Control of Wireless Networks with Rechargeable Batteries

Marios Gatzianas, Leonidas Georgiadis, *Senior Member, IEEE*, and Leandros Tassiulas, *Fellow, IEEE*

***Abstract*—We consider the problem of cross-layer resource allocation for wireless networks operating with rechargeable batteries under general arrival, channel state and recharge processes. The objective is to maximize total system utility, defined as a function of the long-term rate achieved per link, while satisfying energy and power constraints. A policy with decoupled admission control and power allocation decisions is proposed that achieves asymptotic optimality for sufficiently large battery capacity to maximum transmission power ratio (explicit bounds are provided). We present first a downlink resource allocation scenario; the analysis is then extended to multihop networks. The policy is evaluated via simulations and is seen to perform very well even in the non-asymptotic regime. This policy is particularly suitable for sensor networks, which typically satisfy the asymptotic conditions required by our methodology.**

***Index Terms*—Stochastic optimal control, Lyapunov drift, rechargeable batteries, sensor networks.**

I. Introduction

Efficient energy management is a crucial component of wireless network design, since it can lead to increased throughput and network lifetime. The latter concept is meaningful for battery-operated devices that do not have energy-harvesting capabilities and, hence, become inactive once they run out of energy. On the other hand, there exist applications where the wireless transmitters can replenish their batteries, two common examples being solar-paneled satellites and sensor networks [1]. Such rechargeable systems are usually regarded as having practically infinite lifetime [2], so that long-term performance metrics become appropriate.

We initially consider control of an L user downlink operating in discrete time under rechargeable batteries and later generalize the analysis to multihop networks. The relevant literature has greatly expanded in recent years, with most works being based on a dynamic programming (DP) and/or Markov decision process (MDP) approach. Finite horizon control problems are studied in [3], for a single satellite downlink subject to stochastic power demands and rewards (however, no packet dynamics are included), and [4], for a multihop network where the nodes have knowledge of the future short-term recharge process (an assumption we dispense with in this work). The model in [4] is inspired by the virtual circuit concept and, as such, performs resource allocation on a one-shot basis for each accepted service request (the latter is defined by its source and destination nodes as well as the associated gained revenue). Specifically, a policy is developed that, for each accepted request, computes an appropriate energy-weighted shortest path from the source to the destination node of the request and *simultaneously* reduces the energy of all nodes lying on the selected path by the required cost (provided that the gained revenue is larger than the expended cost). Since the policy drops any requests that cannot be immediately served, queueing effects are ignored. Additionally, the link-based energy costs do not depend on channel variations.

In [5], a rechargeable group of cells under realistic battery fatigue is examined, where the objective is the maximization of the energy delivered from the battery. Open and closed loop policies for finite and infinite horizon control also appear in [6]. Especially for sensor networks, [7] examines a scenario where the derived utility (namely, probability of event detection) depends only on the number of active sensors. The sensors are either recharging or transmitting (but not both) and, once completely drained of energy, can only activate themselves when fully recharged. Distributed threshold policies are proposed with a multiplicative guaranteed bound of 1/2 w. r. t. the optimal policy. In [8], the sensors can be activated even when partially recharged and an asymptotically optimal (w. r. t. battery capacity) policy is proposed for Poisson and exponential recharge/discharge processes, respectively. Finally, [9] focuses on the temporal correlation between the events sensed by a single sensor and studies, in terms of an MDP, the structure of the optimal policy under various observability conditions. Efficient suboptimal policies are also proposed and numerically evaluated for multiple node networks in [10]. References [7]–[10] focus exclusively on the sensing aspect of sensor networks rather than the network flow of information (i. e. what happens *after* an event is detected), which is the main focus of this paper.

Our model is distinct from the aforementioned works and is directly influenced by the cross-layer stochastic optimization framework of [11], [12], which, in turn, was inspired by [13]. This framework was applied in [14], which proposed optimal adaptive backpressure policies (collectively referred to as ABP) for non-rechargeable-battery networks under long-

Manuscript received July 11, 2008; revised February 11, 2009 and October 7, 2009; accepted October 19, 2009. The associate editor coordinating the review of this paper and approving it for publication was Y. Fang.

This research was supported by the GSRT project #05NON-EU-160, “Resource Allocation Techniques for Efficient Control and Management in Wireless Networks”, as well as by the European Commission through STREP project NET-REFOUND (FP6-IST-034413) and the ARO through grant W911NF-04-1-0306. Part of this work was presented at the IWCMC 2008 Conference, August 6-8, Crete, Greece.

M. Gatzianas and L. Georgiadis are with the Department of Electrical and Computer Engineering, Division of Telecommunications, Aristotle University of Thessaloniki, Thessaloniki, 54 124, Greece (e-mail: {mgkatzia, leonid}@auth.gr).

L. Tassiulas is with the Computer Engineering and Telecommunications Department, University of Thessaly, Volos, 38 221, Greece (e-mail: leandros@uth.gr).

Digital Object Identifier 10.1109/TWC.2010.080903

term average power constraints, and introduced the concept of virtual queues to handle the latter, while [15] addressed fairness issues in a similar setting. The above methodology relies crucially on the assumption that available controls depend only on current channel states and are independent of prior history, and this condition is violated in our model, as will become apparent later. Hence, this approach is not directly applicable and a new one is required. Our rechargeable model takes into account wireless channel variations and packet dynamics and aims in maximizing a suitably defined measure of user satisfaction (in the form of a utility function) while imposing minimal assumptions on the arrival, recharge and channel processes. Since the lifetime of rechargeable systems is practically infinite, the inclusion of packet dynamics forces us to study the stability properties of any policy under consideration. An adaptive stabilizing policy is proposed with guaranteed performance bounds that imply asymptotic optimality for sufficiently large battery capacity.

An outline of the results in this paper is as follows. Initially, the rechargeable battery model is related to a model of an infinite-capacity (non-rechargeable) battery operating under an average power constraint, and the utility of the optimal policy in the latter model is shown to be an upper bound to *any* policy (including the optimal one) of the former. Since the optimal utility in the infinite battery scenario can be achieved arbitrarily close by a well-known ABP-based policy, we state the intuitive claim that a carefully crafted modification of this policy will perform nearly optimally (when applied to the rechargeable battery) provided that the battery level rarely falls below a certain threshold. The bulk of the paper is devoted to rigorously proving the above claim and analytically extracting guaranteed performance bounds for the modified policy (a crucial point in the above derivation is the construction of appropriately connected virtual queues and the proof that all queues remain finite under the application of the new policy). The original results in this paper are presented in Theorem 1, Corollary 1 of Section IV and Corollary 2 of Section V for single-hop and multihop networks, respectively.

