

# Two-Scale Stochastic Control for Integrated Multipoint Communication Systems with Renewables

Xin Wang, *Senior Member, IEEE*, Xiaojing Chen, *Student Member, IEEE*, Tianyi Chen, *Student Member, IEEE*, Longbo Huang, *Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

**Abstract**—Increasing threats of global warming and climate changes call for an energy-efficient and sustainable design of future wireless communication systems. To this end, a novel two-scale stochastic control framework is put forth for smart-grid powered coordinated multi-point (CoMP) systems. Taking into account renewable energy sources (RES), dynamic pricing, two-way energy trading facilities and imperfect energy storage devices, the energy management task is formulated as an infinite-horizon optimization problem minimizing the time-averaged energy transaction cost, subject to the users’ quality of service (QoS) requirements. Leveraging the Lyapunov optimization approach as well as the stochastic subgradient method, a two-scale online control (TS-OC) approach is developed for the resultant smart-grid powered CoMP systems. Using only historical data, the proposed TS-OC makes online control decisions at two timescales, and features a provably feasible and asymptotically near-optimal solution. Numerical tests further corroborate the theoretical analysis, and demonstrate the merits of the proposed approach.

**Index Terms**—Two-scale control, ahead-of-time market, real-time market, battery degeneration, CoMP systems, smart grids, Lyapunov optimization.

## I. INTRODUCTION

Interference is a major obstacle in wireless communication systems due to their broadcast nature, and becomes more severe in next-generation spectrum- and energy-constrained cellular networks with smaller cells and more flexible frequency reuse [1]. With ever increasing demand for energy-efficient transmissions, coordinated multi-point processing (CoMP) has been proposed as a promising paradigm for efficient inter-cell interference management in heterogeneous networks (HetNets) [2]. In CoMP systems, base stations (BSs) are partitioned into clusters, where BSs per cluster perform coordinated beamforming to serve the users [3]–[5]. As the number of BSs in

HetNets increases, their electricity consumption constitutes a major part of the operational expenditure of cellular networks, and contributes a considerable portion of the global *carbon footprint* [6]. Fortunately, contemporary advances in smart grids offer great opportunity of achieving both energy-efficient and environment friendly communication solutions by using new grid infrastructure features; e.g., high penetration of renewable energy sources (RES), two-way energy trading, and dynamic pricing based demand-side management (DSM) [9]–[11]. In this context, energy-efficient “green” communication solutions are advocated for their economic and ecological merits [3]–[6]. Driven by the need of sustainable “green communications,” manufacturers and network operators such as Ericsson, Huawei, Vodafone and China Mobile have started developing the “green” BSs that can be jointly supplied by the persistent power sources from the main electric grid as well as the harvested renewable energy (e.g., solar and wind) [7], [8]. It is expected that renewable powered BSs will be widely used in the future cellular systems.

A few recent works have considered the smart-grid powered CoMP transmissions [12]–[15]. Assuming that the energy harvested from RES is accurately available *a priori* through e.g., forecasting, [12] and [13] considered the energy-efficient resource allocation for RES-powered CoMP downlinks. Building on realistic models, our recent work dealt with robust energy management and transmit-beamforming designs that minimize the worst-case energy transaction cost for the CoMP downlink with RES and DSM [14]. Leveraging novel stochastic optimization tools [16]–[18], we further developed an efficient approach to obtain a feasible and asymptotically optimal online control scheme for smart-grid powered CoMP systems, without knowing the distributions of the random variables involved [15].

A salient assumption in [12]–[15] is that all involved resource allocation tasks are performed in a single time scale. However, RES and wireless channel dynamics typically evolve over different time scales in practice. Development of two-scale control schemes is then well motivated for CoMP systems with RES. In relevant contexts, a few stochastic optimization based two-scale control schemes were recently proposed and analyzed in [19]–[22]. Extending the traditional Lyapunov optimization approach [16]–[18], [19] introduced a two-scale control algorithm that makes distributed routing and server management decisions to reduce power cost for large-scale data centers. Based on a similar approach, [20] developed

Work in this paper was supported by the China Recruitment Program of Global Young Experts, the Program for New Century Excellent Talents in University, the Innovation Program of Shanghai Municipal Education Commission; US NSF grants ECCS-1509005, ECCS-1508993, CCF-1423316, CCF-1442686, and ECCS-1202135.

X. Wang and X. Chen are with the Key Laboratory for Information Science of Electromagnetic Waves (MoE), Department of Communication Science and Engineering, Fudan University, 220 Han Dan Road, Shanghai, China. Emails: {xwang11, 13210720095}@fudan.edu.cn.

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. Emails: {chen3827, georgios}@umn.edu.

L. Huang is with the Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, China. Email: longbohuang@tsinghua.edu.cn.

X. Wang, X. Chen, and T. Chen contributed equally to this work.

the so-called MultiGreen algorithm for data centers, which allows cloud service providers to make energy transactions at two time scales for minimum operational cost. In the context of wireless communication, the algorithm in [21] performed joint precoder assignment and user association and channel resource scheduling for HetNets with non-ideal backhaul; while [22] studied the two-timescale approach for network selection and subchannel allocation in the integrated cellular and Wi-Fi networks with an emphasis on using predictive future information. Note that neither [21] nor [22] considers the diversity of energy prices in fast/slow-timescale energy markets, and the energy leakage effects in the energy management task.

In the present paper, we develop a two-scale online control (TS-OC) approach for smart-grid powered CoMP systems considering RES, dynamic pricing, two-way energy trading facilities and imperfect energy storage devices. Suppose that the RES harvesting occurs at the BSs over a slow timescale relative to the coherence time of wireless channels. The proposed scheme performs an ahead-of-time (e.g., 15-minute ahead, or, hour-ahead) energy planning upon RES arrivals, while deciding real-time energy balancing and transmit-beamforming schedules per channel coherence time slot. Specifically, the TS-OC determines the amount of energy to trade (purchase or sell) with the ahead-of-time wholesale market based on RES generation, as the basic energy supply for all the time slots within a RES harvesting interval. On the other hand, it decides the amount of energy to trade with the real-time market, energy charging to (or discharging from) the batteries, as well as the coordinated transmit-beamformers to guarantee the users' quality of service (QoS) per time slot. Generalizing the Lyapunov optimization techniques in [19]–[23], we propose a synergetic framework to design and analyze such a two-scale dynamic management scheme to minimize the long-term time-averaged energy transaction cost of the CoMP transmissions, without knowing the distributions of the random channel, RES, and energy price processes. The main contributions of our work are summarized as follows.

- Leveraging the ahead-of-time and real-time electricity markets along with generalized system models and formulations in [14], [15], a novel two-scale optimization framework is developed to facilitate the dynamic resource management for smart-grid powered CoMP systems with RES and channel dynamics at different time scales.
- While methods in [15], [19] and [20] cannot capture the practical battery degeneration (energy leakage), we integrate the modified Lyapunov optimization technique [23] into the two-scale stochastic optimization approach [19], [20] to leverage the diversity of energy prices along with the energy leakage effects on the dynamic energy management task.
- Using only past channel and energy-price realizations, we propose a novel stochastic subgradient approach to solve the ahead-of-time energy planning (sub-)problem, which is suitable for a general family of continuous distributions, and avoids constructing the tedious histogram for high-dimensional vector compared with the empirical-pdf based approach in [19], [20].

- Rigorous analysis is presented to justify the feasibility and quantify the optimality gap for the proposed two-scale online control algorithm.

The rest of the paper is organized as follows. The system models are described in Section II. The proposed dynamic resource management scheme is developed in Section III. Analysis of the guaranteed performance is the subject of Section IV. Numerical tests are provided in Section V, followed by concluding remarks in Section VI.

*Notation.* Boldface lower (upper) case letters represent vectors (matrices);  $\mathbb{C}^N$  and  $\mathbb{R}^{N \times M}$  stand for spaces of  $N \times 1$  complex vectors and  $N \times M$  real matrices, respectively;  $(\cdot)'$  denotes transpose, and  $(\cdot)^H$  conjugate transpose;  $\text{diag}(a_1, \dots, a_M)$  denotes a diagonal matrix with diagonal elements  $a_1, \dots, a_M$ ;  $|\cdot|$  the magnitude of a complex scalar; and  $\mathbb{E}$  denotes expectation.

## II. SYSTEM MODELS

Consider a cluster-based CoMP downlink setup, where a set  $\mathcal{I} := \{1, \dots, I\}$  of distributed BSs (e.g., macro/micro/pico BSs) is selected to serve a set  $\mathcal{K} := \{1, \dots, K\}$  of mobile users, as in e.g., [14], [15]. Each BS is equipped with  $M \geq 1$  transmit antennas, whereas each user has a single receive antenna. Suppose that through the smart-grid infrastructure conventional power generation is available, but each BS can also harvest RES (through e.g., solar panels and/or wind turbines), and it has an energy storage device (i.e., battery) to save the harvested energy. Relying on a two-way energy trading facility, the BS can also buy energy from or sell energy to the main grid at dynamically changing market prices. For the CoMP cluster, there is a low-latency backhaul network connecting the set of BSs to a central controller [4], which coordinates energy trading as well as cooperative communication. This central entity can collect both communication data (transmit messages, channel state information) from each BS through the cellular backhaul links, as well as the energy information (energy purchase/selling prices, energy queue sizes) via smart meters installed at BSs, and the grid-deployed communication/control links connecting them.<sup>1</sup>

As the RES and wireless channel dynamics emerge typically at different time scales in practice, we propose a two-scale control mechanism. As shown in Fig. 1, time is divided in slots of length smaller than the coherence time of the wireless channels. For convenience, the slot duration is normalized to unity; thus, the terms “energy” and “power” can be used interchangeably. On the other hand, we define the (virtual) “coarse-grained” time intervals in accordance with the slow RES harvesting scale, with each coarse-grained interval consisting of  $T$  time slots.

<sup>1</sup>Note that we assume perfect channel state information per slot in the present paper. However, the proposed formulation can readily take into account the channel estimation errors to accommodate a robust beamforming design; see the approaches in our recent works [14], [15]. On the other hand, the proposed approach can be also generalized to address the impact of imperfect energy queue information based on the Lyapunov optimization framework in [20]. In fact, it can be shown that such imperfect information would not substantially affect the effectiveness of the proposed scheme. We do not consider these generalizations in order to prevent distraction from our main focus on development of the two-scale stochastic online control approach for integrated CoMP systems with RES.

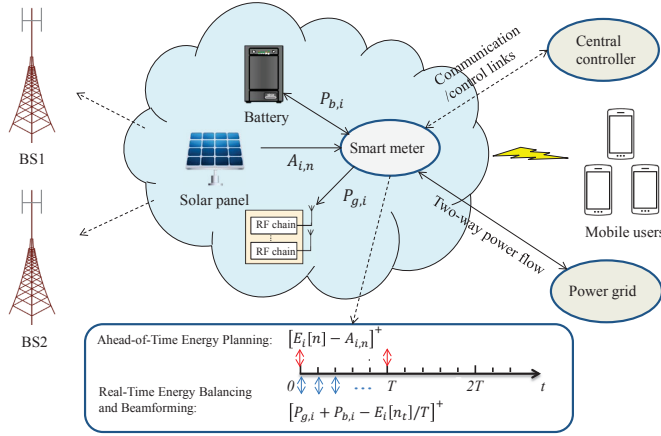


Fig. 1. A smart grid powered CoMP system. Two BSs with local renewable energy harvesting and storage devices implement two-way energy trading with the main grid.

### A. Ahead-of-Time Energy Planning

At the beginning of each “coarse-grained” interval, namely at time  $t = nT$ ,  $n = 1, 2, \dots$ , let  $A_{i,n}$  denote the RES amount collected per BS  $i \in \mathcal{I}$ , and  $\mathbf{A}_n := [A_{1,n}, \dots, A_{I,n}]'$ . With  $\mathbf{A}_n$  available, an energy planner at the central unity decides the energy amounts  $E_i[n]$ ,  $\forall i$ , to be used in the next  $T$  slots per BS  $i$ . With a two-way energy trading facility, the BSs then either purchase energy from the main grid according to their shortage, or sell their surplus energy to the grid at a fair price in order to reduce operational costs. Specifically, following the decision, BS  $i$  contributes its RES amount  $A_{i,n}$  to the main grid and requests the grid to supply an average energy amount of  $E_i[n]/T$  per slot  $t = nT, \dots, (n+1)T - 1$ .

