Real-Time Energy Management in Microgrids With Reduced Battery Capacity Requirements

Bingcong Li, Tianyi Chen, *Student Member, IEEE*, Xin Wang, *Senior Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Energy storage units hold promise to transform the electric power industry, since they can supply power to end customers during peak demand times, and operate as customers upon a power surplus. This paper studies online energy management with renewable energy resources and energy storage units. For the problem at hand, the popular approaches rely on stochastic dual (sub)gradient (SDG) iterations for a chosen stepsize μ, which generally require battery capacity $\mathcal{O}(1/\mu)$ to guarantee an $\mathcal{O}(\mu)$-optimal solution. With the goal of achieving optimal energy cost with considerably reduced battery capacity requirements, an online learning-aided management (OLAM) scheme is introduced for energy management, which incorporates statistical learning advances into real-time energy management. To facilitate real-time implementation of the proposed scheme, the alternating direction method of multipliers method is also leveraged to solve the involved subproblems in a distributed fashion. It is analytically established that OLAM incurs an $\mathcal{O}(\mu)$ optimality gap, while only requiring battery capacity $\mathcal{O}(\log^2(\mu)/\sqrt{\mu})$. Simulations on the IEEE power grid benchmark corroborate that OLAM incurs similar average cost relative to that of SDG, while requiring markedly lower battery capacity.**

***Index Terms*—Energy management, smart microgrids, energy storages, stochastic approximation, statistical learning.**

NOMENCLATURE

Indices, Numbers, and Sets

T, t	Number and index of time slots.
N, i, $\mathcal{N}$	Number, index, and set of buses.
E, e, $\mathcal{E}$	Number, index, and set of transmission lines.
k	Iteration indexes of the distributed algorithm.

Constants

$\boldsymbol{\alpha}_t$	Buying/selling energy prices at slot t.
$\mathbf{d}_t$	Power demand of all buses at slot t.
$\mathbf{r}_t$	Renewable generator generated power at slot t.
$\mathbf{s}_t$	Microgrid system state at slot t.
$\underline{\mathbf{g}}$, $\overline{\mathbf{g}}$	Lower/upper limits of power generation on conventional generators.
$\mathbf{q}_0$, $\underline{\mathbf{q}}$, $\overline{\mathbf{q}}$	Initial, minimum and maximum battery levels.
$\overline{\mathbf{f}}$	Capacity of transmission lines.
κ	Charging/discharging efficiency.
$\underline{\mathbf{b}}$, $\overline{\mathbf{b}}$	Maximum charging and discharging rates.
τ	Duration of one slot.
$\mathbf{A}$	Incidence matrix of microgrid.
$\boldsymbol{\alpha}_t$	Power price of all buses at slot t.
μ	Stepsize in OLAM.
$\boldsymbol{\beta}$	Bias control variable in OLAM.

Variables

$\mathbf{p}_t$	Power purchased from/sold to main grid at slot t.
$\mathbf{g}_t$	Conventional generator generated power at slot t.
$\mathbf{f}_t$	Power flow on transmission line at slot t.
$\mathbf{b}_t^+$	Battery charged power on all buses at slot t.
$\mathbf{b}_t^-$	Battery discharged power at slot t.
$\mathbf{b}_t$	Auxiliary charging/discharging variable at slot t.
$\boldsymbol{\theta}_t$	Voltage phases at slot t.
$\mathbf{q}_t$	Battery level at the beginning of slot t.
$\check{\boldsymbol{\lambda}}_t$	Effective dual variable at iteration t.
$\hat{\boldsymbol{\lambda}}_t$	Empirical dual variable at iteration t.

Manuscript received August 13, 2017; revised November 9, 2017; accepted December 9, 2017. Date of publication December 15, 2017; date of current version February 18, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61671154, in part by the National Key Research and Development Program of China under Grant 2017YFB0403402, and in part by the U.S. NSF under Grant 1509005, Grant 1508993, and Grant 1711471. Paper no. TSG-01170-2017.

B. Li was with the State Key Laboratory of ASIC and System, Shanghai Institute for Advanced Communication and Data Science, Department of Communication Science and Engineering, Fudan University, Shanghai 200433, China. He is now with the Department of Electrical and Computer Engineering and the Digital Technology Center, University of Minnesota, Minneapolis, MN 55455 USA. (e-mail: lixx5599@umn.edu).

T. Chen, and G. B. Giannakis are with the Department of Electrical and Computer Engineering and the Digital Technology Center, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: chen3827@umn.edu; georgios@umn.edu).

X. Wang is with the State Key Laboratory of ASIC and System, Shanghai Institute for Advanced Communication and Data Science, Department of Communication Science and Engineering, Fudan University, Shanghai 200433, China (e-mail: xwang11@fudan.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2017.2783894

I. INTRODUCTION

THE CURRENT grid infrastructure is on the verge of a major paradigm shift, migrating from the aging grid to a "smart" one. Among several attractive features, the smart grid will embrace renewable energy sources (RES) that hold great potential to mitigate the increasing threats of global warming and climate change concerns. Equipped with RES and other advanced capabilities, microgrids are also envisioned to be deployed across a limited geographic area (e.g., a campus) [1].

In a microgrid, distributed storage effected with batteries, promises to cope with the intermittent nature of RES (wind energy and solar energy), and the stochasticity of end-user demand. With the rampant proliferation of RES and

TABLE I
A SUMMARY OF RELATED WORKS ON REAL-TIME ENERGY MANAGEMENT

Reference	Storage units	Network integration	Learning-aided
[2], [4]–[6], [8]–[10]	Yes	No	No
[12], [13]	No	Yes	No
[14], [15]	Yes	Yes	No
This work	Yes	Yes	Yes

energy storage units, current cyber-control approaches are challenged by pronounced variability and uncertainty. Prior works on those based on energy management under uncertainty include robust and stochastic optimization [2], [3]. Targeting optimal performance on average, dynamic programming based algorithms for energy management with storage units and stochastic energy pricing have been pursued in [4] and [5]. However, dynamic programming-based approaches usually incur prohibitively high complexity, especially as the number of involved states grows. Thanks to its low complexity and provable optimality, stochastic gradient-based methods such as Lyapunov optimization (LO) have been recently adopted to enable real-time energy management. Specifically, energy management is considered in [6]–[10] to reduce electricity bills of data centers and operational cost of cellular networks by leveraging energy storage. For a more comprehensive microgrid architecture, [11] developed a LO-based algorithm for real-time power balancing with ramping constraints, and elastic as well as inelastic loads. It is worth mentioning that only an aggregate supply-demand model was considered in [6]–[11].

However, the aggregate-based supply-demand model does not account for the underlying power networks, the associated power flow (Kirchhoffs laws), and grid operational constraints. Hence, control decisions generated by [6]–[11] do not guarantee satisfaction of the real-world constraints. Recently, increasing attention has been paid to network-constrained energy management, which results in a more sophisticated control policy as dispatch decisions per bus become coupled. Without storage units, [12], [13] dealt with energy management under network constraints. Considering both storage and network constraints, an LO-based energy management algorithm was reported in [14], while [15] further incorporated a more general battery model with possible energy leakage over time. Yet, the salient tradeoff between optimality gap and battery capacity inherent to LO-based approaches lend themselves to microgrids equipped with batteries that must have sufficiently large capacity to guarantee near-optimality, which leads to high installment overhead [16].

With the goal of achieving optimal energy cost at considerably reduced battery capacity requirements, the present paper targets a novel online microgrid energy management algorithm, which we term online learning aided method (OLAM).

The main contributions can be summarized as follows; see also a summary in Table I.

c1) Combining the power of statistical learning and stochastic optimization, OLAM offers online energy management in smart microgrids.

c2) With μ denoting stepsize, it is established that OLAM yields a cost-battery capacity tradeoff $[\mu, (\log^2(\mu))/\sqrt{\mu}]$ for energy management, which markedly improves the $[\mu, 1/\mu]$ tradeoff of SDG [6], [7], [11], [15], [17], [18].

c3) To facilitate real-time implementation of the proposed scheme, the alternating direction method of multipliers (ADMM) method is adopted to develop a distributed and scalable variant of OLAM.

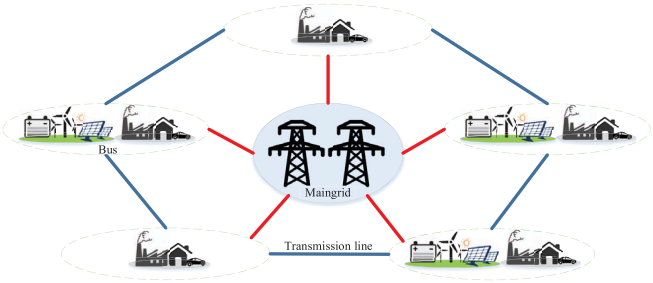


Fig. 1. Diagram of a microgrid connecting the main grid.

Notation: $\mathbb{E}(\mathbb{P})$ denotes the expectation (probability); $(\cdot)^\top$ stands for vector and matrix transposition; and $\|\mathbf{x}\|$ denotes the ℓ_2-norm of a vector $\mathbf{x}$, while the inequalities for vectors, e.g., $\mathbf{x} > \mathbf{0}$, are defined entry-wise. The projection is defined as $[\mathbf{x}]_{\underline{\mathbf{x}}}^{\overline{\mathbf{x}}} := \max\{\underline{\mathbf{x}}, \min\{\overline{\mathbf{x}}, \mathbf{x}\}\}$, also entry-wise.

II. SYSTEM MODELING

In this section, we begin with a microgrid model consisting of renewable energy sources and storage units, and then introduce its real-time energy management task we will deal with.

A. Modeling Preliminaries

Consider a microgrid represented by the graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, e.g., the gird in Fig. 1, where the set of nodes $\mathcal{N}$ comprises to N buses, and the edges in $\mathcal{E}$ correspond to E transmission lines. The $N \times E$ bus-edge incidence matrix capturing the microgrid connectivity is defined entry-wise as

$$\mathbf{A}_{(n,e)} := \begin{cases} 1, & \text{if edge } e \text{ enters bus } i \\ -1, & \text{if edge } e \text{ leaves bus } i \\ 0, & \text{else.} \end{cases} \tag{1}$$

We collect voltage phases in the $N \times 1$ vector $\boldsymbol{\theta}_t$; and stack power flows on all transmission lines in $\mathbf{f}_t \in \mathbb{R}^E$. Under the approximate DC flow model, the aforementioned grid quantities satisfy $\mathbf{f}_t = \mathbf{D}\mathbf{A}^\top\boldsymbol{\theta}_t$, where $\mathbf{D} \in \mathbb{R}^{E\times E}$ is a diagonal matrix, see [19] for detailed derivation. With vector $\bar{\mathbf{f}} \in \mathbb{R}^E$ denoting the transmission-line flow capacities, the network-imposed constraints are $-\bar{\mathbf{f}} \leq \mathbf{f}_t \leq \bar{\mathbf{f}}$. For convenience, $\bar{\mathbf{f}}$ is assumed constant here. A more complex AC power flow model can be also considered, but requires elaborate convexification techniques; see [20] and the references therein.