The paper is structured as follows. Section II contains the system model and problem statement while Section III presents the proposed downlink policy, whose performance and stability properties are analytically determined in Section IV. Section V describes an extension to multihop networks and presents all necessary results needed to guarantee asymptotic optimality. Due to space restrictions, the proofs of the results in Section V are presented in [16] instead. Section VI contains the numerical evaluation of the policy while Section VII concludes the paper and offers directions for future research.

II. System Model and Problem Statement

Although the proposed policy is applicable to generic multihop networks under arbitrary rate-convergent [12] arrival, channel state and (energy) recharge processes, in order to simplify the discussion and gain valuable insight into the operation of the policy, we initially present the analysis for a downlink scenario with iid arrival, channel and recharge processes. As mentioned in Section I, the downlink model is applicable to satellite downlinks where the satellite is powered by rechargeable panels. In Section V, we describe the modifications needed for the general multihop case. Non-iid rate-convergent processes can be handled by the T-slot analysis of [12] and are not covered here.

We consider a time-slotted system where slot t corresponds to the time interval $[t \quad t+1)$. There exists a single base node that transmits to L users (i. e. there exist L wireless downlinks). The link channels are time-varying so that we denote with $\boldsymbol{S}(t)$ the channel state at slot t and assume it is iid distributed over a finite set $\mathcal{S}$, i. e. $\boldsymbol{S}(t) \in \mathcal{S}$ for all t. We also define $\pi_{\boldsymbol{s}} \triangleq \Pr(\boldsymbol{S}(t) = \boldsymbol{s})$. Channel conditions remain constant for the duration of each slot but may change at slot boundaries. The base node has a distinct traffic stream for each user so that $A_l(t)$ stands for the number of exogenously generated bits during slot t destined for user l. These bits enter the transport layer and are stored, in a FIFO manner, in an external queue awaiting admission into the network layer and subsequent transmission. Admission control is necessary for the case of heavy traffic that exceeds the network capacity. When the bits are admitted into the network layer, they are stored in an internal "network" queue, again in FIFO fashion, until transmission occurs. Let $V_l(t), U_l(t)$ be the number of bits stored at time t in the external/internal queue, respectively, of the base node and destined for user l (through link l). Denote with $R_{l,in}(t)$ the number of bits moved from the external to the internal queue (i. e. admitted into the network) of link l at time t. Supply/demand constraints clearly require $R_{l,in}(t) \leq V_l(t) + A_l(t)$ for all l, t. It is further assumed that each exogenous arrival process $A_l(t)$ has a deterministic upper bound $\hat{A}_l$ (i. e. $A_l(t) \leq \hat{A}_l \quad \forall t$) and is iid distributed with expected value λ_l.

The base node is equipped with a rechargeable battery of maximum energy E_{max}. The battery level at time t is denoted as $E(t)$. The battery energy is depleted due to link transmissions but is also replenished due to a recharge process. Specifically, we denote with $B(t)$ the amount of replenished energy during slot t, where $B(t)$ is assumed to be bounded ($B(t) \leq \hat{B} \quad \forall t$) and iid distributed with an expected value of $\bar{B}$. At the beginning of each time slot t, the network controller chooses a power vector $\boldsymbol{P}(t) \triangleq (P_1(t), \ldots, P_L(t))$, where $P_l(t)$ is the selected transmission power in link l during slot t. Available transmission powers may be channel-state dependent, i. e. $\boldsymbol{P}(t) \in \mathcal{P}_{\boldsymbol{S}}$ whenever $\boldsymbol{S}(t) = \boldsymbol{s}$, where $\mathcal{P}_{\boldsymbol{S}}$ is a finite set[1] for all $\boldsymbol{s} \in \mathcal{S}$. The only constraints imposed on $\mathcal{P}_{\boldsymbol{S}}$ are that, for all $\boldsymbol{s} \in \mathcal{S}$, it holds $\mathbf{0} \in \mathcal{P}_{\boldsymbol{S}}$ and $\sum_{l=1}^{L} p_l \leq \hat{P} \quad \forall \boldsymbol{p} \in \mathcal{P}_{\boldsymbol{S}}$ (with $\hat{P} \leq E_{max}$). The first constraint states that the base node may always choose to decline transmission while the second one models hardware limitations or standard regulations. The existence of the finite energy battery further restricts the available power vectors for slot t by the natural condition $\sum_{l=1}^{L} p_l \leq E(t) \quad \forall \boldsymbol{p} \in \mathcal{P}_{\boldsymbol{S}(t)}$.

To facilitate analysis, it is assumed that all values of the recharge process $B(t)$ as well as all members of $\mathcal{P}_{\boldsymbol{S}}$ are integer multiples of an arbitrary constant (i. e. they are quantized), so

[1]although in our case it holds $\mathcal{P}_{\boldsymbol{S}} = \mathcal{P} \quad \forall \boldsymbol{s} \in \mathcal{S}$, since there is no reason for available power choices to depend on the channel state, there exist scenarios where the available powers depend on a properly defined state space (which may include more components than channel conditions only). These cases also fall under the considered model.

that $E(t)$ effectively takes values in a countable set. For a given state $\boldsymbol{S}(t)$ and selected power $\boldsymbol{P}(t)$, the transmission rate $\boldsymbol{\mu}(t) \triangleq (\mu_1(t), \ldots, \mu_L(t))$ in slot t (i. e. the number of bits that can be transmitted in each link) is upper bounded, component-wise, by a vector function $\boldsymbol{c}\left(\boldsymbol{S}(t), \boldsymbol{P}(t)\right)$ with the following properties

- it holds $c_l\left(\boldsymbol{s}, \boldsymbol{p}\right) = 0$ for any $\boldsymbol{s}, \boldsymbol{p}$ such that $p_l = 0$. A consequence of the previous statement is the fact that $\boldsymbol{c}\left(\boldsymbol{s}, \boldsymbol{0}\right) = \boldsymbol{0}$.
- for any $\boldsymbol{s}, \boldsymbol{p}$, and for all l, it holds $c_l\left(\boldsymbol{s}, \boldsymbol{p}\right) \leq c_l\left(\boldsymbol{s}, \tilde{\boldsymbol{p}}\right)$ where $\tilde{p}_k = \delta_{lk} p_k$ and δ_{lk} is Kronecker's delta. The function $c_l\left(\boldsymbol{s}, \tilde{\boldsymbol{p}}\right)$ is non-decreasing, differentiable and concave w. r. t. p_l. Interpreting the above inequality, it is equivalent to saying that the maximum rate of link l decreases (for a fixed p_l) as the other links are assigned non-zero power, which follows from standard interference properties. The second condition is also typical of most rate functions and appears often in literature.
- it holds $\frac{\partial c_l}{\partial p_l}\left(\boldsymbol{s}, \boldsymbol{0}\right) < \infty$ for all l and $\boldsymbol{s}$.