RES is assumed harvested for free after deployment. Given the requested energy  $E_i[n]$  and the harvested energy  $A_{i,n}$ , the shortage energy that is purchased from the grid for BS  $i$  is clearly  $[E_i[n] - A_{i,n}]^+$ ; or, the surplus energy that is sold to the grid is  $[A_{i,n} - E_i[n]]^+$ , where  $[a]^+ := \max\{a, 0\}$ . Depending on the difference  $(E_i[n] - A_{i,n})$ , the BS  $i$  either buys electricity from the grid with the ahead-of-time (i.e., long-term) price  $\alpha_n^{\text{lt}}$ , or sells electricity to the grid with price  $\beta_n^{\text{lt}}$  for profit (the latter leads to a negative cost). Notwithstanding, we shall always set  $\alpha_n^{\text{lt}} > \beta_n^{\text{lt}}$  to avoid meaningless buy-and-sell activities of the BSs for profit. The transaction cost with BS  $i$  for such an energy planning is therefore given by

$$G^{\text{lt}}(E_i[n]) := \alpha_n^{\text{lt}}[E_i[n] - A_{i,n}]^+ - \beta_n^{\text{lt}}[A_{i,n} - E_i[n]]^+. \quad (1)$$

For the sake of a concise presentation, we use a random vector collecting all the random variables evolving at this slow timescale; i.e.,  $\xi_n^{\text{lt}} := \{\alpha_n^{\text{lt}}, \beta_n^{\text{lt}}, \mathbf{A}_n, \forall n\}$ .

### B. CoMP Downlink Transmissions

Per slot  $t$ , let  $\mathbf{h}_{ik,t} \in \mathbb{C}^M$  denote the vector channel from BS  $i$  to user  $k$ ,  $\forall i \in \mathcal{I}, \forall k \in \mathcal{K}$ ; let  $\mathbf{h}_{k,t} := [\mathbf{h}_{1k,t}, \dots, \mathbf{h}_{Ik,t}]'$  collect the channel vectors from all BSs to user  $k$ , and  $\mathbf{H}_t := [\mathbf{h}_{1,t}, \dots, \mathbf{h}_{K,t}]$ . With linear transmit beamforming performed across BSs, the vector signal transmitted to user  $k$  is:  $\mathbf{q}_k(t) = \mathbf{w}_k(t)s_k(t)$ ,  $\forall k$ , where  $s_k(t)$  denotes the information-bearing scalar symbol with unit-energy, and  $\mathbf{w}_k(t) \in \mathbb{C}^{MI}$  denotes

the beamforming vector across the BSs serving user  $k$ . The received vector at slot  $t$  for user  $k$  is therefore

$$\mathbf{y}_k(t) = \mathbf{h}_{k,t}^H \mathbf{q}_k(t) + \sum_{l \neq k} \mathbf{h}_{k,t}^H \mathbf{q}_l(t) + n_k(t) \quad (2)$$

where  $\mathbf{h}_{k,t}^H \mathbf{q}_k(t)$  is the desired signal of user  $k$ ,  $\sum_{l \neq k} \mathbf{h}_{k,t}^H \mathbf{q}_l(t)$  is the inter-user interference from the same cluster, and  $n_k(t)$  denotes additive noise, which may also include the downlink interference from other BSs outside user  $k$ 's cluster. It is assumed that  $n_k(t)$  is a circularly symmetric complex Gaussian (CSCG) random variable with zero mean and variance  $\sigma_k^2$ .

The signal-to-interference-plus-noise ratio (SINR) at user  $k$  can be expressed as

$$\text{SINR}_k(\{\mathbf{w}_k(t)\}) = \frac{|\mathbf{h}_{k,t}^H \mathbf{w}_k(t)|^2}{\sum_{l \neq k} (|\mathbf{h}_{k,t}^H \mathbf{w}_l(t)|^2) + \sigma_k^2}. \quad (3)$$

The transmit power at each BS  $i$  clearly is given by

$$P_{x,i}(t) = \sum_{k \in \mathcal{K}} \mathbf{w}_k^H(t) \mathbf{B}_i \mathbf{w}_k(t) \quad (4)$$

where the matrix

$$\mathbf{B}_i := \text{diag} \left( \underbrace{0, \dots, 0}_{(i-1)M}, \underbrace{1, \dots, 1}_M, \underbrace{0, \dots, 0}_{(I-i)M} \right) \in \mathbb{R}^{MI \times MI}$$

selects the corresponding rows out of  $\{\mathbf{w}_k(t)\}_{k \in \mathcal{K}}$  to form the  $i$ -th BS's transmit-beamforming vector of size  $M \times 1$ .

To guarantee QoS per slot user  $k$ , it is required that the central controller selects a set of  $\{\mathbf{w}_k(t)\}$  satisfying [cf. (3)]

$$\text{SINR}_k(\{\mathbf{w}_k(t)\}) \geq \gamma_k, \quad \forall k \quad (5)$$

where  $\gamma_k$  denotes the target SINR value per user  $k$ .

### C. Real-Time Energy Balancing

For the  $i$ -th BS, the total energy consumption  $P_{g,i}(t)$  per slot  $t$  includes the transmission-related power  $P_{x,i}(t)$ , and the rest that is due to other components such as air conditioning, data processor, and circuits, which can be generally modeled as a constant power,  $P_c > 0$  [13]. We further suppose that  $P_{g,i}(t)$  is bounded by  $P_g^{\text{max}}$ . Namely,

$$P_{g,i}(t) = P_c + \sum_{k \in \mathcal{K}} \mathbf{w}_k^H(t) \mathbf{B}_i \mathbf{w}_k(t) \leq P_g^{\text{max}}, \quad \forall i. \quad (6)$$

Per slot  $t$ , the energy supply available from the ahead-of-time planning may not exactly meet the actual demand at BS  $i$ . Hence, the BS  $i$  is also allowed to perform real-time energy trading with the main grid to balance its supply with demand. Let  $P_i(t)$  denote the real-time energy amount that is purchased from ( $P_i(t) > 0$ ) or sold to ( $P_i(t) < 0$ ) the grid by BS  $i$ . Let  $\alpha_t^{\text{rt}}$  and  $\beta_t^{\text{rt}}$  ( $\alpha_t^{\text{rt}} > \beta_t^{\text{rt}}$ ) denote the real-time energy purchase and selling prices, respectively. Then the real-time energy transaction cost for BS  $i$  is

$$G^{\text{rt}}(P_i(t)) := \alpha_t^{\text{rt}}[P_i(t)]^+ - \beta_t^{\text{rt}}[-P_i(t)]^+. \quad (7)$$

Fig. 2 depicts the day-ahead and real-time energy prices in the Pennsylvania-Jersey-Maryland (PJM) wholesale market [24]. In practice, the average purchase price in the real-time market

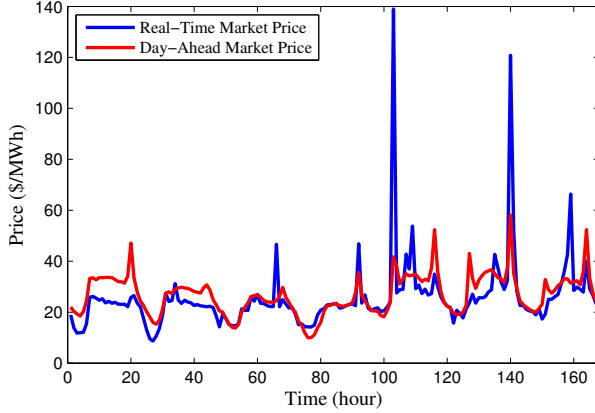


Fig. 2. Hourly price trend for day-ahead and real-time electricity markets during Oct. 01-07, 2015 [24].

tends to be no lower than that in the day-ahead market; that is,  $\mathbb{E}\{\alpha_t^{\text{rt}}\} \geq \mathbb{E}\{\alpha_n^{\text{lt}}\}$ ; similarly, we have  $\mathbb{E}\{\beta_t^{\text{rt}}\} \leq \mathbb{E}\{\beta_n^{\text{lt}}\}$ . Again, we use a random vector  $\xi_t^{\text{rt}} := \{\alpha_t^{\text{rt}}, \beta_t^{\text{rt}}, \mathbf{H}_t, \forall t\}$  collecting all the random variables evolving at the fast timescale.

#### D. Energy Storage with Degeneration

As energy consumption will become a major concern of the future large-scale cellular networks, uninterrupted power supply type storage units can be installed at the BSs to prevent power outages, and provide opportunities to optimize the BSs' electricity bills. Different from the ideal battery models in [12]–[15], [20], we consider here a practical battery with degeneration (i.e., energy leakage over time even in the absence of discharging) as in [23].

For the battery of the  $i$ -th BS, let  $C_i(0)$  denote the initial amount of stored energy, and  $C_i(t)$  its state of charge (SoC) at the beginning of time slot  $t$ . The battery is assumed to have a finite capacity  $C^{\max}$ . Furthermore, for reliability purposes, it might be required to ensure that a minimum energy level  $C^{\min}$  is maintained at all times. Let  $P_{b,i}(t)$  denote the energy delivered to or drawn from the battery at slot  $t$ , which amounts to either charging ( $P_{b,i}(t) > 0$ ) or discharging ( $P_{b,i}(t) < 0$ ). The stored energy then obeys the dynamic equation

$$C_i(t+1) = \eta C_i(t) + P_{b,i}(t), \quad C^{\min} \leq C_i(t) \leq C^{\max}, \quad \forall i \quad (8)$$

where  $\eta \in (0, 1]$  denotes the storage efficiency (e.g.,  $\eta = 0.9$  means that 10% of the stored energy will be “leaked” over a slot, even in the absence of discharging).

The amount of power (dis)charged is also assumed bounded by

$$P_b^{\min} \leq P_{b,i}(t) \leq P_b^{\max}, \quad \forall i \quad (9)$$

where  $P_b^{\min} < 0$  and  $P_b^{\max} > 0$  are introduced by physical constraints.

With  $n_t := \lfloor \frac{t}{T} \rfloor$  and consideration of  $P_{b,i}(t)$ , we have the following demand-and-supply balance equation per slot  $t$ :

$$P_c + \sum_{k \in \mathcal{K}} \mathbf{w}_k^H(t) \mathbf{B}_i \mathbf{w}_k(t) + P_{b,i}(t) = \frac{E_i[n_t]}{T} + P_i(t), \quad \forall i. \quad (10)$$

### III. DYNAMIC RESOURCE MANAGEMENT SCHEME

Note that the harvested RES amounts  $\{\mathbf{A}_n, \forall n\}$ , the ahead-of-time prices  $\{\alpha_n^{\text{lt}}, \beta_n^{\text{lt}}, \forall n\}$ , the real-time prices  $\{\alpha_t^{\text{rt}}, \beta_t^{\text{rt}}, \forall t\}$ , and the wireless channel matrices  $\{\mathbf{H}_t, \forall t\}$  are all random. The smart-grid powered CoMP downlink to be controlled is a stochastic system. The goal is to design an online resource management scheme that chooses the ahead-of-time energy-trading amounts  $\{E_i[n], \forall i\}$  at every  $t = nT$ , as well as the real-time energy-trading amounts  $\{P_i(t), \forall i\}$ , battery (dis)charging amounts  $\{P_{b,i}(t), \forall i\}$ , and the CoMP beamforming vectors  $\{\mathbf{w}_k(t), \forall k\}$  per slot  $t$ , so as to minimize the expected total energy transaction cost, without knowing the distributions of the aforementioned random processes.

According to (1) and (7), define the energy transaction cost for BS  $i$  per slot  $t$  as:

$$\Phi_i(t) := \frac{1}{T} G^{\text{lt}}(E_i[n_t]) + G^{\text{rt}}(P_i(t)). \quad (11)$$

Let  $\mathcal{X} := \{E_i[n], \forall i, n; P_i(t), P_{b,i}(t), C_i(t), \forall i, t; \mathbf{w}_k(t), \forall k, t\}$ . The problem of interest is to find

$$\begin{aligned} \Phi^{\text{opt}} := \min_{\mathcal{X}} \lim_{N \rightarrow \infty} \frac{1}{NT} \sum_{t=0}^{NT-1} \sum_{i \in \mathcal{I}} \mathbb{E}\{\Phi_i(t)\} \\ \text{subject to } (5), (6), (8), (9), (10), \quad \forall t \end{aligned} \quad (12)$$

where the expectations of  $\Phi_i(t)$  are taken over all sources of randomness. Note that here the constraints (5), (6), (8), (9), and (10) are implicitly required to hold for every realization of the underlying random states  $\xi_t^{\text{rt}}$  and  $\xi_n^{\text{lt}}$ .