Generators and batteries: We consider that each bus is connected with a load entity, and some buses are also equipped with a battery, a conventional generator, and a renewable generator. Per slot t, concatenate (random) outputs from renewable generator and energy demand per bus into the $N \times 1$ vectors $\mathbf{r}_t$ and $\mathbf{d}_t$. Let also $\mathbf{g}_t \in \mathbb{R}^N$ collect the generated powers from

all conventional generators. The outputs of conventional generators per time t are confined in the range $\underline{\mathbf{g}} \leq \mathbf{g}_t \leq \overline{\mathbf{g}}$, where the bounds $[\underline{\mathbf{g}}, \overline{\mathbf{g}}]$ depend on the generation capacity per bus. Note that for brevity, generator limits $\underline{\mathbf{g}}$ but $\overline{\mathbf{g}}$ are assumed to be time-invariant here, but generalizing to scenarios with dynamic network reconfigurations is straightforward.

Storage units are installed in modern microgrids to provide extra power supply during the peak demand hours; e.g., charging the battery upon a power surplus, and discharging upon a power deficit. Vectors $\mathbf{q}_t \in \mathbb{R}^N$ collect the state of charge at the beginning of period t. Let $\mathbf{b}_t^+ \in \mathbb{R}^N$ and $\mathbf{b}_t^- \in \mathbb{R}^N$ be charging and discharging power vectors, respectively. Since the battery can be only charging or discharging at one slot, $(b_t^i)^+$ and $(b_t^i)^-$ cannot be positive at the same time. With $\kappa \in (0, 1]$ denoting the charging/discharging efficiency and τ be the duration of each slot t, the energy-level of each battery satisfies the recursion

$$\mathbf{q}_{t+1} = \mathbf{q}_t + \tau\left(\kappa \mathbf{b}_t^+ - \mathbf{b}_t^-/\kappa\right) \tag{2a}$$

$$\underline{\mathbf{q}} \leq \mathbf{q}_t \leq \overline{\mathbf{q}} \tag{2b}$$

$$\mathbf{0} \leq \mathbf{b}_t^+ \leq \overline{\mathbf{b}} \tag{2c}$$

$$\mathbf{0} \leq \mathbf{b}_t^- \leq \underline{\mathbf{b}} \tag{2d}$$

where the constraints (2a) capture the battery dynamics over consecutive time slots as well as the state of charge within the allowable battery levels $[\underline{\mathbf{q}}, \overline{\mathbf{q}}]$, while (2c) and (2d) ensures the amount of charging and discharging stays within the limit.

Microgrid cost: Let vector $\mathbf{p}_t \in \mathbb{R}^N$ collect the buying/selling power from the main grid at slot t, and $\boldsymbol{\alpha}_t \in \mathbb{R}^N$ concatenate the per-unit power price at each bus. With output $\mathbf{g}_t$, the cost of all conventional generators is a convex function $\Phi^c(\mathbf{g}_t)$. To avoid battery degradation and extend the battery lifetime, a convex function $\Phi^b(\mathbf{b}_t^+, \mathbf{b}_t^-)$ is adopted to penalize fast charging or discharging. With vector $\mathbf{x}_t$ collecting all the optimization variables $\{\mathbf{p}_t, \mathbf{g}_t, \mathbf{b}_t^+, \mathbf{b}_t^-, \mathbf{f}_t, \boldsymbol{\theta}_t\}$ except $\mathbf{q}_t$, and vector $\boldsymbol{s}_t$ collecting all random variables $\{\mathbf{r}_t, \mathbf{d}_t, \boldsymbol{\alpha}_t\}$, performance is characterized by the grid-wide cost, given by

$$\Phi_t(\mathbf{x}_t) = \Phi(\mathbf{x}_t; \boldsymbol{s}_t) := \boldsymbol{\alpha}_t^\top \mathbf{p}_t + \Phi^c(\mathbf{g}_t) + \Phi^b\left(\mathbf{b}_t^+, \mathbf{b}_t^-\right) \tag{3}$$

which aggregates the generator and storage costs, along with the cost of external energy sources.

B. Problem Statement

At a high level, the goal of a microgrid operator is to dispatch power in the most economical manner by leveraging storage units, while obeying the transmission line, generation, and storage operational constraints. Concretely, the energy management task amounts to solving the following dynamic stochastic optimization problem in an online fashion

$$\Phi^* := \min_{\{\mathbf{x}_t, \mathbf{q}_t\}_t} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[\Phi_t(\mathbf{x}_t)] \tag{4a}$$

$$\text{s.t.} \quad \mathbf{g}_t + \mathbf{p}_t + \mathbf{r}_t + \mathbf{A}\mathbf{f}_t = \mathbf{d}_t + \kappa \mathbf{b}_t^+ - \frac{\mathbf{b}_t^-}{\kappa}, \ \forall t \tag{4b}$$

$$\mathbf{f}_t = \mathbf{D}\mathbf{A}^\top \boldsymbol{\theta}_t, \ \forall t \tag{4c}$$

$$-\overline{\mathbf{f}} \leq \mathbf{f}_t \leq \overline{\mathbf{f}}, \ \forall t \tag{4d}$$

$$\underline{\mathbf{g}} \leq \mathbf{g}_t \leq \overline{\mathbf{g}}, \ \forall t \tag{4e}$$

$$\mathbf{q}_{t+1} = \mathbf{q}_t + \tau(\kappa \mathbf{b}_t^+ - \mathbf{b}_t^-/\kappa), \quad \forall t \tag{4f}$$

$$\underline{\mathbf{q}} \leq \mathbf{q}_t \leq \overline{\mathbf{q}}, \ \forall t \tag{4g}$$

$$\mathbf{0} \leq \mathbf{b}_t^+ \leq \overline{\mathbf{b}}, \ \forall t \tag{4h}$$

$$\mathbf{0} \leq \mathbf{b}_t^- \leq \underline{\mathbf{b}}, \ \forall t \tag{4i}$$

where the expectation $\mathbb{E}$ in (4a) is taken over the random vector $\boldsymbol{s}_t$ and possible randomness of the control policy $\boldsymbol{\chi}(\boldsymbol{s}_t) := \mathbf{x}_t$; constraints in (4b) correspond to the power supply-demand balance; constraints in (4c) specify the real power flow at each transmission line; constraints in (4d) and (4e) bound power flows and outputs from conventional generators within their feasible regions; constraints in (4f) capture the recursion of energy levels in the batteries; and constraints in (4g)-(4i) enforce battery level and (dis)charging rates within their limits.

Certainly, the long-term cost minimization in (4) cannot be achieved by myopically minimizing $\Phi_t(\mathbf{x}_t)$ per slot. To see this, notice that the (dis)charging decisions $\mathbf{b}_t$ are coupled across time through the battery dynamics (4f). Specifically, given $\mathbf{q}_t$ at time t, the amount of (dis)charge $\mathbf{b}_t^+$, $\mathbf{b}_t^-$ should satisfy $\mathbf{0} \leq \mathbf{b}_t^+ \leq \min\{\overline{\mathbf{b}}, \frac{\overline{\mathbf{q}}-\mathbf{q}_t}{\tau\kappa}\}$, and $\mathbf{0} \leq \mathbf{b}_t^- \leq \min\{\underline{\mathbf{b}}, \frac{\kappa(\mathbf{q}_t-\underline{\mathbf{q}})}{\tau}\}$, where the upper- and lower-bounds depend on the previous decision $\mathbf{q}_t$ thus $\mathbf{b}_{t-1}^+$ and $\mathbf{b}_{t-1}^-$. Clearly, making decisions at slot t can have impact on future decisions over time, which highlights the difficulty of solving (4). Even worse, for the practical case where the knowledge of $\boldsymbol{s}_t$ is causal, finding the optimal solution while accounting for the coupling across time, generally calls for a dynamic programming solver [21], which is known to suffer from the "curse of dimensionality", and it is thus intractable in our online setting. For this reason, we will circumvent this obstacle by relaxing (4f)-(4g) to average constraints, and employing dual decomposition techniques next.

III. Lagrange Duality and SDG

Targeting a low-complexity solver without reliance on dynamic programming, this section first reformulates the dynamic optimization (4) to a stationary one, based on which an online energy management iteration, such as SDG, could be applied.

A. Problem Reformulation

To solve (4), the first step is to combine constraints (4f) and (4g), and argue that in the long-term, battery (dis)charging amounts must satisfy the following necessary condition

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[\kappa \mathbf{b}_t^+ - \frac{\mathbf{b}_t^-}{\kappa}\right] = \mathbf{0}. \tag{5}$$

Indeed, summing (4f) over slots $t = 1, \ldots, T$ and taking expectation yields $\mathbb{E}[\mathbf{q}_{T+1}] = \mathbb{E}[\mathbf{q}_1] + \tau \sum_{t=1}^{T} \mathbb{E}[\kappa \mathbf{b}_t^+ - \mathbf{b}_t^-/\kappa]$. Since both $\mathbf{q}_1$ and $\mathbf{q}_{T+1}$ are bounded due to (4g), dividing both sides by T and taking limits as $T \to \infty$, implies (5). In words, on average all charged power should be discharged during real-time power balancing. To simplify notation, we assume without loss of generality that $\tau = 1$, and introduce

an auxiliary variable $\mathbf{b}_t := \kappa \mathbf{b}_t^+ - \mathbf{b}_t^-/\kappa$ to denote charge and discharge at the same time, which allows rewriting (5) as

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[\mathbf{b}_t] = \mathbf{0}. \tag{6}$$

To assure that $\mathbf{b}_t$ meets constraints (4h) and (4i), it should further satisfy

$$-\underline{\mathbf{b}}/\kappa \leq \mathbf{b}_t \leq \kappa \overline{\mathbf{b}}. \tag{7}$$

Given b_t^i, $(b_t^i)^+$ and $(b_t^i)^-$ can be uniquely recovered as $(b_t^i)^+ = \max\{0, b_t^i/\kappa\}$ and $(b_t^i)^- = -\min\{0, \kappa b_t^i\}$. Note that (5) is a relaxation of (4f) and (4g), so a feasible solution of (6) may not be feasible for (4f) and (4g). Using (6) and (7), a relaxed version of (4) is

$$\begin{aligned} \tilde{\Phi}^* := \min_{\{\mathbf{x}_t\}} \ & \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[\Phi_t(\mathbf{x}_t)] \\ \text{s.t.} \ & \text{(4b)} - \text{(4e)}, \ \text{(6)}, \ \text{and (7)}. \end{aligned} \tag{8}$$