Since all sets encountered so far are finite, we define the bound[2] $\hat{c}_l \triangleq \max_{s \in \mathcal{S}, \boldsymbol{p} \in \mathcal{P}_{\boldsymbol{S}}} c_l(\boldsymbol{s}, \boldsymbol{p})$.

Under the previous assumptions, the queues $\boldsymbol{V}(t), \boldsymbol{U}(t)$ and the battery level $E(t)$ evolve as

$$V_l(t+1) = V_l(t) + A_l(t) - R_{l,in}(t), \tag{1}$$

$$U_l(t+1) = \left[U_l(t) - \mu_l\left(\boldsymbol{S}(t), \boldsymbol{P}(t)\right)\right]^+ + R_{l,in}(t), \tag{2}$$

$$E(t+1) = \min\left(E(t) - \sum_{l=1}^{L} P_l(t) + B(t), E_{max}\right), \tag{3}$$

where $[x]^+ \triangleq \max(x, 0)$, subject to constraints

$$R_{l,in}(t) \leq V_l(t) + A_l(t), \quad \forall\, l,\, t, \tag{4}$$

$$\boldsymbol{P}(t) \in \mathcal{P}_{\boldsymbol{S}(t)}, \quad \sum_{l=1}^{L} P_l(t) \leq \min\left(E(t), \hat{P}\right) \quad \forall\, t, \tag{5}$$

$$\mu_l\left(\boldsymbol{S}(t), \boldsymbol{P}(t)\right) \leq c_l\left(\boldsymbol{S}(t), \boldsymbol{P}(t)\right) \quad \forall\, l,\, t. \tag{6}$$

It can be argued that $\mu_l\left(\boldsymbol{S}(t), \boldsymbol{P}(t)\right)$ should be replaced by $c_l\left(\boldsymbol{S}(t), \boldsymbol{P}(t)\right)$ in (2) since, intuitively, there is no benefit in transmitting fewer bits than those allowed by the allocated power. Although this argument is indeed correct, we choose to keep the current notation and reach the same conclusion through the policy specification itself rather than mere intuition.

Denote with $R_l(t)$ the number of bits *actually* transmitted in link l during slot t so that the corresponding time-average rate up to t is $\bar{r}_l(t) \triangleq \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[R_l(\tau)]$. Each user l derives a satisfaction $f_l\left(\bar{r}_l(t)\right)$ based on its current $\bar{r}_l(t)$, where the utility function $f_l(\cdot)$ is assumed to be non-negative, increasing, differentiable and concave. We also impose the constraints $f_l(0) = 0$, $f_l'(0) < \infty$ (see [16] for the modification required to handle the case $f_l'(0) = \infty$). The total system satisfaction $g\left(\bar{\boldsymbol{r}}(t)\right)$ at time t is $g\left(\bar{\boldsymbol{r}}(t)\right) \triangleq \sum_{l=1}^{L} f_l\left(\bar{r}_l(t)\right)$. We also assume that all bit queues are empty at $t = 0$. No assumption is made regarding the initial battery level, i. e. $E(0)$ can take any value in the set $[0 \quad E_{max}]$. A decision rule is a procedure for selecting the control variables $\boldsymbol{R}_{in}(t), \boldsymbol{P}(t)$ for a specific time slot t subject to all aforementioned constraints, while a policy is a sequence of decision rules for all time slots. We restrict attention to policies that stabilize the network according to the definition of [11]. Specifically, a network queue $\boldsymbol{U}(t)$ is stable iff $\limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}[U_l(\tau)] < \infty \quad \forall\, l$. Hence, we state our problem as

Problem 1: For the downlink scenario described by the queuing evolutions of (1)–(3) and the constraints of (4)–(6), find a stabilizing policy that maximizes $\liminf_{t\to\infty} g\left(\bar{\boldsymbol{r}}(t)\right)$. The optimization is performed over the set of all stabilizing policies, including those with perfect knowledge of future events. Denote the optimal objective value as g_{re}^*.

Determination of the optimal policy in Problem 1 is very challenging, so we instead seek a policy with a system utility that is close to optimal. Since $\hat{P}$, $\hat{B}$ are considered to be system parameters, we can fix their values and restate the problem as

Problem 2: Under the same assumptions as in Problem 1, find a policy that for any $\epsilon > 0$ achieves an objective value no less than $g_{re}^* - \epsilon$ for a sufficiently large E_{max}, i. e. for $E_{max} > E^*$ where E^* depends implicitly on ϵ (the notation $E^*(\epsilon)$ will be used to emphasize the latter fact).

Problems 1, 2 can, in principle, be attacked with dynamic programming (DP) and/or Markov chain techniques. However, these techniques are impractical since they suffer from the dimensionality curse of DP and require extensive knowledge of system parameters, e. g. $\pi_{\boldsymbol{S}}$, $\boldsymbol{\lambda}$, which may not be available to the network controller. On the other hand, the Lyapunov drift framework of [12]–[15] assumes that, at any state $\boldsymbol{s}$, one may always choose any of the available controls in $\mathcal{P}_{\boldsymbol{S}}$. However, this is not the case in our model as evidenced from (5), which implies that, when $E(t) < \hat{P}$, the available powers $\boldsymbol{P}(t)$ depend explicitly on the battery level $E(t)$. Hence, this approach is not directly applicable to our problem. Nevertheless, it hints at the existence of a modified policy that solves Problem 2 when $E_{max} \gg \hat{B}, \hat{P}$. The intuition behind the last statement is explained below.