#### A. Two-Scale Online Control Algorithm

Problem (12) is a stochastic optimization task. We next generalize and integrate the Lyapunov optimization techniques in [19]–[23] to develop a TS-OC algorithm, which will be proven feasible, and asymptotically near-optimal for (12). To start, we assume the following two relatively mild conditions for the system parameters:

$$P_b^{\max} \geq (1 - \eta) C^{\min} \quad (13)$$

$$C^{\max} - C^{\min} \geq \frac{1 - \eta^T}{1 - \eta} (P_b^{\max} - P_b^{\min}). \quad (14)$$

Condition (13) simply implies that the energy leakage of the battery can be compensated by the charging. Condition (14) requires that the allowable SoC range is large enough to accommodate the largest possible charging/discharging over  $T$  time slots of each coarse-grained interval. This then makes the system “controllable” by our two-scale mechanism.

Our algorithm depends on two parameters, namely a “queue perturbation” parameter  $\Gamma$ , and a weight parameter  $V$ . Define  $\bar{\alpha} := \max\{\alpha_t^{\text{rt}}, \forall t\}$  and  $\underline{\beta} := \min\{\beta_t^{\text{rt}}, \forall t\}$ . Derived from the feasibility requirement of the proposed algorithm (see the proof of Proposition 1 in the sequel), any pair  $(\Gamma, V)$  that satisfies the following conditions can be used:

$$\Gamma^{\min} \leq \Gamma \leq \Gamma^{\max}, \quad 0 < V \leq V^{\max} \quad (15)$$



where

$$\Gamma^{\min} := \max_{\tau=1, \dots, T} \left\{ \frac{1}{\eta^\tau} \left( \frac{1-\eta^\tau}{1-\eta} P_b^{\max} - C^{\max} \right) - V \underline{\beta} \right\} \quad (16)$$

$$\Gamma^{\max} := \min_{\tau=1, \dots, T} \left\{ \frac{1}{\eta^\tau} \left( \frac{1-\eta^\tau}{1-\eta} P_b^{\min} - C^{\min} \right) - V \bar{\alpha} \right\} \quad (17)$$

$$V^{\max} := \min_{\tau=1, \dots, T} \left\{ \frac{C^{\max} - C^{\min} - \frac{1-\eta^\tau}{1-\eta} (P_b^{\max} - P_b^{\min})}{\eta^\tau (\bar{\alpha} - \underline{\beta})} \right\}. \quad (18)$$

Note that the interval for  $V$  in (15) is well defined under condition (14), and the interval for  $\Gamma$  is valid when  $V \leq V^{\max}$ .

We now present the proposed TS-OC algorithm:

- **Initialization:** Select  $\Gamma$  and  $V$ , and introduce a virtual queue  $Q_i(0) := C_i(0) + \Gamma, \forall i$ .
- **Ahead-of-time energy planning:** Per interval  $\tau = nT$ , observe a realization  $\xi_n^{\text{lt}}$ , and determine the energy amounts  $\{E_i^*[n], \forall i\}$  by solving

$$\begin{aligned} \min_{\{E_i^*[n]\}} \sum_{i \in \mathcal{I}} \left\{ V \left[ G^{\text{lt}}(E_i[n]) + \sum_{t=\tau}^{\tau+T-1} \mathbb{E}\{G^{\text{rt}}(P_i(t))\} \right] \right. \\ \left. + \sum_{t=\tau}^{\tau+T-1} Q_i(t) \mathbb{E}\{P_{b,i}(t)\} \right\} \\ \text{s. t. (5), (6), (9), (10), } \quad \forall t = \tau, \dots, \tau + T - 1 \quad (19) \end{aligned}$$

where expectations are taken over  $\xi_i^{\text{rt}}$ . Then the BSs trade energy with the main grid based on  $\{E_i^*[n], \forall i\}$ , and request the grid to supply an average amount  $E_i^*[n]/T$  per slot  $t = \tau, \dots, \tau + T - 1$ .

- **Energy balancing and beamforming schedule:** At every slot  $t \in [nT, (n+1)T - 1]$ , observe a realization  $\xi_t^{\text{rt}}$ , and decide  $\{P_i^*(t), P_{b,i}^*(t), \forall i; \mathbf{w}_k^*(t), \forall k\}$  by solving the following problem given  $E_i[n] = E_i^*[n]$

$$\begin{aligned} \min_{\{P_i^*(t), P_{b,i}^*(t), \mathbf{w}_k^*(t)\}} \sum_{i \in \mathcal{I}} \{V G^{\text{rt}}(P_i(t)) + Q_i(nT) P_{b,i}(t)\} \\ \text{s. t. (5), (6), (9), (10).} \quad (20) \end{aligned}$$

The BSs perform real-time energy trading with the main grid based on  $\{P_i^*(t), \forall i\}$ , and coordinated beamforming based on  $\{\mathbf{w}_k^*(t), \forall k\}$ .

- **Queue updates:** Per slot  $t$ , charge (or discharge) the battery based on  $\{P_{b,i}^*(t)\}$ , so that the stored energy  $C_i(t+1) = \eta C_i(t) + P_{b,i}^*(t), \forall i$ ; and update the virtual queues  $Q_i(t) := C_i(t) + \Gamma, \forall i$ .

**Remark 1:** Note that we use queue sizes  $Q_i(\tau)$  instead of  $Q_i(t)$  in problems (19) and (20). This idea in fact inherits from the two-timescale Lyapunov optimization frameworks in [19], [20]. Recall that the main design principle in original Lyapunov optimization is to choose control actions that minimize  $\sum_t \sum_i \mathbb{E}[V \Phi_i(t) + Q_i(t) P_{b,i}(t)]$ . For the ahead-of-time energy planning, this requires a-priori knowledge of the future queue backlogs  $Q_i(t)$  over slots  $[\tau+1, \dots, \tau+T-1]$  at time  $\tau = nT$ . It is certainly not reasonable to assume that this information is always available. For this reason, we simply approximate future queue backlog values as the current value at  $\tau = nT$ , i.e.,  $Q_i(t) \approx Q_i(\tau), \forall t = \tau+1, \dots, \tau+T-1$ , in ahead-of-time energy planning problem (19). To ensure that the real-time energy balancing and beamforming schedule

solves the same problem as the ahead-of-time energy planning, we also use  $Q_i(nT)$  in problem (20) although the real-time battery state of charge  $Q_i(t)$  is available at slot  $t$ . Rigorous analysis shows that the performance penalty incurred by such an approximation would not affect the asymptotic optimality of the proposed stochastic control scheme. On the other hand, using  $Q_i(t)$  in real-time energy balancing can be also suggested in practice. While our feasibility analysis affords such modification, the derivation of the optimality gap needs additional efforts.

Next, we develop efficient solvers of (19) and (20) to obtain the TS-OC algorithm.

### B. Real-Time Energy Balancing and Beamforming

It is easy to argue that the objective (20) is convex. Indeed, with  $\alpha_t^{\text{rt}} > \beta_t^{\text{rt}}$ , the transaction cost with  $P_i(t)$  can be alternatively written as

$$G^{\text{rt}}(P_i(t)) = \max\{\alpha_t^{\text{rt}} P_i(t), \beta_t^{\text{rt}} P_i(t)\} \quad (21)$$

which is clearly convex [25]; and so is the objective in (20).

The SINR constraints in (5) can be actually rewritten into a convex form. Observe that an arbitrary phase rotation can be added to the beamforming vectors  $\mathbf{w}_k(t)$  without affecting the SINRs. Hence, we can choose a phase so that  $\mathbf{h}_{k,t}^H \mathbf{w}_k(t)$  is real and nonnegative. Then by proper rearrangement, the SINR constraints become convex second-order cone (SOC) constraints [27]; that is,

$$\begin{aligned} \sqrt{\sum_{l \neq k} |\mathbf{h}_{k,t}^H \mathbf{w}_l(t)|^2 + \sigma_k^2} \leq \frac{1}{\sqrt{\gamma_k}} \text{Re}\{\mathbf{h}_{k,t}^H \mathbf{w}_k(t)\}, \\ \text{Im}\{\mathbf{h}_{k,t}^H \mathbf{w}_k(t)\} = 0, \quad \forall k. \end{aligned}$$

We can then rewrite the problem (20) as

$$\begin{aligned} \min_{i \in \mathcal{I}} \sum_{i \in \mathcal{I}} \left\{ V G^{\text{rt}}(P_c + \sum_{k \in \mathcal{K}} \mathbf{w}_k^H(t) \mathbf{B}_i \mathbf{w}_k(t) + P_{b,i}(t) - \frac{E_i^*[n_t]}{T}) \right. \\ \left. + Q_i(n_t T) P_{b,i}(t) \right\} \\ \text{s. t. } \sqrt{\sum_{l \neq k} |\mathbf{h}_{k,t}^H \mathbf{w}_l(t)|^2 + \sigma_k^2} \leq \frac{1}{\sqrt{\gamma_k}} \text{Re}\{\mathbf{h}_{k,t}^H \mathbf{w}_k(t)\}, \\ \text{Im}\{\mathbf{h}_{k,t}^H \mathbf{w}_k(t)\} = 0, \quad \forall k \\ P_b^{\min} \leq P_{b,i}(t) \leq P_b^{\max}, \quad \forall i \\ P_c + \sum_{k \in \mathcal{K}} \mathbf{w}_k^H(t) \mathbf{B}_i \mathbf{w}_k(t) \leq P_g^{\max}, \quad \forall i. \quad (22) \end{aligned}$$

As  $G^{\text{rt}}(\cdot)$  is convex and increasing, it is easy to see that  $G^{\text{rt}}(P_c + \sum_k \mathbf{w}_k^H(t) \mathbf{B}_i \mathbf{w}_k(t) + P_{b,i}(t) - E_i^*[n_t]/T)$  is jointly convex in  $(P_{b,i}(t), \{\mathbf{w}_k(t)\})$  [25, Sec. 3.2.4]. It then readily follows that (22) is a convex optimization problem, which can be solved via off-the-shelf solvers.

### C. Ahead-of-Time Energy Planning

To solve (19), the probability distribution function (pdf) of the random state  $\xi_t^{\text{rt}}$  must be known across slots  $t = nT, \dots, (n+1)T - 1$ . However, this pdf is seldom available in practice. Suppose that  $\xi_t^{\text{rt}}$  is independent and identically

distributed (i.i.d.) over time slots, and takes values from a finite state space. It was proposed in [19] to obtain an empirical pdf of  $\xi_t^{\text{rt}}$  from past realizations over a large window comprising  $L$  intervals. This estimate becomes accurate as  $L$  grows sufficiently large; then it can be used to evaluate the expectations in (19). Based on such an empirical pdf, an approximate solution for (19) could be obtained.

