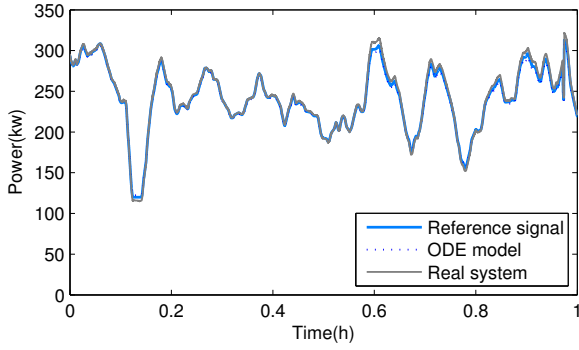


Fig. 6. Tracking a reference signal for the system with noise rejection.

which generalizes (11). A complete design of this kind would require a characterization of the class of reference signals to be tracked, for instance through a frequency weighting function. At this point we will opt for the simpler strategy of combining the feedback and feedforward components from the earlier sections, using a controller of the form

$$\delta u(t) = -K\delta n + \tilde{u}(t),$$

where K has the form (12) and \tilde{u} is a function of the reference input. Substitution into (10) gives

$$\delta \dot{n} = -\frac{1}{\tau}u^*(1-a)\delta n - \frac{1}{\tau}n^*\tilde{u} + v, \quad (18a)$$

$$\delta p = p_0[u^*(1-a)\delta n + n^*\tilde{u}], \quad (18b)$$

which leads in the Laplace transform domain to $\delta p(s) = G_{\tilde{u}p}^a(s)\tilde{u}(s) + G_{vp}^a(s)v(s)$, with

$$G_{\tilde{u}p}^a(s) = \frac{p_0n^*s}{s + \frac{u^*}{\tau}(1-a)}, \quad G_{vp}^a(s) = \frac{p_0u^*(1-a)}{s + \frac{u^*}{\tau}(1-a)}.$$

This suggests choosing the feedforward component \tilde{u} as

$$\tilde{u}(s) = \frac{1}{G_{\tilde{u}p}^a(s)}\delta r(s) = \frac{1}{p_0n^*} \left(1 + \frac{u^*(1-a)}{\tau s} \right) \delta r(s),$$

which results in the following closed loop transfer function from noise to tracking error:

$$e(s) = \delta p(s) - \delta r(s) = G_{vp}^a(s)v(s); \quad (19)$$

therefore the noise penalty on performance will be exactly the one computed in (15). The control effort will also have the noise term of (16), but in addition there is the impact of the reference signal analogous to (6).

To end we show a simulation of the complete system. We use the same signal and parameters of Fig. 2 adding the feedback for noise reduction. Again we choose $a = 0.8$. We can see in Fig. 6 that the tracking is clearly improved.

V. CONCLUSIONS

In this paper, we analyzed a model for a load demand aggregator that manages a large number of consumer deferrable loads and is capable of adjusting the number of active jobs to control its power. The proposed macroscopic model is oblivious to the exact management of the loads and captures the essential behavior of the system through the service

fraction the aggregator provides. Using this model, we were able to analyze the impact of variability in the demands, and design tracking and noise rejection controllers. These simple mechanisms enable a load aggregator to reduce its need for regulation services, and even offer regulation services to others. The results were evaluated through simulation, illustrating the performance of the designed mechanisms.

Several lines of future work remain open. A more thorough controller design to take into account the constraints in the input signal, as well as to cope with the nonlinearities in the system can be performed. From the queuing perspective, it would be interesting to analyze more precise fluid models for the different scheduling mechanisms involved.

REFERENCES

- [1] G. D. Bella, L. Giarrè, M. Ippolito, A. Jean-Marie, G. Neglia, and I. Tinnirello, "Modeling Energy Demand Aggregators for Residential Consumers," in *Proc. of the 52nd IEEE Conference on Decision and Control*, 2013.
- [2] EnergyPool, <http://www.energy-pool.eu/en/>.
- [3] A. Ferragut and F. Paganini, "Content dynamics in P2P networks from queueing and fluid perspectives," in *Proc. of 24th International Teletraffic Congress, Krakow, Poland*, 2012.
- [4] GoodEnergy, <http://www.goodenergy.com>.
- [5] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent, "A Generalized Battery Model of a Collection of Thermostatically Controlled Loads for Providing Ancillary Service," in *Proc. of the 51st Allerton conference on Communication, Control and Computing*, 2013.
- [6] —, "Frequency Regulation from Flexible Loads: Potential, Economics, and Implementation," in *Proc. of the American Control Conference (ACC)*, 2014.
- [7] H. Holttinen and et al., "Design and operation of power systems with large amounts of wind power," IEA wind, Tech. Rep., 2009.
- [8] J. Hong, X. Tan, and D. Towsley, "A performance analysis of minimum laxity and earliest deadline scheduling in a real-time system," *IEEE Transactions on Computers*, vol. 38, pp. 1736–1744, 1989.
- [9] B. Kirby, "Ancillary Services: Technical and Commercial Insights," Oak Ridge National Laboratory, Tech. Rep., 2007.
- [10] S. Koch, J. Mathieu, and D. Callaway, "Modeling and Control of Aggregated Heterogeneous Thermostatically Controlled Loads for Ancillary Services," in *Proc. of the 17th Power Systems Computation Conference*, 2011.
- [11] A. Nayyar, J. Taylor, A. Subramanian, K. Poolla, and P. Varaiya, "Aggregate Flexibility of a Collection of Loads," in *Proc. of the 52nd IEEE Conference on Decision and Control*, 2013.
- [12] New York Independent System Operator, "Ancillary services manual," http://www.nyiso.com/public/webdocs/markets_operations/documents/Manuals_and_Guides/Manuals/Operations/ancserv.pdf.
- [13] F. Paganini and A. Ferragut, "PDE models for population and residual work applied to peer-to-peer networks," in *Proc. of 46th Annual Conference on Information Sciences and Systems, Princeton, NJ*, 2012.
- [14] PJM, "Fast Response Regulation Signal," <http://www.pjm.com/markets-and-operations/ancillary-services/mkt-based-regulation/fast-response-regulation-signal.aspx>.
- [15] F. Rahimi and A. Ipakchi, "Demand Response as a Market Resource Under the Smart Grid Paradigm," *IEEE Transactions on Smart Grid*, vol. 1, pp. 82–88, 2010.
- [16] P. Robert, *Stochastic networks and queues*. Springer-Verlag, 2003.
- [17] J. A. Short, D. Infield, and L. L. Freris, "Stabilization of Grid Frequency Through Dynamic Demand Control," *IEEE Transactions on Power Systems*, vol. 22, pp. 1284–1293, 2007.
- [18] A. Subramanian, M. Garcia, D. Callaway, K. Poolla, and P. Varaiya, "Real-Time Scheduling of Distributed Resources," *IEEE Transactions on Smart Grid*, vol. 4, pp. 2122–2130, 2013.
- [19] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice Hall, 1996.