Compared to (4), problem (8) eliminates the time-coupling across variables $\mathbf{q}_t$ by replacing (4f) and (4g) with (6). Since (8) is a relaxed version of (4), we have $\tilde{\Phi}^* \leq \Phi^*$. Thus, if one solves (8) instead of (4), it will be useful to assess the optimality gap $\tilde{\Phi}^* - \Phi^*$, and also ensure that the schedule obtained by solving (8) is feasible for (4). In addition, using arguments similar to those in [17], it can be shown that if the random process $\{\boldsymbol{s}_t\}$ is stationary, there exists a *stationary* control policy $\boldsymbol{\chi}^*(\cdot)$, which maps the current $\boldsymbol{s}_t$ to the instantaneous decision $\boldsymbol{\chi}^*(\boldsymbol{s}_t) := \mathbf{x}_t^*$ including $\mathbf{b}^*(\boldsymbol{s}_t) := \mathbf{b}_t^*$; vector $\mathbf{x}_t$ satisfies (4b)-(4e) and (7), as well as guarantees that $\mathbb{E}[\Phi(\boldsymbol{\chi}^*(\boldsymbol{s}_t); \boldsymbol{s}_t)] = \tilde{\Phi}^*$ and $\mathbb{E}[\mathbf{b}^*(\boldsymbol{s}_t)] = \mathbf{0}$. Stationarity implies that all the expectations inside the limiting time averages in (8) yield the same quantity; hence, the *dynamic* problem (8) is equivalent to the *stationary* stochastic program

$$\tilde{\Phi}^* := \min_{\boldsymbol{\chi}(\cdot)} \ \mathbb{E}[\Phi(\boldsymbol{\chi}(\boldsymbol{s}_t); \boldsymbol{s}_t)] \tag{9a}$$

$$\text{s.t.} \quad \mathbb{E}[\mathbf{b}(\boldsymbol{s}_t)] = \mathbf{0} \tag{9b}$$

$$\text{(4b)} - \text{(4e)}, \ \text{(4h) and (4i)}, \ \forall \boldsymbol{s}_t \tag{9c}$$

where $\mathbb{E}$ is again taken over the random vector $\boldsymbol{s}_t$; possible randomness of the policy $\boldsymbol{\chi}(\boldsymbol{s}_t)$; and, $\Phi(\boldsymbol{\chi}(\boldsymbol{s}_t); \boldsymbol{s}_t) := \Phi_t(\mathbf{x}_t)$.

Note that the optimization in (9) is with respect to the policy $\boldsymbol{\chi}(\cdot)$, which is an infinite dimensional problem in the primal domain. Observe though, that there are only a finite number of constraints coupling the realizations [see (9b)]. Thus, the dual problem contains a finite number of variables, which motivates solving (9) is more tractable in the dual domain.

B. Lagrange Dual and Optimal Policy

With $\boldsymbol{\lambda} \in \mathbb{R}^N$ denoting the Lagrange multiplier vector associated with (9b), the partial Lagrangian of (8) is

$$\mathcal{L}(\boldsymbol{\chi}, \boldsymbol{\lambda}) := \mathbb{E}[\mathcal{L}_t(\mathbf{x}_t, \boldsymbol{\lambda})] \tag{10}$$

where the instantaneous partial Lagrangian is given by

$$\mathcal{L}_t(\mathbf{x}_t, \boldsymbol{\lambda}) := \ \Phi_t(\mathbf{x}_t) + \boldsymbol{\lambda}^\top \mathbf{b}_t \tag{11}$$

with constraint (9c) remaining implicit.

Correspondingly, the Lagrange dual function is

$$\mathcal{D}(\boldsymbol{\lambda}) := \min_{\{\boldsymbol{\chi}(\boldsymbol{s}_t) \in \mathcal{X}(\boldsymbol{s}_t), \forall \boldsymbol{s}_t\}} \mathcal{L}(\boldsymbol{\chi}, \boldsymbol{\lambda}) := \mathbb{E}\left[\min_{\mathbf{x}_t \in \mathcal{X}_t} \mathcal{L}_t(\mathbf{x}_t, \boldsymbol{\lambda})\right] \tag{12}$$

where set $\mathcal{X}(\boldsymbol{s}_t) := \mathcal{X}_t$ collects constraints in (4b)-(4e), and (4h). Hence, the dual problem of (9) is

$$\max_{\boldsymbol{\lambda} \in \mathbb{R}^N} \mathcal{D}(\boldsymbol{\lambda}) := \mathbb{E}[\mathcal{D}_t(\boldsymbol{\lambda})] \tag{13}$$

where $\mathcal{D}_t(\boldsymbol{\lambda}) := \min_{\mathbf{x}_t \in \mathcal{X}_t} \mathcal{L}_t(\mathbf{x}_t, \boldsymbol{\lambda})$.

If the optimal Lagrange multiplier $\boldsymbol{\lambda}^*$ were known, then optimizing (8) or (9) would be equivalent to minimizing the Lagrangian $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ or all instantaneous $\{\mathcal{L}_t(\mathbf{x}_t, \boldsymbol{\lambda}^*)\}$, over the set $\mathcal{X}_t$ [22, Prop. 3.3.4]. To obtain the aforementioned optimal energy management, $\boldsymbol{\lambda}^*$ must be known. However, the distribution of $\boldsymbol{s}_t$ is usually unknown in practice, which motivates our novel approach developed next.

C. LO-Based Energy Management Schemes

To better illustrate the merits of our approach, we first outline the popular Lyapunov optimization (LO)-based solver, which has been successfully applied to various network control tasks with energy storage units [6], [7], [11], [15], [18]. The major challenge in solving (13) is that the distribution of $\boldsymbol{s}_t$ is usually unknown in practice. Existing approaches rely on LO-based a.k.a. SDG-based techniques to find $\boldsymbol{\lambda}^*$; that is, the subproblem solved per slot is [see (11)]

$$\min_{\mathbf{x}_t \in \mathcal{X}_t} \Phi_t(\mathbf{x}_t) + \boldsymbol{\lambda}_t^\top \mathbf{b}_t, \quad \forall t \tag{14a}$$

$$\boldsymbol{\lambda}_{t+1} = \boldsymbol{\lambda}_t + \mu \mathbf{b}_t, \quad \forall t \tag{14b}$$

for a predefined positive constant μ [6], [7], [11], [15], [18]. In LO-based approaches, the primal recursion (14a) generates real time energy management decisions, while the Lagrange multiplier recursion (14b) is within a scalar identical to (4f); hence, it is often viewed as a scaled virtual queue recursion of $\mathbf{q}_t$. For energy management tasks, the LO-based method entails a optimality-battery capacity tradeoff [7], [11], [17] that we summarize in the ensuing proposition.

Proposition 1: If Φ^* is the optimal cost in (4), the LO-based solver of (14) achieves an $\mathcal{O}(\mu)$ optimal solution satisfying

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[\Phi_t(\mathbf{x}_t(\boldsymbol{\lambda}_t))] \leq \Phi^* + \mathcal{O}(\mu) \tag{15}$$

where $\mathbf{x}_t(\mathbf{q}_t)$ denotes the decisions obtained from (14); and the LO-based solver incurs a steady-state battery level

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\big[\mathbf{q}_t\big] = \mathcal{O}(1/\mu). \tag{16}$$

This $[\mathcal{O}(\mu), \mathcal{O}(1/\mu)]$ performance tradeoff emerges in most LO-based energy management approaches [6]–[9], [11], [15], [18], [23]. Using the virtual queue interpretation of Lagrange multiplier, e.g., $\boldsymbol{\lambda}_t = \mu \mathbf{q}_t$, sometimes it is also referred to $[\mathcal{O}(1/V), \mathcal{O}(V)]$ tradeoff with $\mu = 1/V$. However, batteries with large capacity are costly [16], which renders LO approaches expensive to achieve close-to-optimal performance in practice. This motivates our novel learning-aided stochastic gradient method, termed OLAM, which will be described next.

Algorithm 1 OLAM for Energy Management in Microgrids

Initialize: dual iterates $\check{\boldsymbol{\lambda}}_1$ and $\hat{\boldsymbol{\lambda}}_1$, battery level $\mathbf{q}_1$, bias control variable $\boldsymbol{\beta}$, and proper stepsizes μ and $\{\eta_t, \forall t\}$.
for $t = 1, 2 \ldots$ **do**
  Energy management (1st gradient):
  Construct the hybrid dual variable via (17).
  Observe state $\boldsymbol{s}_t$, and obtain $\mathbf{x}_t(\check{\boldsymbol{\lambda}}_t)$ via (18).
  Update the battery level via (19).
  Sample recourse (2nd gradient):
  Obtain variable $\mathbf{x}_t(\hat{\boldsymbol{\lambda}}_t)$ by solving (21) with sample $\boldsymbol{s}_t$.
  Update the empirical dual variable $\hat{\boldsymbol{\lambda}}_{t+1}$ via (20).
end for

IV. Learning-Aided Energy Management

In this section, we develop our novel OLAM scheme, along with its decentralized implementation, namely d-OLAM.

A. OLAM Algorithm

The performance tradeoff of LO- and SDG-based approaches [11], [15], [18] is mainly due to the direct relation $\boldsymbol{\lambda}_t/\mu = \mathbf{q}_t$. The consequence is coming μ-close to the optimal cost Φ^* with stepsize μ, the battery capacity $\overline{\mathbf{q}}$ must be inversely proportional to μ [see Prop. 1]. Aiming to improve this tradeoff, the key idea behind our approach is to disentangle the direct tie between stochastic multiplier and battery level, while still ensuring optimality and feasibility. Specifically, instead of the single $\boldsymbol{\lambda}_t$ in SDG, our novel energy management is driven by the combined stochastic multiplier

$$\check{\boldsymbol{\lambda}}_t = \hat{\boldsymbol{\lambda}}_t + \mu \mathbf{q}_t - \boldsymbol{\beta} \tag{17}$$

which incorporates an empirical term $\hat{\boldsymbol{\lambda}}_t$ obtained through a learning process, and a scaled instantaneous battery level $\mu\mathbf{q}_t$, along with a bias-control vector variable $\boldsymbol{\beta}$ which will enable OLAM to attain near optimally in the steady state. The disentangled updates of $\hat{\boldsymbol{\lambda}}_t$ and $\mathbf{q}_t$ follow next. At each time slot t, OLAM finds two stochastic gradients using $\boldsymbol{s}_t$. For the first gradient, it minimizes the instantaneous Lagrangian

$$\mathbf{x}_t\left(\check{\boldsymbol{\lambda}}_t\right) = \arg\min_{\mathbf{x}\in\mathcal{X}_t} \mathcal{L}_t\left(\mathbf{x}, \check{\boldsymbol{\lambda}}_t\right) \tag{18}$$

which depends on $\check{\boldsymbol{\lambda}}_t$ [see (17)]. Next, we update the battery level using $\mathbf{b}_t(\check{\boldsymbol{\lambda}}_t) \in \mathbf{x}_t(\check{\boldsymbol{\lambda}}_t)$; that is,

$$\mathbf{q}_{t+1} = \left[\mathbf{q}_t + \mathbf{b}_t\left(\check{\boldsymbol{\lambda}}_t\right)\right]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} \tag{19}$$

which ensures the battery level feasibility in (4g).