Different from [19], here we propose a stochastic gradient approach to solve (19). Suppose that  $\xi_t^{\text{rt}}$  is i.i.d. across time slots (but not necessarily with a finite support). For stationary  $\xi_t^{\text{rt}}$ , we can remove the index  $t$  from all optimization variables, and rewrite (19) as (with short-hand notation  $Q_i[n] := Q_i(nT)$ )

$$\min \sum_{i \in \mathcal{I}} \left\{ V G^{\text{lt}}(E_i[n]) + T \mathbb{E} [V G^{\text{rt}}(P_i(\xi_t^{\text{rt}})) + Q_i[n] P_{b,i}(\xi_t^{\text{rt}})] \right\}$$

$$\text{s. t. } \sqrt{\sum_{l \neq k} |h_k^H \mathbf{w}_l(\xi_t^{\text{rt}})|^2 + \sigma_k^2} \leq \frac{1}{\sqrt{\gamma_k}} \text{Re}\{\mathbf{h}_k^H \mathbf{w}_k(\xi_t^{\text{rt}})\},$$

$$\text{Im}\{\mathbf{h}_k^H \mathbf{w}_k(\xi_t^{\text{rt}})\} = 0, \forall k, \xi_t^{\text{rt}} \quad (23a)$$

$$P_b^{\min} \leq P_{b,i}(\xi_t^{\text{rt}}) \leq P_b^{\max}, \forall i, \xi_t^{\text{rt}} \quad (23b)$$

$$P_c + \sum_{k \in \mathcal{K}} \mathbf{w}_k^H(\xi_t^{\text{rt}}) \mathbf{B}_i \mathbf{w}_k(\xi_t^{\text{rt}}) \leq P_g^{\max}, \forall i, \xi_t^{\text{rt}} \quad (23c)$$

$$P_c + \sum_{k \in \mathcal{K}} \mathbf{w}_k^H(\xi_t^{\text{rt}}) \mathbf{B}_i \mathbf{w}_k(\xi_t^{\text{rt}}) + P_{b,i}(\xi_t^{\text{rt}}) = \frac{E_i[n]}{T} + P_i(\xi_t^{\text{rt}}), \forall i, \xi_t^{\text{rt}}. \quad (23d)$$

Note that this form explicitly indicates the dependence of the decision variables  $\{P_i, P_{b,i}, \mathbf{w}_k\}$  on the realization of  $\xi_t^{\text{rt}}$ .

Since the energy planning problem (19) only determines the optimal ahead-of-time energy purchase  $E_i^*[n]$ , we can then eliminate the variable  $P_i$  and write (23) as an unconstrained optimization problem with respect to the variable  $E_i^*[n]$ , namely

$$\min_{\{E_i[n]\}} \sum_{i \in \mathcal{I}} \left[ V G^{\text{lt}}(E_i[n]) + T \bar{G}^{\text{rt}}(\{E_i[n]\}) \right] \quad (24)$$

where we define

$$\begin{aligned} \bar{G}^{\text{rt}}(\{E_i[n]\}) := & \min_{\{P_i, P_{b,i}, \mathbf{w}_k\}} \sum_{i \in \mathcal{I}} \mathbb{E} \left\{ V \Psi^{\text{rt}}(E_i[n], P_{b,i}(\xi_t^{\text{rt}}), \{\mathbf{w}_k(\xi_t^{\text{rt}})\}) \right. \\ & \left. + Q_i[n] P_{b,i}(\xi_t^{\text{rt}}) \right\} \text{ s. t. (23a), (23b), (23c)} \quad (25) \end{aligned}$$

with the compact notation  $\Psi^{\text{rt}}(E_i, P_{b,i}, \{\mathbf{w}_k\}) := G^{\text{rt}}(P_c + \sum_{k \in \mathcal{K}} \mathbf{w}_k^H \mathbf{B}_i \mathbf{w}_k + P_{b,i} - \frac{E_i}{T})$ . Since  $\mathbb{E}[V \Psi^{\text{rt}}(E_i[n], P_{b,i}(\xi_t^{\text{rt}}), \{\mathbf{w}_k(\xi_t^{\text{rt}})\}) + Q_i[n] P_{b,i}(\xi_t^{\text{rt}})]$  is jointly convex in  $(E_i, P_{b,i}, \{\mathbf{w}_k\})$  [cf. (22)], then the minimization over  $(P_{b,i}, \{\mathbf{w}_k\})$  is within a convex set; thus, (23a)-(23c) is still convex with respect to  $E_i[n]$  [25, Sec. 3.2.5]. In addition, due to  $\alpha_n^{\text{lt}} > \beta_n^{\text{lt}}$ , we can alternatively write  $G^{\text{lt}}(E_i[n]) = \max\{\alpha_n^{\text{lt}}(E_i[n] - A_{i,n}), \beta_n^{\text{lt}}(E_i[n] - A_{i,n})\}$ , which is in the family of convex functions. Hence, (24) is generally a nonsmooth and unconstrained convex problem with respect to  $\{E_i[n]\}$ , which can be solved using the stochastic subgradient iteration described next.

The subgradient of  $G^{\text{lt}}(E_i[n])$  can be first written as

$$\partial G^{\text{lt}}(E_i[n]) = \begin{cases} \alpha_n^{\text{lt}}, & \text{if } E_i[n] > A_{i,n} \\ \beta_n^{\text{lt}}, & \text{if } E_i[n] < A_{i,n} \\ \text{any } x \in [\beta_n^{\text{lt}}, \alpha_n^{\text{lt}}], & \text{if } E_i[n] = A_{i,n}. \end{cases}$$

With  $\{P_{b,i}^E(\xi_t^{\text{rt}}), \mathbf{w}_k^E(\xi_t^{\text{rt}})\}$  denoting the optimal solution for the problem in (25), the partial subgradient of  $\bar{G}^{\text{rt}}(\{E_i[n]\})$  with respect to  $E_i[n]$  is  $\partial_i \bar{G}^{\text{rt}}(\{E_i[n]\}) = V \mathbb{E}\{\partial \Psi^{\text{rt}}(E_i[n], P_{b,i}^E(\xi_t^{\text{rt}}), \{\mathbf{w}_k^E(\xi_t^{\text{rt}})\})\}$ , where

$$\partial \Psi^{\text{rt}}(E_i[n], P_{b,i}^E(\xi_t^{\text{rt}}), \{\mathbf{w}_k^E(\xi_t^{\text{rt}})\}) = \begin{cases} \frac{-\beta_t^{\text{rt}}}{T}, & \text{if } \frac{E_i[n]}{T} > \Delta \\ \frac{-\alpha_t^{\text{rt}}}{T}, & \text{if } \frac{E_i[n]}{T} < \Delta \\ x \in \left[ \frac{-\alpha_t^{\text{rt}}}{T}, \frac{-\beta_t^{\text{rt}}}{T} \right], & \text{else} \end{cases}$$

with  $\Delta := P_c + \sum_k \mathbf{w}_k^H(\xi_t^{\text{rt}}) \mathbf{B}_i \mathbf{w}_k^E(\xi_t^{\text{rt}}) + P_{b,i}^E(\xi_t^{\text{rt}})$ .

Defining  $\bar{g}_i(E_i) := V \partial G^{\text{lt}}(E_i) + T \partial_i \bar{G}^{\text{rt}}(\{E_i\})$ , a standard sub-gradient descent iteration can be employed to find the optimal  $E_i^*[n]$  for (24), as

$$E_i^{(j+1)}[n] = [E_i^{(j)}[n] - \mu^{(j)} \bar{g}_i(E_i^{(j)}[n])]^+, \quad \forall i \quad (26)$$

where  $j$  denotes iteration index, and  $\{\mu^{(j)}\}$  is the sequence of stepsizes.

Implementing (26) essentially requires performing (high-dimensional) integration over the unknown multivariate distribution function of  $\xi_t^{\text{rt}}$  present in  $\bar{g}_i$  through  $\bar{G}^{\text{rt}}$  in (25). To circumvent this impasse, a stochastic subgradient approach is devised based on the past realizations  $\{\xi_\tau^{\text{rt}}, \tau = 0, 1, \dots, nT-1\}$ . Per iteration  $j$ , we randomly draw a realization  $\xi_\tau^{\text{rt}}$  from past realizations, and run the following iteration

$$E_i^{(j+1)}[n] = [E_i^{(j)}[n] - \mu^{(j)} g_i(E_i^{(j)}[n])]^+, \quad \forall i \quad (27)$$

where  $g_i(E_i^{(j)}[n]) := V(\partial G^{\text{lt}}(E_i^{(j)}[n]) + T \partial \Psi^{\text{rt}}(E_i^{(j)}[n], P_{b,i}^E(\xi_\tau^{\text{rt}}), \{\mathbf{w}_k^E(\xi_\tau^{\text{rt}})\}))$  with  $\{P_{b,i}^E(\xi_\tau^{\text{rt}}), \mathbf{w}_k^E(\xi_\tau^{\text{rt}})\}$  obtained by solving a convex problem (25) with  $E_i[n] = E_i^{(j)}[n]$ .

As  $g_i(E_i^{(j)}[n])$  is indeed an unbiased random realization of  $\bar{g}_i(E_i^{(j)}[n]) = \mathbb{E}\{g_i(E_i^{(j)}[n])\}$  [28], if we adopt a sequence of non-summable diminishing stepsizes satisfying  $\lim_{j \rightarrow \infty} \mu^{(j)} = 0$  and  $\sum_{j=0}^{\infty} \mu^{(j)} = \infty$ , the iteration (27) asymptotically converges to the optimal  $\{E_i^*[n], \forall i\}$  as  $j \rightarrow \infty$  [29].

Compared with the empirical pdf based approach in [19], the proposed stochastic subgradient method is particularly tailored for our setting, which does not require the random vector  $\xi_t^{\text{rt}}$  to have discrete and finite support. In addition, as the former essentially belongs to the class of statistical learning based approaches [30], the proposed stochastic method avoids constructing a histogram for learning the distribution and requires a smaller number of samples to reach an accurate estimate of  $E_i^*[n]$ .

**Remark 2:** The computational complexity with the proposed algorithm is fairly low. Specifically, for solving the real-time energy balancing and beamforming problem (22) per slot  $t$ , the off-the-shelf interior-point solver incurs a worst-case complexity  $\mathcal{O}(I^{3.5} K^{3.5})$  to obtain the decisions  $\{P_{b,i}^*(t), \forall i; \mathbf{w}_k^*(t), \forall k\}$  [26]; for solving the ahead-of-time energy planning problem (25) every  $T$  slots, the stochastic subgradient approach needs  $\mathcal{O}(1/\epsilon^2)$  iterations to obtain an  $\epsilon$ -optimal solution, and the per iteration complexity is in the order of  $\mathcal{O}(I^{3.5} K^{3.5})$ . And updating  $E_i^{(j)}[n]$  in (27) requires only linear complexity  $\mathcal{O}(I)$ .

#### IV. PERFORMANCE ANALYSIS

In this section, we show that the TS-OC can yield a feasible and asymptotically (near-)optimal solution for problem (12).

##### A. Feasibility Guarantee

Note that in problems (19) and (20),  $\{C_i(t)\}$  are removed from the set of optimization variables and the constraints in (8) are ignored. While the battery dynamics  $C_i(t+1) = \eta C_i(t) + P_{b,i}(t)$  are accounted for by the TS-OC algorithm (in the step of “Queue updates”), it is not clear whether the resultant  $C_i(t) \in [C^{\min}, C^{\max}]$ ,  $\forall i, t$ . Yet, we will show that by selecting a pair  $(\Gamma, V)$  in (15), we can guarantee that  $C^{\min} \leq C_i(t) \leq C^{\max}$ ,  $\forall i, t$ ; meaning, the online control policy produced by the TS-OC is a feasible one for the original problem (12), under the conditions (13)–(14).

To this end, we first show the following lemma.

**Lemma 1:** If  $\bar{\alpha} := \max\{\alpha_i^{\text{rt}}, \forall t\}$  and  $\underline{\beta} := \min\{\beta_i^{\text{rt}}, \forall t\}$ , the battery (dis)charging amounts  $P_{b,i}^*(t)$  obtained from the TS-OC algorithm satisfy: i)  $P_{b,i}^*(t) = P_b^{\min}$ , if  $C_i(n_t T) > -V\underline{\beta} - \Gamma$ ; and ii)  $P_{b,i}^*(t) = P_b^{\max}$ , if  $C_i(n_t T) < -V\bar{\alpha} - \Gamma$ .

*Proof:* In TS-OC, we determine  $P_{b,i}^*(t)$  by solving (20). From the equivalent problem (22), we can see that the determination of  $P_{b,i}^*(t)$  is decoupled across BSs, and it depends on the first derivative of  $G^{\text{rt}}(\cdot)$ . By (21), the maximum possible gradient for  $G^{\text{rt}}(\cdot)$  is  $V\bar{\alpha}$ . It then follows that if  $V\bar{\alpha} + Q_i(n_t T) < 0$ , we must have  $P_{b,i}^*(t) = P_b^{\max}$ . Similarly, if  $V\underline{\beta} + Q_i(n_t T) > 0$ , we must have  $P_{b,i}^*(t) = P_b^{\min}$ . Given that  $Q_i(t) = C_i(t) + \Gamma$ , the lemma follows readily. ■

Lemma 1 reveals partial characteristics of the dynamic TS-OC policy. Specifically, when the energy queue (i.e., battery SoC) is large enough, the battery must be discharged as much as possible; that is,  $P_{b,i}^*(t) = P_b^{\min}$ . On the other hand, when the energy queue is small enough, the battery must be charged as much as possible; i.e.,  $P_{b,i}^*(t) = P_b^{\max}$ . Alternatively, such results can be justified by the economic interpretation of the virtual queues. Specifically,  $-\frac{Q_i(t)}{V}$  can be viewed as the instantaneous discharging price. For high prices  $-\frac{Q_i(t)}{V} > \bar{\alpha}$ , the TS-OC dictates full charge. Conversely, the battery units can afford full discharge if the price is low.