A. Some intuitive remarks

Conservation of energy, combined with (3), (5), yields

$$\begin{aligned} &\sum_{\tau=0}^{t-1}\sum_{l=1}^{L} P_l(\tau) \leq E_{max} + \sum_{\tau=0}^{t-1} B(\tau) \\ &\Rightarrow \limsup_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left[\sum_{l=1}^{L} P_l(\tau)\right] \leq \bar{B}. \end{aligned} \tag{7}$$

Eq. (7) implies that *any* policy (stabilizing or not) acting on the rechargeable battery satisfies an average power constraint of $\bar{B}$, which, in turn, allows us to perform the following "thought experiment". Consider the same downlink problem as above and replace the rechargeable battery with an infinite capacity battery, which essentially removes (3) and the $E(t)$ constraint in (5). Denote with g_{avg}^* the maximum utility achieved over all policies acting on the infinite battery under an average power constraint $\bar{B}$ (we collectively refer to these as "average"

[2] unless otherwise noted, the accent ^ will always denote an upper bound. Also, the indices $l, \boldsymbol{s}, \boldsymbol{p}$ will hereafter range, respectively, over the sets $\{1, \ldots, L\}, \mathcal{S}, \mathcal{P}_{\boldsymbol{S}}$ (e. g. we write $\forall\, \boldsymbol{s}$ instead of the formally correct $\forall\, \boldsymbol{s} \in \mathcal{S}$, etc.).

policies). This scenario is treatable by the methodology in [14], which proposes an ABP-based policy whose performance can approach g_{avg}^* arbitrarily close. Since it clearly holds $g_{re} \leq g_{avg}^*$, Problem 2 is solved if we can find a policy (for the rechargeable battery) that performs better than $g_{avg}^* - \epsilon$. Since a rechargeable policy that mimics ABP "most of the time" (i. e. for a given channel state, it chooses most of the time the same, or almost the same, transmission power that ABP selects) is expected to perform close to g_{avg}^*, we need to identify any mechanisms through which a rechargeable policy's performance is degraded w. r. t. ABP performance and propose a policy that mitigates these mechanisms.

A policy acting on the rechargeable battery may suffer a performance loss w. r. t. the optimal ABP-based average policy (the one achieving g_{avg}^*) through the following mechanisms. It may happen that for some slot t it holds $E(t) < \hat{P}$, so that the set of available controls $\boldsymbol{P}(t)$ for the rechargeable policy is a proper subset of the corresponding set for the average policy (i. e. the battery does not currently have sufficient energy to fully exploit favorable channel states). At another slot, it may happen that $B(t)$ is large enough to drive the battery over the E_{max} cap, which implies that some energy is permanently lost and cannot be used for future transmissions (the ABP average policy is immune to this effect since there is no battery to overflow in the first place). In both cases, the performance of the rechargeable policy deviates from the optimal value g_{avg}^* by an amount that is not known *a priori* and is uncertain (at this point) whether it can be made arbitrarily small.

The previous observations indicate that a policy closely resembling ABP when $E(t) \geq \hat{P}$ and under which $\Pr(E(t) < \hat{P})$, $\Pr(E_{max} - \hat{B} < E(t) \leq E_{max})$ are small should perform nearly optimally. This is quantified next.

III. Policy Description

We now shift our attention to the rechargeable battery setting and, in the spirit of [11], introduce a virtual queue for each linear long-term constraint. Two such constraints exist in our model, the first one due to the finite arrivals, as modeled in (1), and the second one due to (7). Hence, we define virtual queues $\boldsymbol{Y}(t) \triangleq (Y_1(t), \ldots, Y_L(t))$ and $D(t)$ which evolve as

$$Y_l(t+1) = [Y_l(t) - R_{l,in}(t)]^+ + \gamma_l(t), \tag{8}$$

$$D(t+1) = [D(t) - (1-\delta)B(t)]^+ + \sum_{l=1}^{L} P_l(t), \tag{9}$$

for some $0 < \delta < 1$, where $\gamma_l(t)$ is an auxiliary process introduced for mathematical convenience. Since $\boldsymbol{R}_{in}(t), \boldsymbol{\gamma}(t)$ are determined by the policy, we can impose arbitrary bounds on them. Specifically, we require $0 \leq R_{l,in}(t), \gamma_l(t) \leq \hat{A}_l \; \forall l$. The virtual queues are constructed in such a way that any policy that stabilizes them also satisfies the appropriate long-term constraints (this novel insight was introduced in [14]). Eq. (9) is different from the standard approach of [14], which would replace $(1-\delta)B(t)$ with $\bar{B}$, since the latter is the actual constraint to be satisfied. This modification is an essential ingredient of our approach, as will become apparent later. Fig. 1 presents the interconnections between the various physical and virtual queues, as defined by their evolutions in (1)–(3) and (8), (9).

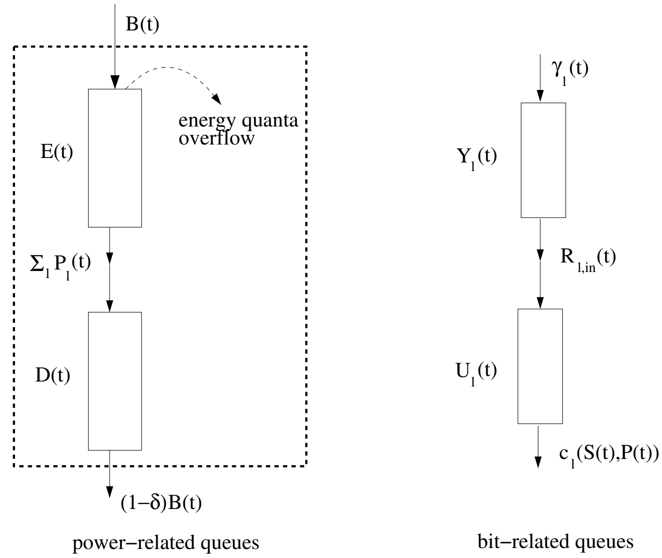


Fig. 1. Queue connections in the downlink scenario with rechargeable battery.

To assist the reader, the policy description is given first, followed by the stability and performance analysis in later Sections. The following policy is similar to the ABP-based policy of [11] for non-rechargeable batteries, with a small but crucial difference that will be noted.