Regarding the second gradient evaluation, OLAM simply adopts a stochastic gradient iteration for (13), namely

$$\hat{\boldsymbol{\lambda}}_{t+1} = \hat{\boldsymbol{\lambda}}_t + \eta_t \mathbf{b}_t\left(\hat{\boldsymbol{\lambda}}_t\right) \tag{20}$$

where $\mathbf{b}_t(\hat{\boldsymbol{\lambda}}_t)$ is obtained as

$$\mathbf{b}_t\left(\hat{\boldsymbol{\lambda}}_t\right) = \arg\min_{\mathbf{b}\in\mathcal{X}_t} \mathcal{L}_t\left(\mathbf{b}, \hat{\boldsymbol{\lambda}}_t\right) \tag{21}$$

with η_t denoting a diminishing stepsize. The intuition behind the combined multiplier $\check{\boldsymbol{\lambda}}_t$ in (17) is that instead of binding statistical learning and battery control through the single multiplier $\boldsymbol{\lambda}_t$, improved performance tradeoff can be achieved by breaking this dual role of the stochastic multiplier. Ideally, if both $\check{\boldsymbol{\lambda}}_t$ and $\hat{\boldsymbol{\lambda}}_t$ can be shown convergent to $\boldsymbol{\lambda}^*$, the steady-state battery level will be $\mathbf{q}_\infty \approx \boldsymbol{\beta}/\mu$, and thus the order of $\mathbf{q}_\infty$ can be reduced by a μ-dependent selection on $\boldsymbol{\beta}$. This is in contrast to a single $\boldsymbol{\lambda}_t$, which leads to $\boldsymbol{q}_\infty \approx \mathcal{O}(1/\mu)$ in Prop. 1. The proposed OLAM is summarized in Algorithm 1.

To recap, OLAM incrementally learns the microgrid state statistics from streaming data $\boldsymbol{s}_t$, and monitors the battery level $\mathbf{q}_t$ simultaneously. Different from [7], [11], [15], and [18], the real-time energy management decisions (18) are obtained by accounting for both learning and allocation effects through the combined multiplier $\check{\boldsymbol{\lambda}}_t$. Similar to the stepsize in LO-based approaches [6], [8], [9], [11], [15], [18], μ in (17) also controls the performance tradeoff. Intuitively, a large μ will allow the policy $\mathcal{X}$ to quickly respond to instantaneous variations so that the policy gains improved control of the battery level, while a small μ puts more weight on statistical learning from historical samples. A formal statement will be established in Section V.

Remark 1: Unlike [24] and [25], which dealt with a generic network resource allocation task that targets only queueing delay reduction, the main challenge for the microgrid setting here is to ensure the battery level also stays within the target physical range $[\underline{\mathbf{q}}, \overline{\mathbf{q}}]$. While the original learn-and-adapt approaches were developed without such bounds [24], OLAM with suitably chosen parameters guarantees an $\mathcal{O}(\mu)$ probability violation, as well as an $\mathcal{O}(\mu)$ extra cost to compensate for this violation.

Remark 2: In order to ensure feasibility of the battery, a projection is taken in (19), and thus the actual (dis)charging amount is adjusted accordingly. This means that the actual (dis)charging amount $\mathbf{b}_t(\check{\boldsymbol{\lambda}}_t) \coloneqq \mathbf{q}_{t+1} - \mathbf{q}_t$ is no longer a subgradient of $\mathcal{D}_t(\boldsymbol{\lambda})$ at $\boldsymbol{\lambda} = \check{\boldsymbol{\lambda}}_t$, which seems discouraging. Due to this mismatch between the planned and the actual (dis)charging amounts, extra cost is introduced for online power balancing through buying from and selling to the external market. Fortunately, our analysis demonstrates that the probability of having this mismatch is $\mathcal{O}(\mu)$, which does not affect the sub-optimality presented in Theorem 2; see also Lemma 2.

B. Distributed Implementation

Compared with SDG, OLAM needs an additional gradient update. Per slot t, the worst-case complexity for solving (18) and (21) is $\mathcal{O}((3N + E)^{3.5})$ when using an interior-point method. Therefore, the complexity of OLAM is two times that of SDG. The computational cost of OLAM grows as N and E increase, and a centralized grid operator with global information is required to perform the designed real-time management. To overcome this limitation, we endow OLAM with a proper ADMM iteration, and develop a decentralized and scalable version of OLAM next.

Since solving (18) and (21) only differ in the dual variable ($\check{\boldsymbol{\lambda}}_t$ in (18) and $\hat{\boldsymbol{\lambda}}_t$ in (21)), we focus on (18) in the following

discussion. For notational brevity, we omit the iteration index t throughout this section. With $E(i)$ denoting all the edges linked with bus i, and $V(e)$ for buses linked by edge e, we use $\mathbf{f}^i \in \mathbb{R}^{|E(i)|}$ to collect all entries in $\mathbf{f}$ that connect with bus i, and $\boldsymbol{\theta}^e \in \mathbb{R}^2$ to represent phases of buses linked by edge e. Further, we introduce auxiliary variables $\hat{\mathbf{f}}^i \in \mathbb{R}^{|E(i)|}$ and $\hat{\boldsymbol{\theta}}^e \in \mathbb{R}^2$ [26], and define $\hat{\boldsymbol{x}}^i := \{\theta^i, g^i, p^i, b^i\}$, $\boldsymbol{y}^i := \{\hat{\boldsymbol{x}}^i, \hat{\mathbf{f}}^i\}$, and $\boldsymbol{z}^e := \{f^e, \hat{\boldsymbol{\theta}}^e\}$. With these notational conventions, it turns out that (18) can be reformulated as

$$\min_{\{\boldsymbol{y}^i, \boldsymbol{z}^e\}} \sum_{i=1}^{N} u^i(\boldsymbol{y}^i) + \sum_{e=1}^{E} v^e(\boldsymbol{z}^e) \tag{22a}$$

$$\text{s.t. } \hat{\mathbf{f}}^i = \mathbf{f}^i, \ \forall i \in \mathcal{N} \tag{22b}$$

$$\hat{\boldsymbol{\theta}}^e = \boldsymbol{\theta}^e, \ \forall e \in \mathcal{E}. \tag{22c}$$

In (22), $u^i(\boldsymbol{y}^i) := \Phi^i(\hat{\boldsymbol{x}}^i) + \check{\lambda}^i b^i$; the domain of $\boldsymbol{y}^i$ is $\{\text{Dom}(\hat{\boldsymbol{x}}^i), g^i + p^i + r^i + \sum_{j\in E(i)} \mathbf{A}^{ij}\hat{f}^{i,j} = b^i + d^i\}$, where $\text{Dom}(\hat{\boldsymbol{x}}^i)$ is the set defined by the i-th inequality in (4e) and (7); $\hat{f}^{i,j}$ denotes the jth entry of $\hat{\mathbf{f}}^i$; and $v^e(\boldsymbol{z}^e)$ is the indicator function of sets $\{\underline{f}^e \le f^e \le \overline{f}^e, \mathbf{D}^{ee}\sum_{j\in V(e)} \mathbf{A}^{je}\hat{\theta}^{e,j} = f_e\}$, where f^e is the eth entry of $\mathbf{f}$, and $\hat{\theta}^{e,j}$ is the jth entry of $\hat{\boldsymbol{\theta}}^e$.

Based on (22), we will rely on the ADMM to reformulate (18) into a fully separable form. To this end, consider the augmented Lagrangian of (22)

$$\begin{aligned} L_\delta(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\rho}, \boldsymbol{\omega}) = &\sum_{i=1}^{N} u^i(\boldsymbol{y}^i) + \sum_{e=1}^{E} v^e(\boldsymbol{z}^e) + \sum_{i=1}^{N} \boldsymbol{\rho}^{i\top}\left(\hat{\mathbf{f}}^i - \mathbf{f}^i\right) \\ &+ \sum_{e=1}^{E} \boldsymbol{\omega}^{e\top}\left(\hat{\boldsymbol{\theta}}^e - \boldsymbol{\theta}^e\right) + \frac{\delta}{2}\sum_{i=1}^{N} \|\hat{\mathbf{f}}^i - \mathbf{f}^i\|^2 \\ &+ \frac{\delta}{2}\sum_{e=1}^{E} \|\hat{\boldsymbol{\theta}}^e - \boldsymbol{\theta}^e\|^2 \end{aligned} \tag{23}$$

where $\boldsymbol{\rho}^i \in \mathbb{R}^{|E(i)|}$, and $\boldsymbol{\omega}^e \in \mathbb{R}^2$ are the Lagrange multipliers for (22b) and (22c), respectively. With k denoting the iteration index, the ADMM updates $\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\rho}, \boldsymbol{\omega}$ recursively as

$$\boldsymbol{y}_{k+1} = \arg\min_{\boldsymbol{y}} L_\delta(\boldsymbol{y}, \boldsymbol{z}_k, \boldsymbol{\rho}_k, \boldsymbol{\omega}_k) \tag{24a}$$

$$\boldsymbol{z}_{k+1} = \arg\min_{\boldsymbol{z}} L_\delta(\boldsymbol{y}_{k+1}, \boldsymbol{z}, \boldsymbol{\rho}_k, \boldsymbol{\omega}_k) \tag{24b}$$

$$\boldsymbol{\rho}^i_{k+1} = \boldsymbol{\rho}^i_k + \delta\left(\hat{\mathbf{f}}^i_{k+1} - \mathbf{f}^i_{k+1}\right), \ \forall i \in \mathcal{N} \tag{24c}$$

$$\boldsymbol{\omega}^e_{k+1} = \boldsymbol{\omega}^e_k + \delta\left(\hat{\boldsymbol{\theta}}^e_{k+1} - \boldsymbol{\theta}^e_{k+1}\right), \ \forall e \in \mathcal{E}. \tag{24d}$$

Clearly, the update for $\boldsymbol{y}_k$ is separable per node, and that for $\boldsymbol{z}_k$ is separable per edge. Hence, (24) can be rewritten as

$$\begin{aligned} \boldsymbol{y}^i_{k+1} = \arg\min_{\boldsymbol{y}^i} &u^i(\boldsymbol{y}^i) + \frac{\delta}{2}\left\|\hat{\mathbf{f}}^i - \mathbf{f}^i_k + \boldsymbol{\rho}^i_k/\delta\right\|^2 \\ &+ \frac{\delta}{2}\sum_{e\in E(i)}\left(\hat{\theta}^{e,n(i)}_k - \theta^i + \omega^{e,n(i)}_k/\delta\right)^2, \ \forall i \in \mathcal{N} \end{aligned} \tag{25a}$$

and likewise

$$\begin{aligned} \boldsymbol{z}^e_{k+1} = \arg\min_{\boldsymbol{z}^e} &v^e(\boldsymbol{z}^e) + \frac{\delta}{2}\left\|\hat{\boldsymbol{\theta}}^e - \boldsymbol{\theta}^e_{k+1} + \boldsymbol{\omega}^i_k/\delta\right\|^2 \\ &+ \frac{\delta}{2}\sum_{i\in V(e)}\left(\hat{f}^{i,n(e)}_{k+1} - f^e + \rho^{i,n(e)}_k/\delta\right)^2, \ \forall e \in \mathcal{E} \end{aligned} \tag{25b}$$

where $\theta^{e,n(i)}$ denotes the entry of $\boldsymbol{\theta}^e$ associated with bus i, and $f^{i,n(e)}$ denotes the entry of $\mathbf{f}^i$ corresponding to edge e.