Based on the structure in Lemma 1, we can thus establish the following result.

**Proposition 1:** Under the conditions (13)–(14), the TS-OC algorithm with any pair  $(\Gamma, V)$  specified in (15) guarantees  $C^{\min} \leq C_i(t) \leq C^{\max}$ ,  $\forall i, \forall t$ .

*Proof:* See Appendix A. ■

**Remark 3:** Note that Proposition 1 is a *sample path* result; meaning, the bounded energy queues  $C_i(t) \in [C^{\min}, C^{\max}]$ ,  $\forall i$ , hold per time slot under *arbitrary*, even non-stationary,  $\{\mathbf{A}_n, \alpha_n^{\text{lt}}, \beta_n^{\text{lt}}, \alpha_t^{\text{rt}}, \beta_t^{\text{rt}}, \mathbf{H}_t\}$  processes. In other words, under the mild conditions (13)–(14), the proposed TS-OC with proper selection of  $(\Gamma, V)$  always yields a feasible control policy for (12).

##### B. Asymptotic Optimality

To facilitate the analysis, we assume that the random process  $\{\xi_n^{\text{lt}}\}$  and  $\{\xi_t^{\text{rt}}\}$  are both i.i.d. over slow and fast timescales,

respectively. Define  $\bar{C}_i := \frac{1}{NT} \sum_{t=0}^{NT-1} \mathbb{E}\{C_i(t)\}$  and  $\bar{P}_{b,i} := \frac{1}{NT} \sum_{t=0}^{NT-1} \mathbb{E}\{P_{b,i}(t)\}$ . Since  $P_{b,i}(t) \in [P_b^{\min}, P_b^{\max}]$  and  $C_i(t+1) = \eta C_i(t) + P_{b,i}(t)$ , it holds that

$$\bar{P}_{b,i} = \frac{1}{NT} \sum_{t=0}^{NT-1} \mathbb{E}\{C_i(t+1) - \eta C_i(t)\} = (1 - \eta)\bar{C}_i. \quad (28)$$

As  $C_i(t) \in [C^{\min}, C^{\max}]$ ,  $\forall t$ , (28) then implies

$$(1 - \eta)C^{\min} \leq \bar{P}_{b,i} \leq (1 - \eta)C^{\max}, \quad \forall i. \quad (29)$$

Consider now the following problem

$$\begin{aligned} \tilde{\Phi}^{\text{opt}} := \min_{\mathcal{X}} \lim_{N \rightarrow \infty} \frac{1}{NT} \sum_{t=0}^{NT-1} \sum_{i \in \mathcal{I}} \mathbb{E}\{\Phi_i(t)\} \\ \text{s. t. } (5), (6), (9), (10), \forall t, \quad (29). \end{aligned} \quad (30)$$

Note that the constraints in (8),  $\forall t$ , are replaced by (29); i.e., the queue dynamics that need to be performed per realization per slot are replaced by a time-averaged constraint per BS  $i$ . The problem (30) is thus a relaxed version of (12) [23]. Specifically, any feasible solution of (12), satisfying (8),  $\forall t$ , also satisfies (29) in (30), due to the boundedness of  $P_{b,i}(t)$  and  $C_i(t)$ . It then follows that  $\tilde{\Phi}^{\text{opt}} \leq \Phi^{\text{opt}}$ .

Variables  $\{C_i(t)\}$  are removed from (30), and other optimization variables are “decoupled” across time slots due to the removal of constraints (8). This problem has an easy-to-characterize stationary optimal control policy as formally stated in the next lemma.

**Lemma 2:** If  $\xi_n^{\text{lt}}$  and  $\xi_t^{\text{rt}}$  are i.i.d., there exists a stationary control policy  $\mathcal{P}^{\text{stat}}$  that is a pure (possibly randomized) function of the current  $(\xi_{n_t}^{\text{lt}}, \xi_t^{\text{rt}})$ , while satisfying (5), (6), (9), (10), and providing the following guarantees per  $t$ :

$$\begin{aligned} \mathbb{E}\left\{\sum_{i \in \mathcal{I}} \Phi_i^{\text{stat}}(t)\right\} = \tilde{\Phi}^{\text{opt}} \\ (1 - \eta)C^{\min} \leq \mathbb{E}\{P_{b,i}^{\text{stat}}(t)\} \leq (1 - \eta)C^{\max}, \quad \forall i \end{aligned} \quad (31)$$

where  $P_{b,i}^{\text{stat}}(t)$  denotes the decided (dis)charging amount,  $\Phi_i^{\text{stat}}(t)$  the resultant transaction cost by policy  $\mathcal{P}^{\text{stat}}$ , and expectations are taken over the randomization of  $(\xi_{n_t}^{\text{lt}}, \xi_t^{\text{rt}})$  and (possibly)  $\mathcal{P}^{\text{stat}}$ .

*Proof:* The proof argument is similar to that in [31, Theorem 4.5]; hence, it is omitted for brevity. ■

Lemma 2 in fact holds for many non-i.i.d. scenarios as well. Generalizations to other stationary processes, or even to non-stationary processes, can be found in [31] and [32].

It is worth noting that (31) not only assures that the stationary control policy  $\mathcal{P}^{\text{stat}}$  achieves the optimal cost for (30), but also guarantees that the resultant expected transaction cost per slot  $t$  is equal to the optimal time-averaged cost (due to the stationarity of  $\xi_{n_t}^{\text{lt}}$ ,  $\xi_t^{\text{rt}}$  and  $\mathcal{P}^{\text{stat}}$ ). This plays a critical role in establishing the following result.

**Proposition 2:** Suppose that conditions (13)–(15) hold. If  $\xi_n^{\text{lt}}$  and  $\xi_t^{\text{rt}}$  are i.i.d. across time, then the time-averaged cost under the proposed TS-OC algorithm satisfies

$$\lim_{N \rightarrow \infty} \frac{1}{NT} \sum_{t=0}^{NT-1} \sum_{i \in \mathcal{I}} \mathbb{E}\{\Phi_i^*(t)\} \leq \Phi^{\text{opt}} + \frac{M_1 + M_2 + M_3}{V}$$

where the constants<sup>2</sup>

$$M_1 := \frac{IT(1-\eta)}{2\eta(1-\eta^T)} M_B \quad (32)$$

$$M_2 := \frac{I[T(1-\eta) - (1-\eta^T)]}{(1-\eta)(1-\eta^T)} M_B \quad (33)$$

$$M_3 := I(1-\eta)M_C \quad (34)$$

with  $M_B$  and  $M_C$  given by

$$M_B := \max\{[(1-\eta)\Gamma + P_b^{\min}]^2, [(1-\eta)\Gamma + P_b^{\max}]^2\}$$

$$M_C := \max\{(\Gamma + C^{\min})^2, (\Gamma + C^{\max})^2\};$$

$\Phi_i^*(t)$  denotes the resultant cost with the TS-OC, and  $\Phi^{opt}$  is the optimal value of (12) under any feasible control algorithm, including the one knowing all future realizations.

*Proof:* See Appendix B. ■

**Remark 4:** Proposition 2 asserts that the proposed TS-OC algorithm yields a time-averaged cost with optimality gap smaller than  $\frac{M_1+M_2+M_3}{V}$ . The proposed TS-OC is in fact a modified version of the classic queue-length based stochastic optimization scheme, where queue lengths play the role of “stochastic” Lagrange multipliers with a dual-subgradient solver to the regularized dual problem by subtracting an  $\ell_2$ -norm of Lagrange multipliers. Intuitively, the gap  $M_1/V$  is inherited from the underlying stochastic subgradient method. The gap  $M_2/V$  is introduced by the inaccurate queue lengths in use (since we use  $Q_i(nT)$ , instead of  $Q_i(t)$ , for all  $t = nT, \dots, (n+1)T - 1$ ), while the gap  $M_3/V$  is incurred by the presence of  $\ell_2$  regularizer in the dual function (a. k. a. the price of battery imperfections).

### C. Main Theorem

Based on Propositions 1 and 2, it is now possible to arrive at our main result.

**Theorem 1:** Suppose that conditions (13)–(15) hold and  $(\xi_n^{\text{lt}}, \xi_t^{\text{rt}})$  are i.i.d. over slots. Then the proposed TS-OC yields a feasible dynamic control scheme for (12), which is asymptotically near-optimal in the sense that

$$\Phi^{opt} \leq \lim_{N \rightarrow \infty} \frac{1}{NT} \sum_{t=0}^{NT-1} \sum_{i \in \mathcal{I}} \mathbb{E}\{\Phi_i^*(t)\} \leq \Phi^{opt} + \frac{M}{V}$$

where  $M := M_1 + M_2 + M_3$ , as specified in Proposition 2.

The asymptotic behavior of the proposed dynamic approach is more complicated than that of existing alternatives due to the nature of multi-scale scheduling and battery imperfections. Interesting comments on the minimum optimality gap with the TS-OC are now in order.

- 1) When  $\eta = 1$  (perfect battery), the optimality gap between the TS-OC and the offline optimal scheduling reduces to

$$\frac{M}{V} = \frac{M_1 + M_2}{V} = \frac{IT}{2V} \max\{(P_b^{\min})^2, (P_b^{\max})^2\}.$$

The typical tradeoff from the stochastic network optimization holds in this case [31]: an  $\mathcal{O}(V)$  battery size

is necessary, when an  $\mathcal{O}(1/V)$  close-to-optimal cost is achieved. Clearly, the minimum optimality gap is given by  $M/V^{\max}$ , which vanishes as  $V^{\max} \rightarrow \infty$ . By (18), such an asymptotic optimality can be achieved when we have very small price difference  $(\bar{\alpha} - \underline{\beta})$ , or very large battery capacities  $C^{\max}$ .

- 2) When  $\eta \in (0, 1)$ , the constants  $M_1$ ,  $M_2$  and  $M_3$  are in fact functions of  $\Gamma$ , whereas the minimum and maximum values of  $\Gamma$  also depend on  $V$  [cf. (16)–(17)], thus the typical tradeoff in the case 1) is no longer correct. For a given  $V^{\max}$ , the minimum optimality gap,  $G^{\min}(V^{\max})$ , can be obtained by solving the following problem:

$$\min_{(V, \Gamma)} \frac{M}{V} = \frac{M_1(\Gamma)}{V} + \frac{M_2(\Gamma)}{V} + \frac{M_3(\Gamma)}{V}, \text{ s. t. (15)}. \quad (35)$$

For  $V \geq 0$ , we know that the quadratic-over-linear functions  $\frac{[(1-\eta)\Gamma + P_b^{\min}]^2}{V}$  and  $\frac{[(1-\eta)\Gamma + P_b^{\max}]^2}{V}$  are jointly convex in  $V$  and  $\Gamma$  [25]. As a point-wise maximum of these two convex functions,  $\frac{M_B(\Gamma)}{V}$  is also convex [25]. Then  $\frac{M_1(\Gamma)}{V}$  and  $\frac{M_2(\Gamma)}{V}$  are clearly convex by (32)–(33); and likewise for  $\frac{M_3(\Gamma)}{V}$ . Since the objective is convex and the constraints in (15) are linear, problem (35) is a convex program which can be efficiently solved by general interior-point methods. Note that  $G^{\min}(V^{\max})$  no longer monotonically decreases with respect to  $V^{\max}$  (or  $C^{\max}$ ); see also [23]. This makes sense intuitively because for a large battery capacity, the impact of using inaccurate queue lengths (battery SoC) and the dissipation loss due to battery imperfections will also be enlarged. The smallest possible optimality gap can be numerically computed by one dimensional search over  $G^{\min}(V^{\max})$  with respect to  $V^{\max}$ .