Downlink Rechargeable Adaptive Backpressure Policy (DRABP)

1) at the beginning of slot t, observe queues $\boldsymbol{U}(t), \boldsymbol{Y}(t), \boldsymbol{V}(t)$ and select $\boldsymbol{R}_{in}(t)$ bits for admission into the network layer according to

$$\boldsymbol{R}_{in}(t) = \arg\max \sum_{l=1}^{L} (Y_l(t) - U_l(t))\, r_l$$
$$\text{s.t.} \quad 0 \leq r_l \leq \min\left(V_l(t) + A_l(t), \hat{A}_l\right), \quad \forall\, l, \tag{10}$$

2) select $\boldsymbol{\gamma}(t)$ as the solution to the following problem

$$\boldsymbol{\gamma}(t) = \arg\max \sum_{l=1}^{L} [M f_l(\gamma_l) - Y_l(t)\gamma_l]$$
$$\text{s.t.} \quad 0 \leq \gamma_l \leq \hat{A}_l \quad \forall\, l, \tag{11}$$

3) observe $\boldsymbol{S}(t)$ and select $\boldsymbol{P}(t)$ according to

$$\boldsymbol{P}(t) = \arg\max \sum_{l=1}^{L} [U_l(t)\mu_l(\boldsymbol{S}(t), \boldsymbol{p}) - D(t)p_l]$$
$$= \arg\max \sum_{l=1}^{L} [U_l(t)c_l(\boldsymbol{S}(t), \boldsymbol{p}) - D(t)p_l]$$
$$\text{s.t.} \quad \boldsymbol{p} \in \mathcal{P}_{\boldsymbol{S}(t)}, \quad \sum_{l=1}^{L} p_l \leq \min\left(E(t), \hat{P}\right),$$
$$0 \leq \mu_l(\boldsymbol{S}(t), \boldsymbol{p}) \leq c_l(\boldsymbol{S}(t), \boldsymbol{p}) \quad \forall\, l, \tag{12}$$

4) update queues $\boldsymbol{V}(t), \boldsymbol{Y}(t), \boldsymbol{U}(t), E(t), D(t)$, in that order, according to the appropriate evolution.

The previous policy will be denoted as DRABP(δ, M) to emphasize its dependence on these two parameters (where

$M > 0$). Its difference from the policy of [11] lies in the presence of the term $E(t)$ in the constraint of (12). As mentioned in Section II-A, this term causes the performance analysis of [11] to be inapplicable to our model. Hence, a performance bound for DRABP is still required, and this will be pursued next.

As a final note, (10) accepts a typical "bang-bang" solution (specifically, $R_{l,in}(t) = 0$ when $Y_l(t) < U_l(t)$), while the concave maximization problem in (11) is separable and has a closed form solution if $(f_l')^{-1}(\gamma_l)$ is analytically known (the $^{-1}$ superscript denotes functional inverse). Also, DRABP requires no knowledge of system parameters (e. g. $\pi_{\boldsymbol{S}}$, $\boldsymbol{\lambda}$) and the intuitive rule of transmitting with peak rate for a given power arises naturally through (12).

IV. Performance Analysis

This section contains the derivation of a lower bound for the performance of DRABP. The analysis consists of quite a few intermediate steps so that, for the reader's convenience, we first present a brief outline. Since DRABP closely resembles the ABP policy of [11], we initially state a known result (Lemma 1) that describes the latter's performance. The starting point for deriving a performance bound for DRABP is a modification of another known result based on the notion of Lyapunov drift (Lemma 2). An algebraic computation of the Lyapunov drift justifies, in retrospect, the selection of the steps (10)–(12) in DRABP and, combined with the fact that all queues are finitely bounded under DRABP (Lemma 3), leads to a simplified Lyapunov drift expression that contains the probability of the battery level falling below $\hat{P}$ (eq. (24)). Using queueing-theoretic techniques, an upper bound for the latter probability is computed in Lemma 5 and combined with the simplified Lyapunov drift expression to produce the final DRABP performance bound of Theorem 1. More details are provided below.

A. Notation and preliminary results

For $0 \leq \delta < 1$, we define the set

$$\mathcal{R}_\delta \triangleq \Bigg\{ \boldsymbol{r} : \exists \pi_{\boldsymbol{p}}^{\boldsymbol{s}} \geq 0 \text{ s.t. } \sum_{\boldsymbol{s},\boldsymbol{p},l} p_l \pi_{\boldsymbol{p}}^{\boldsymbol{s}} \pi_{\boldsymbol{s}} \leq (1-\delta)\bar{B}, \quad 0 \leq r_l \leq \sum_{\boldsymbol{s},\boldsymbol{p}} c_l(\boldsymbol{s},\boldsymbol{p})\, \pi_{\boldsymbol{p}}^{\boldsymbol{s}} \pi_{\boldsymbol{s}} \;\forall l, \; \sum_{\boldsymbol{p}} \pi_{\boldsymbol{p}}^{\boldsymbol{s}} = 1 \;\forall \boldsymbol{s} \Bigg\}, \tag{13}$$

as the set of rates that can be stabilized by a randomized policy subject to an average power constraint of $(1-\delta)\bar{B}$. The quantity $\pi_{\boldsymbol{p}}^{\boldsymbol{s}}$ (which is a discrete pdf on the set $\mathcal{P}_{\boldsymbol{S}}$) is interpreted in [11] as the probability with which a randomized policy selects power $\boldsymbol{p}$ when the current channel state is $\boldsymbol{s}$. We also define $g_\delta^* \triangleq \max_{\boldsymbol{r}\in\Lambda_\delta} \sum_{l=1}^L f_l(r_l)$, where $\Lambda_\delta = \{\boldsymbol{r} \in \mathcal{R}_\delta : \; \boldsymbol{0} \leq \boldsymbol{r} \leq \boldsymbol{\lambda}\}$. The following Lemma, which is proved in [11], characterizes the optimal solution to the infinite battery model with average power constraint and, as stated in Section II-A, provides an upper bound for *any* rechargeable policy.

Lemma 1: No stabilizing policy acting on an infinite capacity battery with an average power constraint of $\bar{B}$ can achieve an objective value larger than $g_{avg}^* = g_0^*$. It also holds $\lim_{\delta\downarrow 0} g_\delta^* = g_0^*$.