In a nutshell, per iteration, the following steps are performed in a distributed manner.

(S1) Each node $i \in \mathcal{N}$ updates $\boldsymbol{y}^i$ via (25a), and then sends $\boldsymbol{\theta}^e_{k+1}$ and $\hat{\mathbf{f}}^i_{k+1}$ over to all connected edges;

(S2) Each edge $e \in \mathcal{E}$ performs (25b), and then sends a message including $\mathbf{f}^i_{k+1}$ and $\hat{\boldsymbol{\theta}}^e_{k+1}$ to its neighboring nodes;

(S3) Each node $i \in \mathcal{N}$ updates ρ^i via (24c); and,

(S4) Each edge $e \in \mathcal{E}$ updates ω^e via (24d).

The complexity of solving (25a) per node and (25b) per edge is $\mathcal{O}((|E(i)|+4)^{3.5})$ and $\mathcal{O}(3^{3.5})$, respectively, when relying on an interior point solver. Meanwhile, in S1 and S2, communication only takes place between adjacent nodes and edges in the order of $\mathcal{O}(|E(i)|+2)$, thus avoiding voluminous communication overheads. In addition, we have [26].

Proposition 2: Under mild conditions, the sequence of objective values under $\{\boldsymbol{y}^i_k\}$ and $\{\boldsymbol{z}^e_k\}$ converges to the optimal objective value of (18) at sub-linear rate $\mathcal{O}(1/k)$.

Remark 3: While a DC power flow model is considered in this paper, the proposed algorithm can readily tackle the linearized AC power flow model, with the same theoretical guarantees. Regarding the nonconvex full AC model, the corresponding nonconvex subproblem efficiently [see (22)] is challenging to solve, and constitutes an interesting future research direction.

V. Optimality and Stability of OLAM

In this section, we analytically establish the performance of OLAM, and highlight its attractive properties relative to existing schemes. Before assessing the performance, we adopt the following assumptions.

(as1) The microgrid state $\boldsymbol{s}_t$ is independent and identically distributed (i.i.d.) over time t with bounded support.

(as2) $\Phi_t(\mathbf{x}_t)$ is convex, and has Lipschitz continuous gradient.

(as3) There exists a stationary policy $\chi(\cdot)$ satisfying $\chi(\boldsymbol{s}_t) \in \mathcal{X}$ for all $\boldsymbol{s}_t$ and $\mathbb{E}[\mathbf{b}(\boldsymbol{s}_t)] \le -\zeta$, where $\zeta > \mathbf{0}$.

(as4) The dual function $\mathcal{D}_t(\boldsymbol{\lambda})$ is ϵ-strongly concave, and $\nabla\mathcal{D}_t(\boldsymbol{\lambda})$ is L-Lipschitz continuous.

Assumptions (as1)-(as2) are typical in stochastic optimization and LO-based approaches; (as3) is Slater's condition which assures the existence of a bounded optimal Lagrange multiplier; and (as4) requires the dual function to be well behaved. If the dual function is not strongly concave, one can subtract an ℓ_2-regularizer from it to obtain a suboptimal solution, which would not affect the order of sub-optimality [27, Lemma 3.2]. Under (as1)-(as4), we first show that the empirical dual variables in $\hat{\boldsymbol{\lambda}}_t$ converge to their optimal values.

Theorem 1: If the stepsize satisfies $\{\eta_t = \frac{\alpha D}{M\sqrt{t}}, \forall t\}$, where α is a scaling factor, D is the diameter of the feasible set of $\hat{\boldsymbol{\lambda}}_t$, and $M = \max_t \|\mathbf{b}_t\|$, the empirical dual variable will converge to its optimal value with probability 1; that is

$$\lim_{t\to\infty} \hat{\boldsymbol{\lambda}}_t = \boldsymbol{\lambda}^*, \ \text{w.p.1.} \tag{26}$$

Proof: See [28, Proposition 8.2.13]. ■

For the battery level behavior, since the transient stage is rather ephemeral compared with the steady state, we could omit it, and focus only on the steady state. Define a biased residual of the second gradient, namely, $\tilde{\boldsymbol{\beta}}_t := \boldsymbol{\lambda}^* - \hat{\boldsymbol{\lambda}}_t + \boldsymbol{\beta}$. It follows from Theorem 1 that $\lim_{t\to\infty} \tilde{\boldsymbol{\beta}}_t = \boldsymbol{\beta}$. To establish feasibility and optimality of OLAM, the following lemma is handy.

Lemma 1: With $\boldsymbol{\beta}/\mu \leq \overline{\mathbf{q}}$, there exists a constant $B = \Theta(1/\sqrt{\mu})$, and a finite time T_β depending on β, such that for all $t > T_\beta$, when $\|\mathbf{q}_t - \tilde{\boldsymbol{\beta}}/\mu\| > B$, it holds that

$$\mathbb{E}\left[\left\|\mathbf{q}_{t+1} - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\| |\mathbf{q}_t\right] \leq \left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\| - \sqrt{\mu}. \tag{27}$$

Proof: See Appendix A. ∎

Lemma 1 asserts that the battery state always tracks a time-varying variable $\tilde{\boldsymbol{\beta}}/\mu$, which converges to $\boldsymbol{\beta}/\mu$ as $t \to \infty$. Building on Lemma 1, we next establish that the battery state will not violate the battery capacity constraint (4g) frequently.

Lemma 2: Define the instantaneous violation of the battery's lower bound violation $\boldsymbol{l}_t$, and the upper bound violation $\boldsymbol{u}_t$ as

$$\boldsymbol{l}_t := \max\{\mathbf{q}_t - (\mathbf{q}_{t-1} + \mathbf{b}_{t-1}), \mathbf{0}\} \tag{28}$$

$$\boldsymbol{u}_t := \max\{(\mathbf{q}_{t-1} + \mathbf{b}_{t-1}) - \mathbf{q}_t, \mathbf{0}\}. \tag{29}$$

If we set $\boldsymbol{\beta} = \mathcal{O}(\sqrt{\mu}\log^2(\mu)) \cdot \mathbf{1}$ with $2\boldsymbol{\beta}/\mu \leq \overline{\mathbf{q}} + \underline{\mathbf{q}}$, then the time-average violations are bounded by

$$\left\|\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbf{u}_t\right\| \leq \left\|\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbf{l}_t\right\| = \mathcal{O}(\mu). \tag{30}$$

Proof: See Appendix B. ∎

Lemma 2 demonstrates that the magnitude of the limiting time-averages of $\boldsymbol{l}_t$ and $\boldsymbol{u}_t$ is on the order of $\mathcal{O}(\mu)$, which implies that by choosing $2\boldsymbol{\beta}/\mu \leq \overline{\mathbf{q}} + \underline{\mathbf{q}}$, OLAM guarantees $\mathcal{O}(\mu)$ violation of the battery capacity constraint. This result is instructive because to ensure feasibility of the battery, the projected update is taken in (19), and the actual (dis)charging amount will be adjusted accordingly. This suggests buying extra energy from the market when violating the lower bound, and selling to the market when violating the upper bound. As this introduces extra cost for power balancing (buying from and selling to the market), Lemma 2 establishes that the limiting time-average of the adjusted amount is $\mathcal{O}(\mu)$. Together with the Lipschitz continuity condition of the objective in (as2), this implies that the extra buying and selling costs due to the battery projection will *not affect the order of sub-optimality*, if we are after an $\mathcal{O}(\mu)$-optimal online solution. We formally establish this main result in the next theorem.

Theorem 2: Let Φ^* be the optimal objective value of (4) under any feasible policy with distribution information about the random states available. If the control variable is chosen as $\boldsymbol{\beta} = \mathcal{O}(\sqrt{\mu}\log^2(\mu)) \cdot \mathbf{1}$ to satisfy $2\boldsymbol{\beta}/\mu \leq \overline{\mathbf{q}} + \underline{\mathbf{q}}$, then with a proper μ, OLAM yields a near optimal solution; that is

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\Big[\Phi_t\Big(\mathbf{x}_t\Big(\check{\boldsymbol{\lambda}}_t\Big)\Big)\Big] \leq \Phi^* + \mathcal{O}(\mu) \tag{31}$$

where $\mathbf{x}_t(\check{\boldsymbol{\lambda}}_t)$ is the online solution obtained from (18).

Proof: See Appendix C. ∎

Upon selecting $\boldsymbol{\beta} = \mathcal{O}(\sqrt{\mu}\log^2(\mu)) \cdot \mathbf{1}$, the battery capacity in Theorem 2 will satisfy $\overline{\mathbf{q}} = \mathcal{O}(\log^2(\mu)/\sqrt{\mu}) \cdot \mathbf{1}$ [24, Th. 2]. Theorem 2 states that under proper choice of parameters, OLAM incurs an $\mathcal{O}(\mu)$ suboptimality gap in the steady state, while the order of aggregated battery levels is $\mathcal{O}(\log^2(\mu)/\sqrt{\mu})$, which markedly improves upon the standard cost-capacity tradeoff $[\mathcal{O}(\mu), \mathcal{O}(1/\mu)]$ of SDG [6], [11], [14], [18] in the near-optimal regime, i.e., with a sufficiently small μ. This suggests deploying low-capacity batteries with reduced installment cost, while implementing online energy management schemes with near-optimal network energy cost.

VI. Numerical Tests

This section presents numerical tests to confirm the analytical claims and demonstrate the merits of OLAM. All the tests are implemented with MATLAB 2017 on a desktop with Intel Core i7-4790 3.60GHz CPU. The proposed energy management scheme is numerically tested on the IEEE 14-bus test system with $N = 14$ buses, and $E = 20$ transmission lines [29]. Buses 1, 2, 3, 6 and 8 have a battery, a renewable generator, and a conventional generator. We test the performance of the energy management schemes in terms of the time-averaged instantaneous cost in (3), namely,

$$\Phi(\mathbf{x}_t; \boldsymbol{s}_t) = \sum_{i=1}^{N} \alpha_t^i p_t^i + \sum_{i=1}^{N} c_1^i \left(g_t^i\right)^2 + \sum_{i=1}^{N} \left(c_2^i \left(b_t^i\right)^2 + c_3^i b_t^i\right) \tag{32}$$

where $c_1^i = 1.2, c_2^i = 0.5, c_3^i = 1$, $\forall i \in \{1, 2, 3, 6, 8\}$, and $c_1^i = c_2^i = c_3^i = 0$ for other i, these parameters are all zero. The price α_t^i is uniformly distributed over [12, 18]; samples of the renewable supply are uniformly generated over $[\underline{r_t^i}, \overline{r_t^i}]$, in which $\underline{r_t^i}$ is generated uniformly over [6.15, 10.7], and $\overline{r_t^i}$ is generated from a [12, 15] for bus $i \in \{1, 2, 3, 6, 8\}$. The capacity of each line e is set as $\overline{f}^e = 6$. For bus $i \in \{1, 2, 3, 6, 8\}$, the capacity of each generator is in [0, 30]. The demand d_t^i for bus i , stays within [28, 40]. Finally, for all buses equipped with batteries, the limits of charging and discharging are $\overline{b}^i = 7$, and $\underline{b}^i = 7$; the (dis)charging coefficient is set to $\kappa = 0.9975$, which corresponds to NaS batteries. The battery limits are $\overline{q}^i = 30$ and $\underline{q}^i = 3$.