## V. NUMERICAL TESTS

In this section, simulated tests are presented to evaluate our proposed TS-OC algorithm, and justify the analytical claims in Section IV.

### A. Experiment setup

The considered CoMP network includes  $I = 2$  BSs each with  $M = 2$  transmit antennas, and  $K = 3$  mobile users. The system bandwidth is 1 MHz, and each element in channel vectors  $\mathbf{h}_{ik,t}$ ,  $\forall i, k, t$ , is a zero-mean complex-Gaussian random variable with unit variance. Each coarse-grained interval consists of  $T = 5$  time slots. The limits of  $P_{g,i}$ ,  $P_{b,i}$  and  $C_i$ , as well as the values of the initial SoC  $C_i(0)$  and  $P_c$  are listed in Table I. The battery storage efficiency is  $\eta = 0.95$ . The ahead-of-time and real-time energy purchase prices  $\alpha_n^{\text{lt}}$  and  $\alpha_t^{\text{rt}}$  are generated from folded normal distributions, with  $\mathbb{E}\{\alpha_n^{\text{lt}}\} = 1.15$  and  $\mathbb{E}\{\alpha_t^{\text{rt}}\} = 2.3$ . The selling prices are set as  $\beta_n^{\text{lt}} = 0.9 \times \alpha_n^{\text{lt}}$  and  $\beta_t^{\text{rt}} = 0.3 \times \alpha_t^{\text{rt}}$ . The harvested energy  $A_{i,n}$  is also generated from a folded normal distribution. Finally, the Lyapunov control parameter  $V$  is chosen as  $V = V^{\max}$ . The proposed TS-OC algorithm is compared with three baseline schemes to benchmark its performance. ALG 1 is a one-scale scheme without ahead-of-time energy planning; ALG 2 performs two-scale online control without leveraging

<sup>2</sup>Note that  $\lim_{\eta \rightarrow 1} \frac{1-\eta^T}{1-\eta} = T$ , and  $\lim_{\eta \rightarrow 1} \frac{T(1-\eta) - (1-\eta^T)}{(1-\eta)(1-\eta^T)} = \frac{T-1}{2}$ .



TABLE I  
PARAMETER VALUES. ALL UNITS ARE KWH.

$P_c$	$P_g^{\max}$	$P_b^{\min}$	$P_b^{\max}$	$C^{\min}$	$C^{\max}$	$C_i(0)$
10	50	-2	2	0	80	0

the renewable energy or energy storage devices; and the offline benchmark is an ideal scheme with a-priori knowledge of all future channel state, energy price and RES arrival realizations.

### B. Numerical results

Fig. 3 shows the running-average transaction costs of the proposed algorithm, ALGs 1-2, as well as the offline benchmark. It is shown that within 500 time slots, the proposed approach converges closest to the lower bound, while ALGs 1-2 incur about 71% and 31% larger costs than the proposed one. However, note that the optimal offline counterpart cannot work in practice due to the lack of future information, and the optimality gap could be reduced if the battery efficiency  $\eta$  approaches to 1. Among online schemes, the TS-OC algorithm intelligently takes advantage of the ahead-of-time energy planning, and the renewable energy and batteries, to hedge against future potential high energy cost, while ALGs 1-2, due to the lack of these features, have to purchase much more expensive energy from the real-time energy market and result in a higher transaction cost.

The theoretical optimality-gap [cf. (35)] between the TS-OC and the offline optimal scheduling is depicted in Fig. 4 under different battery capacities  $C^{\max}$ . As analyzed after Theorem 1, the optimality-gap  $M/V$  for  $\eta = 1$  diminishes as  $C^{\max}$  (or  $V^{\max}$ ) grows; whereas the gaps for  $\eta = 0.9$  and  $\eta = 0.95$  are no longer monotonically decreasing. Specifically, both of them first decrease and then increase, reaching the lowest points (where the optimality gaps are minimized) at  $C^{\max} = 40$  kWh and  $C^{\max} = 55$  kWh, respectively. As expected, the gap for the worst storage efficiency  $\eta = 0.9$  remains the largest across the entire spectrum of battery capacity.

In Fig. 5, the average transaction cost of the TS-OC is compared under different battery efficiencies  $\eta = 0.9, 0.95, 1$ . Clearly, the average costs monotonically decrease as  $C^{\max}$  grows. The BSs with imperfect batteries ( $\eta = 0.9, 0.95$ ) require larger budgets for energy purchase than the ones with perfect batteries ( $\eta = 1$ ), thus compensating for the battery degeneration losses. In particular, when  $C^{\max} = 120$  kWh, the costs for  $\eta = 0.9$  and  $\eta = 0.95$  are 41.8% and 33.8% larger than that of the perfect battery case, respectively.

The evolutions of battery SoC  $C_1(t)$  with different storage efficiencies  $\eta$  are compared in Fig. 6. Clearly, all the three lines fluctuate within the feasible region; i.e.,  $C^{\min} \leq C_1(t) \leq C^{\max}$ . Among the three cases, the battery with  $\eta = 1$  maintains the highest energy level, followed by those with  $\eta = 0.95$  and  $\eta = 0.9$ . Intuitively speaking, keeping a high energy level in an imperfect battery results in much higher energy dissipation losses. As a result, the TS-OC algorithm tends to maintain a low energy level in such cases (e.g., around 30 kWh for  $\eta = 0.9$ ) to reduce average energy loss, and (dis)charge the battery less frequently.

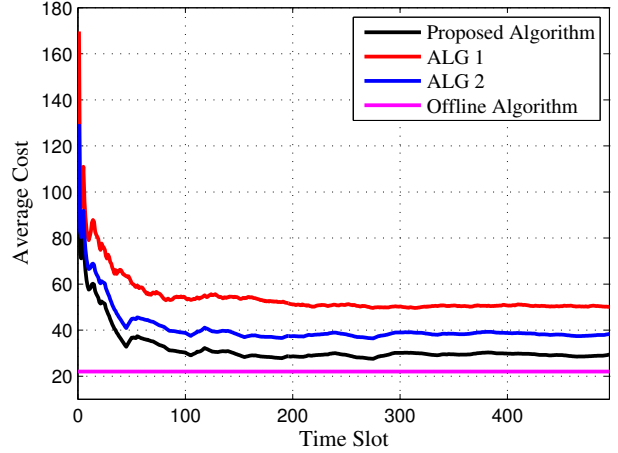


Fig. 3. Comparison of average transaction cost.

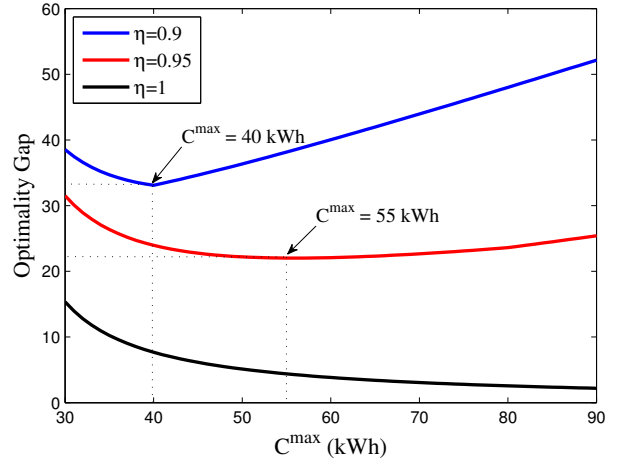
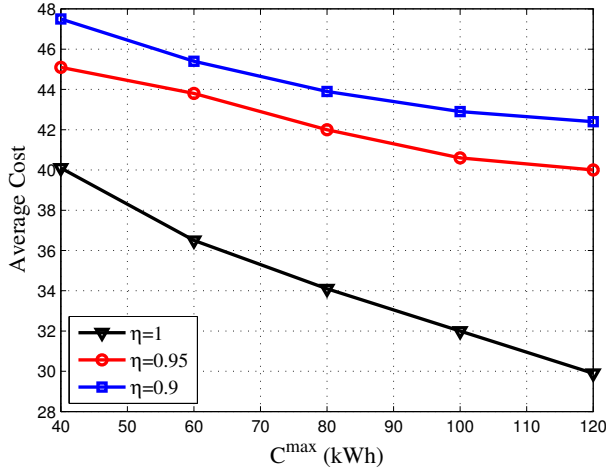
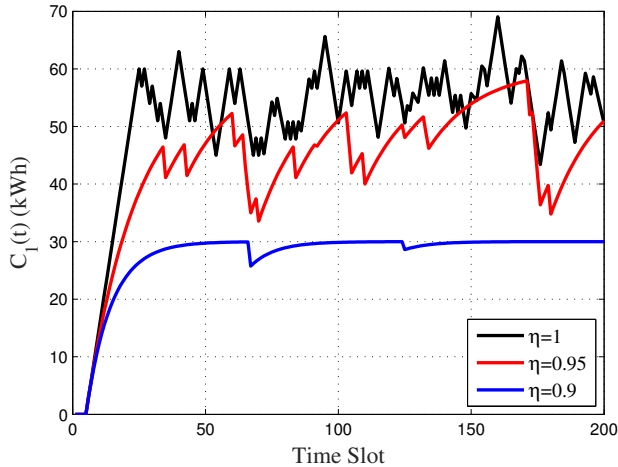


Fig. 4. Optimality-gap versus battery capacity  $C^{\max}$ .

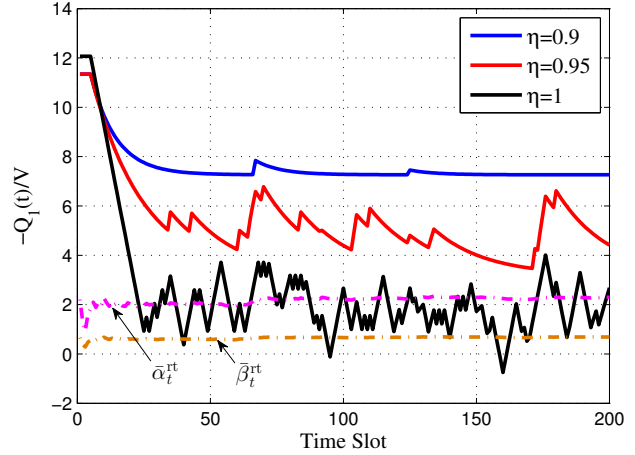
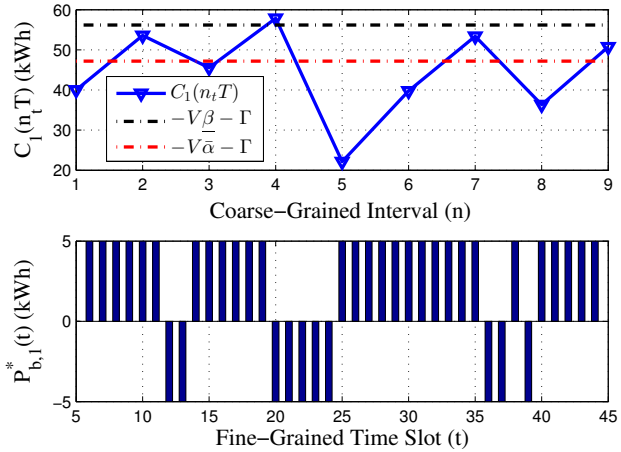
The previous remarks are further substantiated by Fig. 7, where the instantaneous discharging price, or, the “stochastic” Lagrange multiplier  $-\frac{Q_1(t)}{V}$  is compared with the running-average purchase and selling prices  $\bar{\alpha}_t^{\text{rt}} := (1/t) \sum_{\tau=1}^t \alpha_{\tau}^{\text{rt}}$  and  $\bar{\beta}_t^{\text{rt}} := (1/t) \sum_{\tau=1}^t \beta_{\tau}^{\text{rt}}$ . It is interesting to observe that with a perfect battery ( $\eta = 1$ ), the instantaneous discharging price  $-\frac{Q_1(t)}{V}$  is hovering between the average purchase and selling prices, which features a frequent (dis)charging operation. For  $\eta = 0.95$  or  $\eta = 0.9$ ,  $-\frac{Q_1(t)}{V}$  is relatively high compared to the average purchase and selling prices, which discourages frequent (dis)charging; see also Fig. 6. Note that the evolution of  $-\frac{Q_1(t)}{V}$  can be further linked to the standard results from sensitivity analysis, which implies that the subdifferential of the objective  $\lim_{N \rightarrow \infty} \frac{1}{NT} \sum_{t=0}^{NT-1} \sum_i \mathbb{E}\{\Phi_i(t)\}$  with respect to  $P_{b,i}(t)$  (the convex hull of average purchase and selling prices) coincides with the negative of the optimal dual variable corresponding to (28) [25]. Building upon this claim, the asymptotic optimality can be easily verified for  $\eta = 1$  since the “stochastic” Lagrange multiplier  $-\frac{Q_1(t)}{V}$  converges to a neighborhood of the optimal dual variable; and a large optimality gap is also as expected for the imperfect batteries  $\eta < 1$  due to the distance between  $-\frac{Q_1(t)}{V}$  and the


 Fig. 5. Average transaction cost versus battery capacity  $C^{\max}$ .