The starting point for deriving a performance bound for any rechargeable policy is the following variation of a known result [11, Lemma 5.3 and Theorem 5.4]

Lemma 2: Consider a concave function $g(\cdot)$ and stochastic processes $Q(t)$, $\boldsymbol{X}(t)$, $\boldsymbol{Z}(t)$. If there exists a function $Ly(\boldsymbol{X}(t)) \geq 0$ (referred to as a Lyapunov function) such that $\mathbb{E}[Ly(\boldsymbol{X}(0))] < \infty$ and the following relation is satisfied for all t and some arbitrary constants $C, M, g^* > 0$

$$\Delta(\boldsymbol{X}(t)) - M\mathbb{E}\left[g(\boldsymbol{Z}(t))|\, \boldsymbol{X}(t)\right] \leq C - Mg^* + Q(t), \tag{14}$$

it then holds $\liminf_{t\to\infty} g(\bar{\boldsymbol{z}}(t)) \geq g^* - (C+\bar{Q})/M$, where $\bar{\boldsymbol{z}}(t)$ is the corresponding time-average of $\boldsymbol{Z}(t)$, $\bar{Q} \triangleq \limsup_{t\to\infty} \frac{1}{t}\sum_{j=0}^{t-1} \mathbb{E}[Q(j)]$ and $\Delta(\boldsymbol{X}(t)) \triangleq \mathbb{E}\left[Ly(\boldsymbol{X}(t+1)) - Ly(\boldsymbol{X}(t))\,|\boldsymbol{X}(t)\right]$.

Denote the composite queue $\boldsymbol{X}(t) \triangleq (\boldsymbol{U}(t), \boldsymbol{Y}(t), D(t))$ and set $Ly(\boldsymbol{X}(t)) \triangleq \sum_{l=1}^L \left(U_l^2(t) + Y_l^2(t)\right) + D^2(t)$. Squaring (2), (8), (9), applying Lemma 4.3 of [11], summing the resulting inequalities over l and subtracting the term $2M\mathbb{E}\left[\sum_l f_l(\gamma_l(t))|\, \boldsymbol{X}(t)\right]$ from both sides yields after some algebra (see [16] for details)

$$\begin{aligned} &\Delta(\boldsymbol{X}(t)) - 2M\,\mathbb{E}\left[\sum_{l=1}^L f_l(\gamma_l(t))\middle|\, \boldsymbol{X}(t)\right] \leq \dot{C} \\ &- 2\,\mathbb{E}\left[\sum_{l=1}^L \left[U_l(t)\mu_l(\boldsymbol{S}(t), \boldsymbol{P}(t)) - D(t)P_l(t)\right]\middle|\, \boldsymbol{X}(t)\right] \\ &- 2\,\mathbb{E}\left[\sum_{l=1}^L \left[Y_l(t) - U_l(t)\right] R_{l,in}(t)\middle|\, \boldsymbol{X}(t)\right] - 2(1-\delta)\bar{B}D(t) \\ &- 2\,\mathbb{E}\left[\sum_{l=1}^L \left[Mf_l(\gamma_l(t)) - Y_l(t)\gamma_l(t)\right]\middle|\, \boldsymbol{X}(t)\right], \end{aligned} \tag{15}$$

where $\dot{C} = \hat{B}^2 + \hat{P}^2 + \sum_{l=1}^L \left(\hat{c}_l^2 + 3\hat{A}_l^2\right)$. The rationale behind the DRABP specification is now clear; (10)–(12) are selected so as to minimize the corresponding term in the RHS of (15). Our intention is to bring (15) into the form of (14), so that a DRABP performance bound can be derived from Lemma 2.

B. Stability and performance properties of DRABP

The next result, proved in [16], provides all necessary information for the stability properties of DRABP.

Lemma 3: Under DRABP, queues $\boldsymbol{Y}(t), \boldsymbol{U}(t), D(t)$ are deterministically bounded for all t as follows

$$\begin{aligned} Y_l(t) &\leq Vf_l'(0) + \hat{A}_l \triangleq \hat{Y}_l \quad \forall l, \\ U_l(t) &\leq \hat{Y}_l + \hat{A}_l \triangleq \hat{U}_l \quad \forall l, \\ D(t) &\leq \max_{1\leq l\leq L}\left(\hat{U}_l\hat{C}_l\right) + \hat{P} \triangleq \hat{D}, \end{aligned} \tag{16}$$

where $\hat{C}_l \triangleq \max_{\boldsymbol{s}\in\mathcal{S}} \frac{\partial c_l}{\partial p_l}(\boldsymbol{s}, \boldsymbol{0})$.

The following remarks will be useful. For a given $\boldsymbol{X}(t)$ (which implies that $\boldsymbol{U}(t), D(t)$ are known) and $\boldsymbol{S}(t) = \boldsymbol{s}$, the solution to the optimization in (12) does not depend on $E(t)$ when $E(t) \geq \hat{P}$. We denote this solution as $\boldsymbol{p}_{\boldsymbol{S}}^*$ and use the notation $\boldsymbol{P}(\boldsymbol{s}, e)$ for the solution to the same problem when $E(t) =$

$e < \hat{P}$. Since $\boldsymbol{p} = \mathbf{0}$ is always an allowable selection, it follows from optimality of (12)

$$\begin{aligned}&\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{P}\left(\boldsymbol{s},e\right)\right)-D(t)P_l\left(\boldsymbol{s},e\right)\right]\\&\geq\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\mathbf{0}\right)-D(t)\cdot 0\right]=0\quad\forall\,\boldsymbol{s},\forall\,e,\forall\,t.\end{aligned}\tag{17}$$

Let $\mathcal{E}(t)$ be the countable set of values that $E(t)$ may take under the DRABP policy. Define the joint probability $q_{\boldsymbol{s},e}(t) \triangleq \Pr\left(\boldsymbol{S}(t)=\boldsymbol{s},\, E(t)=e|\,\boldsymbol{X}(t)\right)$, when DRABP is applied. It now holds

$$\begin{aligned}&\mathbb{E}\left[\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right)-D(t)P_l(t)\right]\middle|\,\boldsymbol{X}(t)\right]\\
&=\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{e\in\mathcal{E}(t)}\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{P}(\boldsymbol{s},e)\right)-D(t)P_l\left(\boldsymbol{s},e\right)\right]q_{\boldsymbol{s},e}(t)\\
&\geq\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{\substack{e\in\mathcal{E}(t)\\e\geq\hat{P}}}\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}^*_{\boldsymbol{s}}\right)-D(t)p^*_{l,\boldsymbol{s}}\right]q_{\boldsymbol{s},e}(t)\\
&=\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{e\in\mathcal{E}(t)}\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}^*_{\boldsymbol{s}}\right)-D(t)p^*_{l,\boldsymbol{s}}\right]q_{\boldsymbol{s},e}(t)\\
&-\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{\substack{e\in\mathcal{E}(t)\\e<\hat{P}}}\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}^*_{\boldsymbol{s}}\right)-D(t)p^*_{l,\boldsymbol{s}}\right]q_{\boldsymbol{s},e}(t)\\
&\geq\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}^*_{\boldsymbol{s}}\right)-D(t)p^*_{l,\boldsymbol{s}}\right]\pi_{\boldsymbol{s}}\\
&-\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{l=1}^{L}U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}^*_{\boldsymbol{s}}\right))\Pr\left(E(t)<\hat{P},\,\boldsymbol{S}(t)=\boldsymbol{s}\middle|\,\boldsymbol{X}(t)\right)\\
&=\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{l=1}^{L}\left[U_l(t)c_l(\boldsymbol{s},\boldsymbol{p}^*_{\boldsymbol{s}})-D(t)p^*_{l,s}\right]\pi_{\boldsymbol{s}}\\
&-\sum_{l=1}^{L}\hat{c}_l\hat{U}_l\Pr\left(E(t)<\hat{P}\middle|\,\boldsymbol{X}(t)\right),\end{aligned}\tag{18}$$