Regarding the parameters of OLAM, $\mu = 0.05$, the bias control variable is $\boldsymbol{\beta} = 0.22\sqrt{\mu}\log^2(\mu) \cdot \mathbf{1}$, and the stepsize $\eta_t = 1.2/\sqrt{t}$. The distributed version of OLAM (d-OLAM) is simulated with $\delta = 1$. For comparison, we also consider three alternative schemes: i) SDG [14] with stepsize $\mu = 0.05$; and, ii) the greedy one that myopically minimizes the per-slot cost.

Fig. 2(a) compares the time-averaged costs for all three schemes. We observe that SDG and OLAM have almost identical cost, which is significantly (50.7%) less than that of greedy algorithm. It is also shown that OLAM enjoys a faster convergence rate than SDG, as OLAM benefits from its statistical learning module that accelerates convergence of the effective multiplier $\check{\boldsymbol{\lambda}}_t$ to the neighborhood of $\boldsymbol{\lambda}^*$.

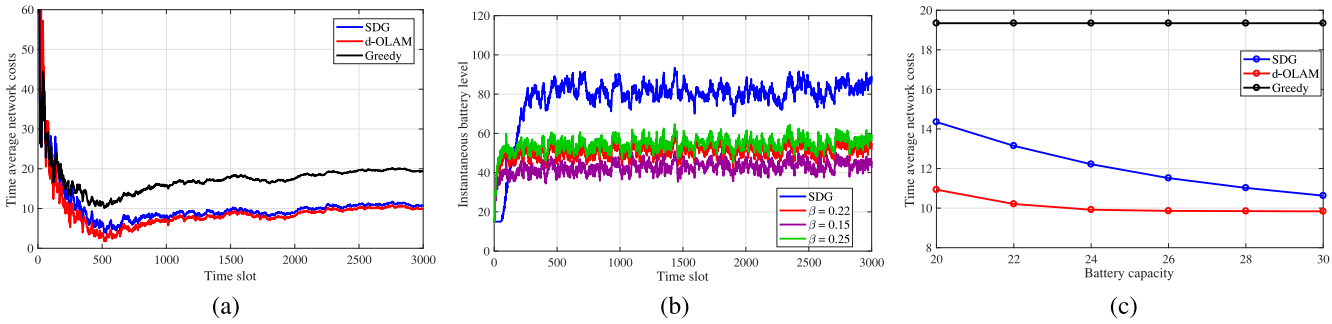


Fig. 2. (a) Comparison of time-average network costs; (b) Battery energy levels averaged over all nodes; (c) Time-average cost versus battery capacity.

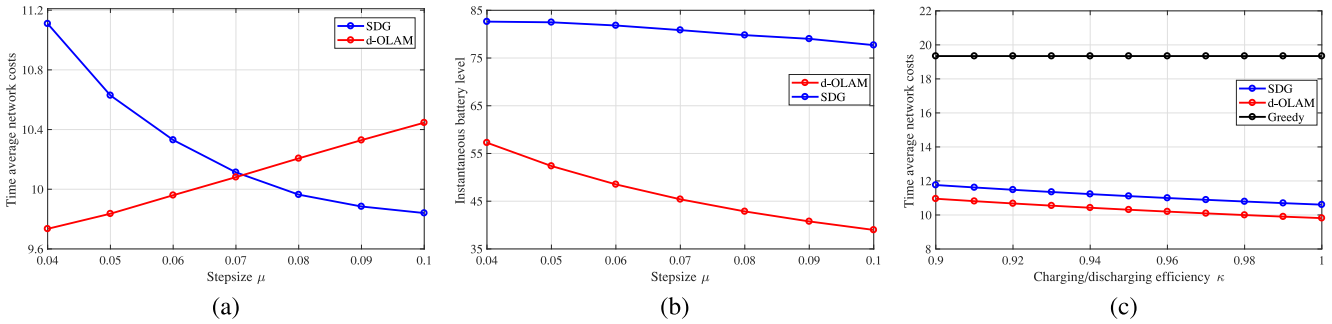


Fig. 3. (a) Time-average network cost versus stepsize μ; (b) Time-average battery level versus stepsize μ; (c) Network cost versus (dis)charging efficiency.

Fig. 2(b) depicts the real-time battery levels with the d-OLAM and SDG schemes. It is observed that the batteries must store a larger amount of energy in the steady state under SDG, which is not efficient. In contrast, the maintained energy amount with d-OLAM is only 63.4% of that with SDG. Hence, OLAM can afford small-capacity storage devices with lower installation cost for efficient energy management.

To further demonstrate the merits of OLAM, Fig. 2(c) compares the time-average costs with SDG and d-OLAM when different battery capacities (energy storage upper bounds) are employed. For SDG, the cost quickly scales up as the battery capacity decreases. On the other hand, when the battery capacity $\overline{\mathbf{q}}$ decreases, the probability of battery over- and under-flow with d-OLAM increases only slightly. Thus, it is possible for d-OLAM to adopt batteries with constant battery capacity $\overline{q} = 20$ units or even low $\overline{\mathbf{q}}$, while only incurring a slightly rise in cost. From an economical viewpoint, d-OLAM can achieve close-to-optimal performance without costly large batteries, thereby facilitating low-cost deployment of microgrids.

Performance of d-OLAM is further tested under various choices of the parameters; see Fig. 3. Fig. 3(a) shows the relation between μ and time-average costs, which suggests that when μ is less than 0.07, d-OLAM outperforms SDG. Fig. 3(b) compares μ versus battery level after $3,000$ iterations, where d-OLAM's battery level is markedly reduced comparing with SDG. Furthermore, Fig. 3 (c) depicts how κ affects time-average costs. Clearly, d-OLAM incurs significantly lower cost than that of the greedy algorithm, and slightly lower compared with SDG. Regarding running time, the centralized SDG takes 0.63s to run one iteration on average, while d-OLAM requires 2.73s for one iteration in our test.

VII. Conclusion

Real-time energy management was investigated for smart microgrids with RES and energy storage units. A novel learning-aided energy management algorithm, that we termed OLAM was developed for both centralized and distributed settings. Different from the celebrated SDG iterations, OLAM leverages a statistical learning process to acquire the needed statistics and mitigate the associated uncertainties. Thanks to the statistical learning step, OLAM was proven to obtain an $\mathcal{O}(\mu)$ sub-optimality gap while requiring an $\mathcal{O}(\log^2(\mu)/\sqrt{\mu})$ battery capacity, which markedly outperforms the standard $[\mathcal{O}(\mu), \mathcal{O}(1/\mu)]$ cost-capacity tradeoff of SDG. To facilitate its practical implementation, a decentralized version of OLAM that we termed d-OLAM was introduced based on ADMM. Using the IEEE 14-bus power grid benchmark, numerical tests demonstrated that the d-OLAM attains slightly lower cost than SDG, while requiring much lower battery capacity.

Our novel OLAM-based energy management approach opens up interesting research directions, which include designing efficient algorithms for solving inner-loop problems (e.g., Lagrangian minimization) to reduce communication overhead, and providing theoretical justification for finite-horizon operation. Finally, incorporating the full AC power flow models and multi-timescale dispatch, and properly handling generator ramping constraints are intriguing research directions too.

Appendix A
Proof of Lemma 1

For a convergent $\hat{\boldsymbol{\lambda}}_t$, there always exists a time T_β, such that for all $t > T_\beta$, $\|\hat{\boldsymbol{\lambda}}_t - \boldsymbol{\lambda}^*\| \leq \beta$ holds; hence, we have

$\mathbf{0} < \tilde{\boldsymbol{\beta}}_t < 2\boldsymbol{\beta} \le 2\mu\overline{\mathbf{q}}$. We will show that

$$\left\| \left[\tilde{\boldsymbol{\beta}}_t/\mu\right]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - \tilde{\boldsymbol{\beta}}_t/\mu \right\| \le \|\overline{\mathbf{q}}\|. \tag{33}$$

Indeed, letting $\tilde{\beta}_t^i$, $\underline{q}^i$, and $\overline{q}^i$ denote the ith entry of $\tilde{\boldsymbol{\beta}}_t$, $\underline{\mathbf{q}}$, and $\overline{\mathbf{q}}$, respectively, consider the following three cases:

i) if $\frac{\tilde{\beta}_t^i}{\mu} \le \underline{q}^i$, we have $|[\frac{\tilde{\beta}_t^i}{\mu}]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - \frac{\tilde{\beta}_t^i}{\mu}| \le \underline{q}^i \le \overline{q}^i$;

ii) if $\underline{q}^i \le \frac{\tilde{\beta}_t^i}{\mu} \le \overline{q}^i$, we have $|[\frac{\tilde{\beta}_t^i}{\mu}]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - \frac{\tilde{\beta}_t^i}{\mu}| = 0 \le \overline{q}^i$;

iii) if $\frac{\tilde{\beta}_t^i}{\mu} > \overline{q}^i$, $|[\frac{\tilde{\beta}_t^i}{\mu}]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - \frac{\tilde{\beta}_t^i}{\mu}| \le |2\overline{q}^i - \overline{q}^i| = \overline{q}^i$.

Cases i)–iii) together imply (33).