 Fig. 6. TS-OC based schedule of the battery SoC  $C_1(t)$ .

average purchase and selling prices.

Taking a deeper look, the battery SoC  $C_1(n_t T)$  and the real-time battery (dis)charging  $P_{b,1}^*(t)$  are jointly depicted in Fig. 8 to reveal the (dis)charging characteristics stated in Lemma 1. It can be observed that the TS-OC dictates the full discharge  $P_{b,1}^*(t) = P_b^{\min}$  in the incoming 5 fine-grained slots  $t \in [20, 24]$  when  $C_1(n_t T) > -V\beta - \Gamma$  at  $n = 4$ , while the battery is fully charged  $P_{b,1}^*(t) = P_b^{\max}$  when  $C_1(n_t T) < -V\bar{\alpha} - \Gamma$  at  $n = 1, 3, 5, 6, 8$ . In addition, when  $C_1(n_t T) \in [-V\bar{\alpha} - \Gamma, -V\beta - \Gamma]$  at  $n = 2, 7$ ,  $P_{b,1}^*(t)$  must be obtained by solving (22) numerically.

Fig. 9 shows the optimal energy planning  $E_1^*[n]$  over a 100-slot period, and the fluctuation of ahead-of-time energy purchase prices  $\alpha_n^{\text{lt}}$  is also plotted to illustrate the resultant online policy. One observation is that the ahead-of-time energy purchase  $E_1^*[n]$  highly depends on the long-term price  $\alpha_n^{\text{lt}}$ . Specifically, the proposed scheme tends to request more energy for the future  $T$  slots when  $\alpha_n^{\text{lt}}$  is lower (e.g.,  $n = 10, 13, 17$ ), and tends to purchase less energy when  $\alpha_n^{\text{lt}}$  is higher (e.g.,  $n = 2, 11$ ).


 Fig. 7. The evolution of  $-Q_1(t)/V$  and running-average of energy prices.

 Fig. 8. TS-OC based schedule of the battery SoC  $C_1(n_t T)$  and battery (dis)charging actions  $P_{b,1}^*(t)$ , where  $P_b^{\max} = 5$  kWh and  $P_b^{\min} = -5$  kWh.

## VI. CONCLUSIONS

A two-scale dynamic resource allocation task was considered for RES-integrated CoMP transmissions. Taking into account the variability of channels, RES and ahead-of-time/real-time electricity prices, as well as battery imperfections, a stochastic optimization problem was formulated to minimize the long-term average energy transaction cost subject to the QoS requirements. Capitalizing on the Lyapunov optimization technique and the stochastic subgradient iteration, a two-scale online algorithm was developed to make control decisions ‘on-the-fly.’ It was analytically established that the novel approach yields feasible and asymptotically near-optimal resource schedules without knowing any statistics of the underlying stochastic processes. Simulated tests confirmed the merits of the proposed approach and highlighted the effect of battery imperfections on the proposed online scheme. This novel two-scale optimization framework opens up some interesting research directions, which include incorporating the power network constraints and/or transmission losses in our formulation, pursuing a fast convergent approach by learning from the historical system statistics, and reducing the battery size

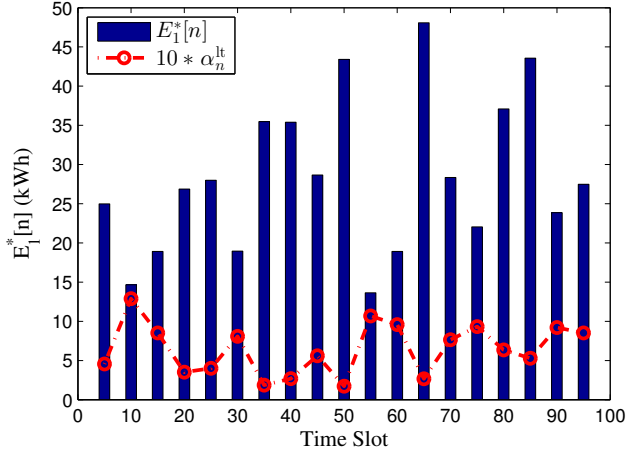


Fig. 9. TS-OC based schedule of the optimal energy planning  $E_1^*[n]$ .

leveraging the so-called predictive scheduling with predicted future information.

## APPENDIX

### A. Proof of Proposition 1

The proof proceeds by induction. First, set  $C_i(0) \in [C^{\min}, C^{\max}]$ ,  $\forall i$ , and suppose that this holds for all  $C_i(nT)$  at slot  $nT$ . We will show the bounds hold for  $C_i(t)$ ,  $\forall t = nT + 1, \dots, (n+1)T$ , as well as in subsequent instances.

By  $C_i(t+1) = \eta C_i(t) + P_{b,i}^*(t)$ , we have

$$C_i(t) = \eta^{t-nT} C_i(nT) + \sum_{\tau=nT}^{t-1} [\eta^{t-1-\tau} P_{b,i}^*(\tau)],$$

$$\forall t = nT + 1, \dots, (n+1)T. \quad (36)$$

Note that by the definitions of  $\Gamma^{\min}$  and  $\Gamma^{\max}$  in (16)-(17), we have  $C^{\min} \leq -V\bar{\alpha} - \Gamma < -V\underline{\beta} - \Gamma \leq C^{\max}$ . We then consider the following three cases.

c1) If  $C_i(nT) \in [C^{\min}, -V\bar{\alpha} - \Gamma)$ , then Lemma 1 implies that  $P_{b,i}^*(t) = P_b^{\max}$ ,  $\forall t = nT, \dots, (n+1)T - 1$ . From (36), we have,  $\forall t = nT + 1, \dots, (n+1)T$ ,

i)  $C_i(t) \geq \eta^{t-nT} C^{\min} + \frac{1-\eta^{t-nT}}{1-\eta} P_b^{\max} \geq C^{\min}$ , due to the condition (13);

ii)  $C_i(t) \leq \eta^{t-nT} (-V\bar{\alpha} - \Gamma) + \frac{1-\eta^{t-nT}}{1-\eta} P_b^{\max} \leq \eta^{t-nT} (-V\underline{\beta} - \Gamma) + \frac{1-\eta^{t-nT}}{1-\eta} P_b^{\max} \leq C^{\max}$ , due to  $\underline{\beta} < \bar{\alpha}$ ,  $\Gamma \geq \Gamma^{\min}$ , and the definition of  $\Gamma^{\min}$  in (16).

c2) If  $C_i(nT) \in [-V\bar{\alpha} - \Gamma, -V\underline{\beta} - \Gamma]$ , then  $P_{b,i}^*(t) \in [P_b^{\min}, P_b^{\max}]$ . We have,  $\forall t = nT + 1, \dots, (n+1)T$ ,

i)  $C_i(t) \geq \eta^{t-nT} (-V\bar{\alpha} - \Gamma) + \frac{1-\eta^{t-nT}}{1-\eta} P_b^{\min} \geq C^{\min}$ , due to  $\Gamma \leq \Gamma^{\max}$  and the definition of  $\Gamma^{\max}$  in (17);

ii)  $C_i(t) \leq \eta^{t-nT} (-V\underline{\beta} - \Gamma) + \frac{1-\eta^{t-nT}}{1-\eta} P_b^{\max} \leq C^{\max}$ , as with c1-ii); and

c3) If  $C_i(nT) \in (-V\underline{\beta} - \Gamma, C^{\max}]$ , it follows from Lemma 1 that  $P_{b,i}^*(t) = P_b^{\min}$ ,  $\forall t = nT, \dots, (n+1)T - 1$ . We have,  $\forall t = nT + 1, \dots, (n+1)T$

i)  $C_i(t) \geq \eta^{t-nT} (-V\underline{\beta} - \Gamma) + \frac{1-\eta^{t-nT}}{1-\eta} P_b^{\min} \geq \eta^{t-nT} (-V\bar{\alpha} - \Gamma) + \frac{1-\eta^{t-nT}}{1-\eta} P_b^{\min} \geq C^{\min}$ , due to  $\underline{\beta} < \bar{\alpha}$  and c2-i);

ii)  $C_i(t) \leq \eta^{t-nT} C^{\max} + \frac{1-\eta^{t-nT}}{1-\eta} P_b^{\min} \leq C^{\max}$ , due to  $\eta \leq 1$ , and  $P_b^{\min} < 0$ .

Cases c1)–c3) together prove the proposition.

### B. Proof of Proposition 2

The evolution of  $Q_i(t)$  in the TS-OC is given by  $Q_i(t+1) = C_i(t+1) + \Gamma = \eta C_i(t) + P_{b,i}^*(t) + \Gamma = \eta Q_i(t) + (1-\eta)\Gamma + P_{b,i}^*(t)$ . Hence, we have

$$\begin{aligned} [Q_i(t+1)]^2 &= [\eta Q_i(t) + (1-\eta)\Gamma + P_{b,i}^*(t)]^2 \\ &= \eta^2 [Q_i(t)]^2 + 2\eta Q_i(t) [(1-\eta)\Gamma + P_{b,i}^*(t)] \\ &\quad + [(1-\eta)\Gamma + P_{b,i}^*(t)]^2 \\ &\leq \eta^2 [Q_i(t)]^2 + 2\eta Q_i(t) [(1-\eta)\Gamma + P_{b,i}^*(t)] \\ &\quad + \max\{[(1-\eta)\Gamma + P_b^{\min}]^2, [(1-\eta)\Gamma + P_b^{\max}]^2\} \end{aligned}$$

where the last inequality holds due to (9).

With  $\mathbf{Q}(t) := [Q_1(t), \dots, Q_I(t)]'$ , consider the Lyapunov function  $L(\mathbf{Q}(t)) := \frac{1}{2} \sum_i [Q_i(t)]^2$ . Using the short-hand notation  $\mathbf{Q}[n] := \mathbf{Q}(nT)$ , it readily follows that

$$\begin{aligned} \Delta_T(\mathbf{Q}[n]) &:= L(\mathbf{Q}[n+1]) - L(\mathbf{Q}[n]) \\ &\leq -\frac{1}{2}(1-\eta^2) \sum_{t=nT}^{(n+1)T-1} \sum_{i \in \mathcal{I}} [Q_i(t)]^2 + \frac{IT}{2} M_B \\ &\quad + \sum_{t=nT}^{(n+1)T-1} \sum_{i \in \mathcal{I}} \{\eta Q_i(t) [(1-\eta)\Gamma + P_{b,i}^*(t)]\} \\ &\leq \frac{IT}{2} M_B + \sum_{t=nT}^{(n+1)T-1} \sum_{i \in \mathcal{I}} \{\eta Q_i(t) [(1-\eta)\Gamma + P_{b,i}^*(t)]\}. \end{aligned}$$

Since  $Q_i(t+1) = \eta Q_i(t) + (1-\eta)\Gamma + P_{b,i}^*(t)$  and  $P_b^{\min} \leq P_{b,i}^*(t) \leq P_b^{\max}$ , we have:  $\forall t = nT, \dots, (n+1)T - 1$ ,

$$\begin{aligned} \eta^{t-nT} Q_i[n] + \frac{1-\eta^{t-nT}}{1-\eta} [(1-\eta)\Gamma + P_b^{\min}] &\leq Q_i(t) \\ &\leq \eta^{t-nT} Q_i[n] + \frac{1-\eta^{t-nT}}{1-\eta} [(1-\eta)\Gamma + P_b^{\max}]. \end{aligned}$$