where the first inequality is due to (17) and the final line follows from the bounds on $\boldsymbol{U}(t)$ and $c_l\left(\boldsymbol{s},\boldsymbol{p}\right)$. We also used the independence of $\boldsymbol{p}^*_{\boldsymbol{s}}$ from e (the value of $E(t)$) and the iid property to write $\sum_{e\in\mathcal{E}(t)} q_{\boldsymbol{s},e}(t) = \Pr\left(\boldsymbol{S}(t)=\boldsymbol{s}|\,\boldsymbol{X}(t)\right) = \pi_{\boldsymbol{s}}$ in the transition from the fourth to the sixth line. We now pick any pdf $\pi^{\boldsymbol{s}}_{\boldsymbol{p}}$ acting on the set $\mathcal{P}_{\boldsymbol{s}}$. We have, from optimality of (12),

$$\begin{aligned}&\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}^*_{\boldsymbol{s}}\right)-D_l(t)p^*_{l,\boldsymbol{s}}\right]\\&\geq\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}\right)-D(t)p_l\right]\end{aligned}\tag{19}$$

for all $t, \boldsymbol{s}, \boldsymbol{p}$ so that multiplying both sides of the above inequality with $\pi^{\boldsymbol{s}}_{\boldsymbol{p}}\pi_{\boldsymbol{s}}$, summing over $\boldsymbol{s}, \boldsymbol{p}$ and exploiting the pdf property $\sum_{\boldsymbol{p}\in\mathcal{P}_{\boldsymbol{s}}}\pi^{\boldsymbol{s}}_{\boldsymbol{p}} = 1$ yields

$$\begin{aligned}&\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}^*_{\boldsymbol{s}}\right)-D(t)p^*_{l,\boldsymbol{s}}\right]\pi_{\boldsymbol{s}}\\&\geq\sum_{\boldsymbol{s}\in\mathcal{S}}\sum_{\boldsymbol{p}\in\mathcal{P}_{\boldsymbol{s}}}\sum_{l=1}^{L}\left[U_l(t)c_l\left(\boldsymbol{s},\boldsymbol{p}\right)-D(t)p_l\right]\pi^{\boldsymbol{s}}_{\boldsymbol{p}}\pi_{\boldsymbol{s}}.\end{aligned}\tag{20}$$

Also, pick any $\boldsymbol{r}\in\Lambda_\delta$. It then holds $r_l \leq \lambda_l\ \forall\, l$, whence it follows $A_l(t)\,\frac{r_l}{\lambda_l} \leq A_l(t) \leq \min\left(V_l(t)+A_l(t),\hat{A}_l\right)$. Hence, $A_l(t)r_l/\lambda_l$ belongs to the constraint set of the optimization in (10), so that it holds for all t

$$\sum_{l=1}^{L}\left(Y_l(t)-U_l(t)\right)R_{l,in}(t)\geq\sum_{l=1}^{L}\left(Y_l(t)-U_l(t)\right)A_l(t)\,\frac{r_l}{\lambda_l}\tag{21}$$

where $R_{l,in}(t)$ is selected by the DRABP policy. The last expression contains the random variables $R_{l,in}(t), A_l(t)$ so that taking conditional expectations upon $\boldsymbol{X}(t)$ and using the iid nature of $A_l(t)$ produces

$$\mathbb{E}\left[\sum_{l=1}^{L}\left(Y_l(t)-U_l(t)\right)R_{l,in}(t)\middle|\,\boldsymbol{X}(t)\right]\geq\sum_{l=1}^{L}\left(Y_l(t)-U_l(t)\right)r_l\tag{22}$$

for any $\boldsymbol{r}\in\Lambda_\delta$, since $\mathbb{E}[A_l(t)|\boldsymbol{X}(t)] = \lambda_l$. Also, the above $\boldsymbol{r}$ satisfies $0 \leq r_l \leq \hat{A}_l$ so that from optimality of (11) we have for any $\boldsymbol{r}\in\Lambda_\delta$

$$\sum_{l=1}^{L}\left[Mf_l\left(\gamma_l(t)\right)-Y_l(t)\gamma_l(t)\right]\geq\sum_{l=1}^{L}\left[Mf_l(r_l)-Y_l(t)r_l\right]\tag{23}$$

where $\gamma_l(t)$, which is selected according to the DRABP policy, is a random variable. Clearly, a similar inequality is produced for the conditional expectation upon $\boldsymbol{X}(t)$ of the LHS of (23).

We set $\boldsymbol{r} = \boldsymbol{r}^*_\delta$ in (23), (22) and use its corresponding, through (13), pdf $\pi^{\boldsymbol{s}}_{\boldsymbol{p}}$ in (20). Inserting the resulting expressions into (15) produces after some manipulations (see [16] for details)

$$\begin{aligned}&\Delta\left(\boldsymbol{X}(t)\right)-2M\,\mathbb{E}\left[\sum_{l=1}^{L}f_l\left(\gamma_l(t)\right)\middle|\,\boldsymbol{X}(t)\right]\leq\dot{C}-2Mg^*_\delta\\&+2\sum_{l=1}^{L}\hat{c}_l\hat{U}_l\Pr\left(E(t)<\hat{P}\middle|\,\boldsymbol{X}(t)\right)\quad\forall\,t.\end{aligned}\tag{24}$$