For a constant $M_1 \coloneqq \|\overline{\mathbf{q}}\|^2 + 2\|\overline{\mathbf{q}}\|\|\overline{\mathbf{q}} - \underline{\mathbf{q}}\|$, we have

$$\begin{aligned}
\left\|\mathbf{q}_{t+1} - \tilde{\boldsymbol{\beta}}_t/\mu\right\|^2 &= \left\|\left[\mathbf{q}_t + \mathbf{b}_t\right]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - \tilde{\boldsymbol{\beta}}_t/\mu\right\|^2 \\
&= \left\|\left[\mathbf{q}_t + \mathbf{b}_t\right]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - \left[\tilde{\boldsymbol{\beta}}_t/\mu\right]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} + \left[\tilde{\boldsymbol{\beta}}_t/\mu\right]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - \tilde{\boldsymbol{\beta}}_t/\mu\right\|^2 \\
&\overset{(a)}{\le} \left\|\left[\mathbf{q}_t + \mathbf{b}_t\right]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - \left[\tilde{\boldsymbol{\beta}}_t/\mu\right]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}}\right\|^2 + M_1^2 \\
&\overset{(b)}{\le} \left\|\mathbf{q}_t + \mathbf{b}_t - \tilde{\boldsymbol{\beta}}_t/\mu\right\|^2 + M_1^2
\end{aligned} \tag{34}$$

where (a) is due to (33), and $\|[\mathbf{q}_t + \mathbf{b}_t]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}} - [\frac{\tilde{\boldsymbol{\beta}}_t}{\mu}]_{\underline{\mathbf{q}}}^{\overline{\mathbf{q}}}\| \le \|\overline{\mathbf{q}} - \underline{\mathbf{q}}\|$; and (b) follows from the non-expansive property of projection. Expanding the norm square, we arrive at

$$\begin{aligned}
\left\|\mathbf{q}_{t+1} - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2 &\overset{(c)}{\le} \left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2 + 2\mathbf{b}_t^\top\left(\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right) + M^2 \\
&= \left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2 + \frac{2}{\mu}\mathbf{b}_t^\top\left(\check{\boldsymbol{\lambda}}_t - \boldsymbol{\lambda}^*\right) + M^2 \\
&\overset{(d)}{\le} \left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2 + \frac{2}{\mu}\left[\mathcal{D}_t\left(\check{\boldsymbol{\lambda}}_t\right) - \mathcal{D}_t\left(\boldsymbol{\lambda}^*\right)\right] + M^2
\end{aligned} \tag{35}$$

where (c) uses $M^2 \coloneqq \max\{\mu^2, M_1^2\} + \max\{\|\kappa\overline{\mathbf{b}}\|^2, \|\underline{\mathbf{b}}/\kappa\|^2\}$; and, (d) holds since $\mathbf{b}_t$ is the gradient of the concave function $\mathcal{D}_t(\boldsymbol{\lambda})$ at $\check{\boldsymbol{\lambda}}_t$. By the strong concavity of $\mathcal{D}_t(\boldsymbol{\lambda})$, we also have

$$\begin{aligned}
\mathcal{D}_t\left(\check{\boldsymbol{\lambda}}_t\right) - \mathcal{D}_t\left(\boldsymbol{\lambda}^*\right) &\le \nabla\mathcal{D}_t\left(\boldsymbol{\lambda}^*\right)^\top\left(\check{\boldsymbol{\lambda}}_t - \boldsymbol{\lambda}^*\right) - \frac{\epsilon}{2}\left\|\check{\boldsymbol{\lambda}}_t - \boldsymbol{\lambda}^*\right\|^2 \\
&= -\frac{\epsilon}{2}\left\|\check{\boldsymbol{\lambda}}_t - \boldsymbol{\lambda}^*\right\|^2
\end{aligned} \tag{36}$$

where the last equality holds since $\nabla\mathcal{D}_t(\boldsymbol{\lambda}^*) = \mathbf{0}$ [24, Prop. 3].

Substituting (36) into (35), and taking expectation, we have

$$\mathbb{E}\left[\left\|\mathbf{q}_{t+1} - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2\right] \le \left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2 - \mu\epsilon\left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2 + M^2. \tag{37}$$

By Vieta's formula, there always exists

$$B = \frac{-\sqrt{\mu} + \sqrt{\mu + \mu\epsilon\left(M^2 - \mu\right)}}{\mu\epsilon} = \boldsymbol{\Theta}\left(\frac{1}{\sqrt{\mu}}\right) \tag{38}$$

such that when $\|\mathbf{q}_t - \tilde{\boldsymbol{\beta}}/\mu\| > B$, we have

$$-\mu\epsilon\left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2 + M^2 \le -2\sqrt{\mu}\left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\| + \mu. \tag{39}$$

Due to (39) and the convexity of the quadratic function, it follows from (37) that

$$\begin{aligned}
\mathbb{E}\left[\left\|\mathbf{q}_{t+1} - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|\right]^2 &\le \mathbb{E}\left[\left\|\mathbf{q}_{t+1} - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\|^2\right] \\
&\le \left(\left\|\mathbf{q}_t - \frac{\tilde{\boldsymbol{\beta}}_t}{\mu}\right\| - \sqrt{\mu}\right)^2
\end{aligned} \tag{40}$$

from which the lemma follows readily.

Appendix B
Proof of Lemma 2

Note that $2\boldsymbol{\beta}/\mu \le \overline{\mathbf{q}} + \underline{\mathbf{q}}$ clearly implies $\boldsymbol{\beta}/\mu \le \overline{\mathbf{q}}$; hence, Lemma 1 holds.

Based on Lemma 1, we can use arguments similar to those in [30, Th. 1(b)] to establish the following bound

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{P}\left[\left\|q_t^i - \frac{\beta}{\mu}\right\| > \tilde{B} + d\right] \le D_1 e^{-D_2 d} \tag{41}$$

with $\tilde{B} \coloneqq B + \theta = \Theta(1/\sqrt{\mu})$, where θ is a constant satisfying $\sqrt{\mu} - 2\theta < 0$, $D_1 = \Theta(1/\mu)$, and $D_2 = \Theta(\sqrt{\mu})$. Based on (41), we have for every battery, it follows that

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{P}\left[q_t^i > \frac{\beta}{\mu} + \tilde{B} + d\right] \le D_1 e^{-D_2 d} \tag{42}$$

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{P}\left[q_t^i < \frac{\beta}{\mu} - \tilde{B} - d\right] \le D_1 e^{-D_2 d}. \tag{43}$$

Inequalities (42) and (43) indicate that the fraction of slots that the battery level is above or below a certain value is bounded.

It is easy to see that

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbf{b}_t = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{u}_t - \boldsymbol{l}_t) + \lim_{T\to\infty}\frac{\mathbf{q}_T - \mathbf{q}_0}{T}. \tag{44}$$

Since $\mathbf{q}_t$ and $\mathbf{q}_0$ are both bounded, the last term in the right-hand-side of (44) approaches 0, which implies that

$$-\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{l}_t \le \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbf{b}_t \le \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{u}_t. \tag{45}$$

Upon defining $b_{\max}^i \coloneqq \max_t |b_t^i|$ and plugging (42) and (43) into (45), we deduce that

$$\begin{aligned}
\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}u_t^i &\overset{(a)}{\le} b_{\max}^i \cdot \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{P}\left[q_t^i > \overline{q}^i - b_{\max}^i\right] \\
&\overset{(b)}{\le} b_{\max}^i D_1 \exp\left[-D_2\left(-\frac{\beta}{\mu} - \tilde{B} - b_{\max}^i + \overline{q}^i\right)\right]
\end{aligned} \tag{46}$$

where (a) follows since projection on the queue's upper bound happens only when $q_t^i > \overline{q}^i - b_{\max}^i$, and the largest amount of projection is $b_{\max}^i$. Similarly, we also have

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} l_t^i \le b_{\max}^i D_1 \exp\left[-D_2\left(\frac{\beta}{\mu}-\tilde{B}-b_{\max}^i-\underline{q}^i\right)\right]. \tag{47}$$

Comparing (46) with (47), we deduce that when $2\beta/\mu \le \overline{q}^i+\underline{q}^i$, the bound of $\sum_{t=1}^T l_t^i$ dominates that of $\sum_{t=1}^T u_t^i$. Since we are seeking high energy efficiency, i.e., low battery level at steady state, we should choose $2\boldsymbol{\beta}/\mu \le \overline{\mathbf{q}}+\underline{\mathbf{q}}$, such that the bound with l^i dominates. Then selecting a small μ such that $\exp[-D_2(\frac{\beta}{\mu}-\tilde{B}-b_{max}^i-\underline{q}^i)] \le \mu^2$, we have $b_{max}^i D_1 \exp[-D_2(-\tilde{B}-b_{max}^i-\underline{q}^i)] = \mathcal{O}(\mu)$, which implies

$$\left\|\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} l_t^i\right\| = \mathcal{O}(\mu). \tag{48}$$

and completes the proof.

Appendix C
Proof of Theorem 2

With $\Delta(\mathbf{q}_t) := (1/2)\mathbb{E}[(\|\mathbf{q}_{t+1}\|^2 - \|\mathbf{q}_t\|^2)|\mathbf{q}_t]$, using arguments similar to those in (35), we have

$$\begin{aligned}\mathbb{E}\left[\|\mathbf{q}_{t+1}\|^2|\mathbf{q}_t\right] &\le \mathbb{E}\left[\|\mathbf{q}_t\|^2\right] + \mathbb{E}\left[\mathbf{b}_t^\top\mathbf{q}_t\right] + \frac{1}{2}M_2^2 \\ &\le \mathbb{E}\left[\|\mathbf{q}_t\|^2\right] + \frac{1}{\mu}\mathbb{E}\left[\mathbf{b}_t^\top\left(\check{\boldsymbol{\lambda}}_t-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\beta}\right)\right] + \frac{1}{2}M_2^2\end{aligned} \tag{49}$$

where $M_2 := \sqrt{\max\{\|\kappa\overline{\mathbf{b}}\|^2, \|\underline{\mathbf{b}}/\kappa\|^2\} + \|\overline{\mathbf{q}}\|^2}$.

Building on (49), we find

$$\begin{aligned}&\mu\Delta(\mathbf{q}_t) + \mathbb{E}\left[\Phi_t\left(\mathbf{x}_t\left(\check{\boldsymbol{\lambda}}_t\right)\right)|\mathbf{q}_t\right] \\ &= \mathbb{E}\left[\Phi_t\left(\mathbf{x}_t\left(\check{\boldsymbol{\lambda}}_t\right)|\mathbf{q}_t\right]\right] + \mathbb{E}\left[\mathbf{b}_t^\top\left(\check{\boldsymbol{\lambda}}_t-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\beta}\right)\right] + \frac{\mu}{2}M_2^2 \\ &\overset{(a)}{=} \mathbb{E}\left[\mathcal{L}_t\left(\mathbf{x}_t\left(\check{\boldsymbol{\lambda}}_t\right),\check{\boldsymbol{\lambda}}_t\right)\right] + \mathbb{E}\left[\mathbf{b}_t^\top\left(-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\beta}\right)|\mathbf{q}_t\right] + \frac{\mu}{2}M_2^2 \\ &\overset{(b)}{=} \mathcal{D}\left(\check{\boldsymbol{\lambda}}_t\right) + \mathbb{E}\left[\mathbf{b}_t^\top\left(-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\beta}\right)|\mathbf{q}_t\right] + \frac{\mu}{2}M_2^2 \\ &\overset{(c)}{\le} \Phi^* + \mathbb{E}\left[\mathbf{b}_t^\top\left(-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\beta}\right)|\mathbf{q}_t\right] + \frac{\mu}{2}M_2^2\end{aligned} \tag{50}$$

where (a) comes from the definition of the Lagrangian $\mathcal{L}_t(\mathbf{x}_t, \check{\boldsymbol{\lambda}}_t)$; (b) is due to the fact that $\mathbf{x}_t(\check{\boldsymbol{\lambda}}_t)$ minimizes $\mathcal{L}_t(\mathbf{x}_t, \check{\boldsymbol{\lambda}}_t)$; and (c) follows since $\mathcal{D}(\check{\boldsymbol{\lambda}}_t) \le \tilde{\Phi}^* \le \Phi^*$.