This implies that  $\forall t = nT, \dots, (n+1)T - 1$ ,

$$\begin{aligned} Q_i(t) [(1-\eta)\Gamma + P_{b,i}^*(t)] &\leq \eta^{t-nT} Q_i[n] [(1-\eta)\Gamma + P_{b,i}^*(t)] \\ &\quad + \frac{1-\eta^{t-nT}}{1-\eta} \max\{[(1-\eta)\Gamma + P_b^{\min}]^2, [(1-\eta)\Gamma + P_b^{\max}]^2\}. \end{aligned}$$

Consequently, it follows that

$$\begin{aligned} \Delta_T(\mathbf{Q}[n]) &\leq \frac{IT}{2} M_B + \sum_{t=nT}^{(n+1)T-1} \sum_{i \in \mathcal{I}} \left\{ \frac{\eta(1-\eta^{t-nT})}{1-\eta} M_B \right. \\ &\quad \left. + \sum_{t=nT}^{(n+1)T-1} \sum_{i \in \mathcal{I}} \{\eta^{t-nT+1} Q_i[n] [(1-\eta)\Gamma + P_{b,i}^*(t)]\} \right\} \\ &\leq \frac{IT}{2} M_B + \frac{I\eta[T(1-\eta) - (1-\eta^T)]}{(1-\eta)^2} M_B \\ &\quad + \sum_{t=nT}^{(n+1)T-1} \sum_{i \in \mathcal{I}} \{\eta^{t-nT+1} Q_i[n] [(1-\eta)\Gamma + P_{b,i}^*(t)]\}. \end{aligned}$$

Taking expectations and adding  $\sum_{t=nT}^{(n+1)T-1} \sum_i [\eta^{t-nT+1} V\mathbb{E}\{\Phi_i^*(t)\}]$  to both sides, we arrive at (with short-hand notation  $M_\Delta := \frac{I\eta}{2} M_B + \frac{I\eta[T(1-\eta)-(1-\eta^T)]}{(1-\eta)^2} M_B$ ):

$$\begin{aligned}
& \mathbb{E}\{\Delta_T(\mathbf{Q}[n])\} + \sum_{t=nT}^{(n+1)T-1} \eta^{t-nT+1} \sum_{i \in \mathcal{I}} [V\mathbb{E}\{\Phi_i^*(t)\}] \\
& \leq M_\Delta + \sum_{t=nT}^{(n+1)T-1} [\eta^{t-nT+1} \sum_{i \in \mathcal{I}} Q_i[n](1-\eta)\Gamma] \\
& \quad + \sum_{t=nT}^{(n+1)T-1} [\eta^{t-nT+1} \sum_{i \in \mathcal{I}} \mathbb{E}\{V\Phi_i^*(t) + Q_i[n]P_{b,i}^*(t)\}] \\
& = M_\Delta + \sum_{t=nT}^{(n+1)T-1} [\eta^{t-nT+1} \sum_{i \in \mathcal{I}} Q_i[n](1-\eta)\Gamma] \\
& \quad + \frac{\eta(1-\eta^T)}{(1-\eta)T} \sum_{t=nT}^{(n+1)T-1} \sum_{i \in \mathcal{I}} [\sum \mathbb{E}\{V\Phi_i^*(t) + Q_i[n]P_{b,i}^*(t)\}] \\
& \leq M_\Delta + \sum_{t=nT}^{(n+1)T-1} [\eta^{t-nT+1} \sum_{i \in \mathcal{I}} Q_i[n](1-\eta)\Gamma] \\
& \quad + \frac{\eta(1-\eta^T)}{(1-\eta)T} \sum_{t=nT}^{(n+1)T-1} \sum_{i \in \mathcal{I}} [\sum \mathbb{E}\{V\Phi_i^{stat}(t) + Q_i[n]P_{b,i}^{stat}(t)\}] \\
& = M_\Delta + \sum_{t=nT}^{(n+1)T-1} [\eta^{t-nT+1} \sum_{i \in \mathcal{I}} \mathbb{E}\{V\Phi_i^{stat}(t)\}] \\
& \quad + \sum_{t=nT}^{(n+1)T-1} [\eta^{t-nT+1} Q_i[n]\{(1-\eta)\Gamma + P_{b,i}^{stat}(t)\}] \\
& \leq M_\Delta + I\eta(1-\eta^T)M_C + \frac{\eta(1-\eta^T)}{1-\eta} V\tilde{\Phi}^{opt}
\end{aligned}$$

where the two equalities hold since both  $\sum_{i \in \mathcal{I}} \mathbb{E}\{V\Phi_i^*(t) + Q_i[n]P_{b,i}^*(t)\}$  for the TS-OC and  $\sum_{i \in \mathcal{I}} \mathbb{E}\{V\Phi_i^{stat}(t) + Q_i[n]P_{b,i}^{stat}(t)\}$  for  $\mathcal{P}^{stat}$  are in fact the same for slots  $t = nT, \dots, (n+1)T-1$ , when  $\xi_i^{st}$  is i.i.d. over slots; the second inequality is because the TS-OC algorithm minimizes the third term  $\sum_i \mathbb{E}\{V\Phi_i(t) + Q_i[n]P_{b,i}(t)\}$  among all policies satisfying (5), (6), (9), and (10), including  $\mathcal{P}^{stat}$ ; and the last inequality is due to (31) and  $Q_i[n] \in [C^{\min} + \Gamma, C^{\max} + \Gamma]$  under conditions (13)–(15) per Proposition 1.

Again, note that  $\sum_i [V\mathbb{E}\{\Phi_i^*(t)\}]$  for the TS-OC is the same for slots  $t = nT, \dots, (n+1)T-1$ , when  $\xi_i^{st}$  is i.i.d. over slots. Summing over all  $n = 1, 2, \dots$ , we then have

$$\begin{aligned}
& \sum_{n=0}^{N-1} \mathbb{E}\{\Delta_T(\mathbf{Q}[n])\} + \sum_{n=0}^{N-1} \sum_{t=nT}^{(n+1)T-1} \eta^{t-nT+1} \sum_{i \in \mathcal{I}} [V\mathbb{E}\{\Phi_i^*(t)\}] \\
& = \mathbb{E}[L(\mathbf{Q}[N])] - L(\mathbf{Q}[0]) + \frac{\eta(1-\eta^T)}{(1-\eta)T} \sum_{t=0}^{NT-1} \sum_{i \in \mathcal{I}} [V\mathbb{E}\{\Phi_i^*(t)\}] \\
& \leq N[M_\Delta + I\eta(1-\eta^T)M_C] + \frac{\eta(1-\eta^T)}{1-\eta} V\tilde{\Phi}^{opt}
\end{aligned}$$

which leads to

$$\begin{aligned}
& \frac{1}{NT} \sum_{t=0}^{NT-1} \mathbb{E}[\sum_{i \in \mathcal{I}} \mathbb{E}\{\Phi_i^*(t)\}] \\
& \leq \tilde{\Phi}^{opt} + \frac{M_1 + M_2 + M_3}{V} + \frac{(1-\eta)}{\eta(1-\eta^T)} \frac{L(\mathbf{Q}[0])}{NV} \\
& \leq \Phi^{opt} + \frac{M_1 + M_2 + M_3}{V} + \frac{(1-\eta)}{\eta(1-\eta^T)} \frac{L(\mathbf{Q}[0])}{NV}
\end{aligned}$$

and the proposition follows by taking the limit as  $N \rightarrow \infty$ .

## REFERENCES

- [1] I. Hwang, B. Song, and S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 20–27, Jun. 2013.
- [2] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.
- [3] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.
- [4] J. Zhang, R. Chen, J. Andrews, A. Ghosh, and R. W. Heath, "Networked MIMO with clustered linear precoding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1910–1921, Apr. 2009.
- [5] C. Ng and H. Huang, "Linear precoding in cooperative MIMO cellular networks with limited coordination clusters," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1446–1454, Dec. 2010.
- [6] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 56–61, Jun. 2011.
- [7] Huawei, Mobile networks go green. [Online]. Available: <http://www1.huawei.com/en/about-huawei/publications/communicate/hw-082734.htm>.
- [8] D. Li, W. Saad, I. Guvenc and A. Mehdodniya, "Decentralized energy allocation for wireless networks with renewable energy powered base stations," *IEEE Trans. on Commun.*, vol. 63, no. 6, pp. 2126–2142, Jun. 2015.
- [9] 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.
- [10] X. Liu and W. Xu, "Economic load dispatch constrained by wind power availability: A here-and-now approach," *IEEE Trans. Sustain. Energy*, vol. 1, no. 1, pp. 2–9, Apr. 2010.
- [11] G. B. Giannakis, V. Kekatos, N. Gatsis, S. Kim, H. Zhu, and B. Wollenberg, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 107–128, Sept. 2013.
- [12] J. Xu, Y. Guo, and R. Zhang, "CoMP meets energy harvesting: A new communication and energy cooperation paradigm," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2476–2488, Jun. 2015.
- [13] J. Xu and R. Zhang, "Cooperative energy trading in CoMP systems powered by smart grids," *IEEE Trans. Veh. Technol.*, to appear, 2016.
- [14] X. Wang, Y. Zhang, G. B. Giannakis, and S. Hu, "Robust smart-grid powered cooperative multipoint systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6188–6199, Nov. 2015.
- [15] 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. 6, Jun. 2016.
- [16] R. Uргаonkar, B. Uргаonkar, M. Neely, and A. Sivasubramaniam, "Optimal power cost management using stored energy in data centers," in *Proc. of ACM SIGMETRICS*, pp. 221–232, San Jose, CA, June 2011.
- [17] S. Lakshminaryana, H. V. Poor, and T. Quek, "Cooperation and storage trade-offs in power grids with renewable energy resources," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1–12, Jul. 2014.
- [18] S. Sun, M. Dong, and B. Liang, "Real-time power balancing in electric grids with distributed storage," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 6, pp. 1167–1181, Dec. 2014.
- [19] Y. Yao, L. Huang, A. Sharma, L. Golubchik, and M. Neely, "Data centers power reduction: A two time scale approach for delay tolerant workloads," in *Proc. of IEEE INFOCOM*, pp. 1431–1439, Orlando, FL, Mar. 2012.



- [20] W. Deng, F. Liu, H. Jin, C. Wu, and X. Liu, "Multigreen: Cost-minimizing multi-source datacenter power supply with online control," in *Proc. of ACM e-Energy*, pp. 149–160, Berkeley, CA, May 2013.
- [21] N. Prasad, M. Arslan, and S. Rangarajan, "A two time scale approach for coordinated multi-point transmission and reception over practical backhaul," in *Proc. of COMSNETS*, pp. 1–8, Bangalore, Jan. 2014.
- [22] H. Yu, M. Cheung, L. Huang and J. Huang, "Power-delay tradeoff with predictive scheduling in integrated cellular and Wi-Fi networks," *IEEE J. Sel. Areas Commun.*, to appear 2016.
- [23] J. Qin, Y. Chow, J. Yang, and R. Rajagopal, "Online modified greedy algorithm for storage control under uncertainty," *IEEE Trans. Power Syst.*, to appear 2016, [Online]. Available: <http://arxiv.org/pdf/1405.7789v4.pdf>.
- [24] "Pennsylvania-New Jersey-Maryland interconnection (PJM) hourly prices" [Online]. Available: <http://www.pjm.com/markets-and-operations/energy/realtime/monthlylmp.aspx>.
- [25] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [26] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," Sep. 2012. [Online]. Available: <http://cvxr.com/cvx>
- [27] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [28] H. Robbins and S. Monro, "A stochastic approximation method," *Annals of Mathematical Statistics*, pp. 400-407, 1951.
- [29] D. P. Bertsekas, *Convex Optimization Theory*, Athena Scientific, 2009.
- [30] L. Huang, "Fast-convergent learning-aided control in energy harvesting networks," preprint, available: <http://arxiv.org/abs/1503.05665>.
- [31] M. Neely, *Stochastic Network Optimization with Applications to Communication and Queueing Systems*, Morgan & Claypool, 2010.
- [32] M. Neely, "Stability and capacity regions for discrete time queueing networks," ArXiv Technical Report: arXiv:1003.3396v1, Mar. 2010.