Taking expectation over $\mathbf{q}_t$ on both sides of (50), and summing across the time horizon, we arrive at

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\Phi_t\left(\mathbf{x}_t\left(\check{\boldsymbol{\lambda}}_t\right)\right)\right] \\ &\le \Phi^* + \frac{\mu}{2}M_2^2 + \lim_{T\to\infty}\frac{\mu\|\mathbf{q}_0\|^2}{2T} + \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\mathbf{b}_t^\top\left(\boldsymbol{\beta}-\hat{\boldsymbol{\lambda}}_t\right)\right] \\ &= \Phi^* + \frac{\mu}{2}M_2^2 + \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\mathbf{b}_t^\top\left(-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\beta}\right)\right]\end{aligned} \tag{51}$$

where the last equality holds because $\mathbf{q}_0$ is bounded. By Lemma 2 and (45), we clearly have

$$\left\|\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} b_t^i\right\| = \mathcal{O}(\mu). \tag{52}$$

Now check the last term in (51) to obtain

$$\begin{aligned}&\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\mathbf{b}_t^\top\left(-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\beta}\right)\right] \\ &= \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\mathbf{b}_t^\top\left(-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\lambda}^*+\boldsymbol{\beta}-\boldsymbol{\lambda}^*\right)\right] \\ &\le \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\mathbf{b}_t^\top\left(\boldsymbol{\beta}-\boldsymbol{\lambda}^*\right)\right] + \overline{\mathbf{b}}^\top\mathbb{E}\left[\left(-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\lambda}^*\right)\right] \\ &\overset{(d)}{\le} \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\mathbf{b}_t^\top\left(\boldsymbol{\beta}-\boldsymbol{\lambda}^*\right)\right] + o(\mu) \\ &\overset{(e)}{\le} \left\|\boldsymbol{\beta}-\boldsymbol{\lambda}^*\right\|\cdot\left\|\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\mathbf{b}_t]\right\| + o(\mu)\end{aligned} \tag{53}$$

where (d) holds since convergence of the empirical variable implies that for any $\sigma = o(\mu)$, there exists a finite T_σ such that $\|-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\lambda}^*\| \le \sigma$ for any $t > T_\sigma$; and (i) follows from Cauchy-Schwarz's inequality. Combining (52) and (53), we obtain

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\mathbf{b}_t^\top\left(-\hat{\boldsymbol{\lambda}}_t+\boldsymbol{\beta}\right)\right] = \mathcal{O}(\mu). \tag{54}$$

Plugging (54) into (51), we find

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\Phi_t\left(\mathbf{x}_t\left(\check{\boldsymbol{\lambda}}_t\right)\right)\right] \le \Phi^* + \frac{\mu}{2}M_2^2 + \mathcal{O}(\mu). \tag{55}$$

References

[1] N. Hatziargyriou, H. Asano, R. Iravani, and C. Marnay, "Microgrids," *IEEE Power Energy Mag.*, vol. 5, no. 4, pp. 78–94, Jul./Aug. 2007.

[2] Y. Zhang, N. Gatsis, and G. B. Giannakis, "Robust energy management for microgrids with high-penetration renewables," *IEEE Trans. Sustain. Energy*, vol. 4, no. 4, pp. 944–953, Oct. 2013.

[3] N. Li, L. Chen, and S. H. Low, "Optimal demand response based on utility maximization in power networks," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Detroit, MI, USA, Jul. 2011, pp. 1–8.

[4] I. Koutsopoulos, V. Hatzi, and L. Tassiulas, "Optimal energy storage control policies for the smart power grid," in *Proc. IEEE SmartGridCom*, Brussels, Belgium, Oct. 2011, pp. 475–480.

[5] J. Qin, R. Sevlian, D. Varodayan, and R. Rajagopal, "Optimal electric energy storage operation," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, San Diego, CA, USA, 2012, pp. 1–6.

[6] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam, "Optimal power cost management using stored energy in data centers," in *Proc. ACM SIGMETRICS*, San Jose, CA, USA, Jun. 2011, pp. 221–232.

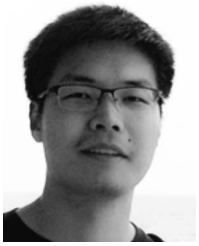
[7] Y. Guo and Y. Fang, "Electricity cost saving strategy in data centers by using energy storage," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 6, pp. 1149–1160, Jun. 2013.

[8] X. Wang, Y. Zhang, T. Chen, and G. B. Giannakis, "Dynamic energy management for smart-grid-powered coordinated multipoint systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1348–1359, May 2016.

[9] X. Wang, X. Chen, T. Chen, L. Huang, and G. B. Giannakis, "Two-scale stochastic control for integrated multipoint communication systems with renewables," *IEEE Trans. Smart Grid*, to be published, doi: 10.1109/TSG.2016.2601073.

[10] S. Gupta and V. Kekatos, "Real-time operation of heterogeneous energy storage units," in *Proc. GlobalSIP Conf.*, Washington, DC, USA, Dec. 2016, pp. 916–920.
[11] S. Sun, M. Dong, and B. Liang, "Distributed real-time power balancing in renewable-integrated power grids with storage and flexible loads," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2337–2349, Sep. 2016.
[12] Q. Li *et al.*, "Networked and distributed control method with optimal power dispatch for islanded microgrids," *IEEE Trans. Ind. Electron.*, vol. 64, no. 1, pp. 493–504, Jan. 2017.
[13] F. Chen *et al.*, "Cost-based droop schemes for economic dispatch in islanded microgrids," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 63–74, Jan. 2017.
[14] W. Shi, N. Li, C.-C. Chu, and R. Gadh, "Real-time energy management in microgrids," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 228–238, Jan. 2017.
[15] J. Qin, Y. Chow, J. Yang, and R. Rajagopal, "Distributed online modified greedy algorithm for networked storage operation under uncertainty," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1106–1118, Mar. 2016.
[16] J. P. Barton and D. G. Infield, "Energy storage and its use with intermittent renewable energy," *IEEE Trans. Energy Convers.*, vol. 19, no. 2, pp. 441–448, Jun. 2004.
[17] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," in *Synthesis Lectures on Communication Networks*, vol. 3. San Rafael, CA, USA: Morgan & Claypool, 2010, pp. 1–211.
[18] J. Qin, Y. Chow, J. Yang, and R. Rajagopal, "Online modified greedy algorithm for storage control under uncertainty," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 1729–1743, May 2016.
[19] G. B. Giannakis *et al.*, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 107–128, Sep. 2013.
[20] S. H. Low, "Convex relaxation of optimal power flow—Part I: Formulations and equivalence," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 15–27, Mar. 2014.
[21] W. B. Powell and S. Meisel, "Tutorial on stochastic optimization in energy—Part II: An energy storage illustration," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1468–1475, Mar. 2016.
[22] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Sci., 1999.
[23] T. Chen, X. Wang, and G. B. Giannakis, "Cooling-aware energy and workload management in data centers via stochastic optimization," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 2, pp. 402–415, Mar. 2016.
[24] T. Chen, Q. Ling, and G. B. Giannakis, "Learn-and-adapt stochastic dual gradients for network resource allocation," *IEEE Trans. Control Netw. Syst.*, to be published, doi: 10.1109/TCNS.2017.2774043. [Online]. Available: arxiv.org/abs/1703.01673
[25] L. Huang, X. Liu, and X. Hao, "The power of online learning in stochastic network optimization," in *Proc. ACM SIGMETRICS*, vol. 42. Austin, TX, USA, Jun. 2014, pp. 153–165.
[26] G. B. Giannakis, Q. Ling, G. Mateos, I. D. Schizas, and H. Zhu, "Decentralized learning for wireless communications and networking," in *Splitting Methods in Communication and Imaging, Science and Engineering*. Cham, Switzerland: Springer Int., 2016.
[27] J. Koshal, A. Nedić, and U. V. Shanbhag, "Multiuser optimization: Distributed algorithms and error analysis," *SIAM J. Optim.*, vol. 21, no. 3, pp. 1046–1081, 2011.
[28] D. P. Bertsekas, A. Nedi, and A. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Sci., 2003.
[29] *Power Systems Test Case Archive*, Univ. Washington, Seattle, WA, USA, Jan. 25, 2014. [Online]. Available: http://www.ee.washington.edu/research/pstca
[30] L. Huang and M. J. Neely, "Delay reduction via Lagrange multipliers in stochastic network optimization," *IEEE Trans. Autom. Control*, vol. 56, no. 4, pp. 842–857, Apr. 2011.

Bingcong Li received the B.Eng. degree (with Highest Hons.) in communication science and engineering from Fudan University in 2017. He is currently pursuing the Ph.D. degree with UMN. His research interests lie in online learning, online convex optimization, and reinforcement learning with applications to cyber physical systems. He was a recipient of the National Scholarship twice from China in 2014 and 2015 and UMN ECE Department Fellowship in 2017.



Tianyi Chen (S'14) received the B.Eng. degree (with Highest Hons.) in communication science and engineering from Fudan University, and the M.Sc. degree in electrical and computer engineering from the University of Minnesota (UMN), in 2014 and 2016, respectively, where he is currently pursuing the Ph.D. degree with UMN. His research interests lie in online learning, online convex optimization, and stochastic network optimization with applications to smart grids, sustainable cloud networks, and Internet-of-Things. He was in the Best Student Paper Award finalist of the Asilomar Conference on Signals, Systems, and Computers. He was a recipient of the National Scholarship from China in 2013, the UMN ECE Department Fellowship in 2014, and the UMN Doctoral Dissertation Fellowship in 2017.

Xin Wang (SM'09) received the B.Sc. and M.Sc. degrees from Fudan University, Shanghai, China, in 1997 and 2000, respectively, and the Ph.D. degree from Auburn University, Auburn, AL, USA, in 2004, all in electrical engineering.

From 2004 to 2006, he was a Post-Doctoral Research Associate with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis. In 2006, he joined the Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA, as an Assistant Professor, and then an Associate Professor from 2010. He is currently a Distinguished Professor with the Department of Communication Science and Engineering, Fudan University. His research interests include stochastic network optimization, energy-efficient communications, cross-layer design, and signal processing for communications. He served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and as an Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Georgios B. Giannakis (F'97) received the Diploma degree in electrical engineering from the National Technical University of Athens, Greece, 1981, and the M.Sc. degree in electrical engineering, the M.Sc. degree in mathematics, and the Ph.D. degree in electrical engineering from the University of Southern California in 1983, 1986, and 1986, respectively, where he was from 1982 to 1986. He was with the University of Virginia from 1987 to 1998, and since 1999, he has been a Professor with the University of Minnesota, where he holds an Endowed Chair in wireless telecommunications, the University of Minnesota McKnight Presidential Chair in ECE, and serves as the Director of the Digital Technology Center.

His general interests span the areas of communications, networking, and statistical signal processing—subjects on which he has published over 400 journal papers, 700 conference papers, 25 book chapters, two edited books, and two research monographs with an H-index of 128. Current research focuses on learning from big data, wireless cognitive radios, and network science with applications to social, brain, and power networks with renewables. He is the (co-)inventor of 30 patents issued. He was a (co-)recipient of nine best paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications, the Technical Achievement Awards from the SP Society in 2000 and from EURASIP in 2005, the Young Faculty Teaching Award, the G. W. Taylor Award for Distinguished Research from the University of Minnesota, and the IEEE Fourier Technical Field Award in 2015. He is a fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.