We now invoke Lemma 2 to prove the following

Lemma 4: Under DRABP, it holds

$$\begin{aligned}\liminf_{t\to\infty}\,\sum_{l=1}^{L}f_l\left(\bar{\gamma}_l(t)\right)\geq&g^*_\delta-\frac{\dot{C}}{2M}-\frac{\sum_{l=1}^{L}\hat{c}_l\hat{U}_l}{M}\\&\times\limsup_{t\to\infty}\frac{1}{t}\sum_{j=0}^{t-1}\Pr\left(E(j)<\hat{P}\right).\end{aligned}\tag{25}$$

Eq. (25) is not very informative since our original intention was to provide a bound for the liminf of $\sum_{l=1}^{L} f_l\left(\bar{r}_l(t)\right)$ and not $\sum_{l=1}^{L} f_l\left(\bar{\gamma}_l(t)\right)$. Additionally, we need to estimate the limsup appearing in (25). The first issue is handled in

[11] (see also [16]), which exploits the properties of *any* stabilizing policy to prove that $\liminf_{t\to\infty}\sum_{l=1}^{L} f_l\left(\bar{\gamma}_l(t)\right) = \liminf_{t\to\infty}\sum_{l=1}^{L} f_l\left(\bar{r}_l(t)\right)$. The following statement, proved in Appendix A, provides an estimation of the limsup appearing in (25).

Lemma 5: Under DRABP, it holds

$$\begin{aligned}&\limsup_{t\to\infty}\frac{1}{t}\sum_{j=0}^{t-1}\Pr\left(E(j)<\hat{P}\right)\\&\leq\Pr\left(\sum_{k=0}^{\sigma-1}B(k)\leq\frac{\hat{P}+\hat{D}-E_{max}}{\delta}\right),\end{aligned}\tag{26}$$

where $\sigma \triangleq \lceil E_{max}/\hat{P}-1\rceil$.

Combining the liminf equality with Lemma 5 and (25) yields the main result for the downlink case

Theorem 1: DRABP stabilizes the system and achieves a performance bound of

$$\begin{aligned}&\liminf_{t\to\infty}\sum_{l=1}^{L} f_l\left(\bar{r}_l(t)\right)\geq g_\delta^*-\dot{C}_1-\dot{C}_2\\&\times\Pr\left(\sum_{k=0}^{\sigma-1}B(k)\leq\frac{\hat{P}+\hat{D}-E_{max}}{\delta}\right),\end{aligned}\tag{27}$$

where $\dot{C}_1=\dot{C}/(2M)$ and $\dot{C}_2=\sum_{l=1}^{L}(\hat{c}_l\hat{U}_l)/M$.
An examination of Lemma 3 reveals that $\hat{U}_l, \hat{D} = \Theta(M)$, $\dot{C}, \dot{C}_2 = \Theta(1)$ and $\dot{C}_1 = O(1/M)$. Theorem 1 now implies the following result

Corollary 1: For a given policy parameter $M > 0$, a selection of E_{max} such that $E_{max} > \hat{P}+\hat{D}$ implies that DRABP satisfies $\liminf_{t\to\infty}\sum_{l=1}^{L} f_l\left(\bar{r}_l(t)\right)\geq g_\delta^*-\dot{C}_1$.
Since $\dot{C}_1 = O(1/M)$ and $g_\delta^* \to g_{avg}^*$ as $\delta\to 0$, the above RHS can approach g_{avg}^* arbitrarily close for sufficiently small δ and sufficiently large M, as long as it holds $E_{max} > \hat{P}+\hat{D}$, which implies the asymptotic optimality of DRABP (see [16] for a calculation of the probability in (27) in the regime $\hat{P}+\hat{D}-\sigma\delta\bar{B} < E_{max} \leq \hat{P}+\hat{D}$).

V. Extension to Multihop Networks

A careful examination of the proofs of the various statements in Section IV reveals that, apart from algebraic manipulations, the optimality of DRABP rests on the following crucial point: queues $\boldsymbol{Y}(t)$, $\boldsymbol{U}(t)$, $D(t)$ are finitely bounded (in fact, the boundedness of each queue is used to prove the boundedness of the subsequent ones) and, by construction of $D(t)$, the battery overflows infinitely often. Hence, we expect a policy that has the previous properties, when applied to a general network, to achieve similar performance to DRABP. This is studied next.

By definition of single-hop networks, the bits exit the network layer once they are transmitted from their source node. Hence, a single-hop network is essentially an aggregate of interacting downlinks so that the analysis of Section IV can be repeated verbatim to produce a "vectorized" version of DRABP (in fact, an asymptotically optimal policy for a single-hop network is produced by replacing all terms $\sum_l$ in (10)–(12) with $\sum_{n,l}$, where n is the node index). Since this is a straightforward procedure, it will not be pursued any further (see [16] for details).

The crucial difference between single-hop and multihop networks is that in the latter case intra-node traffic, in addition to exogenous arrivals, is explicitly allowed. This suggests that a vectorization of DRABP to a multihop setting will be more involved. In fact, new notation is required to capture the fact that a packet may need many hops to reach its destination. We adopt the model in [11], which views a multihop network as a standard digraph $(\mathcal{N},\mathcal{L})$ (where $\mathcal{N}=\{1,\ldots,N\}$), and use the commodity concept to assume that each bit belongs to a packet with an associated commodity $c\in\mathcal{K}$ (which minimally defines the packet destination but may contain additional information). Hence, we denote with $A_n^{(c)}(t)$ the number of commodity c bits exogenously generated at node n and slot t (we assume $A_n^{(c)}(t)$ to be iid with expectation $\lambda_n^{(c)}$ and an upper bound of $\hat{A}_n^{(c)}$) with similar interpretations for the internal/external queues $U_n^{(c)}(t)$, $V_n^{(c)}(t)$, respectively. We also denote with $R_{n,in}^{(c)}(t)$ the number of externally admitted commodity c bits at node n and slot t, while $\mu_{ab}^{(c)}(t)$ is the number of commodity c bits transmitted on link (a,b) during slot t. The total number of bits (for all commodities) transmitted over a link is upper bounded by a vector function $\boldsymbol{c}\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right)$ with the properties mentioned in Section II, so that it holds $\sum_c \mu_{ab}^{(c)}\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right)\leq c_{ab}\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right)$. We also define $\hat{c}_{ab}\triangleq\max_{\boldsymbol{s},\boldsymbol{p}} c_{ab}\left(\boldsymbol{s},\boldsymbol{p}\right)$. The objective is to maximize[3] $\liminf_{t\to\infty}\sum_{n,c} f_{nc}\left(\bar{r}_{n,in}^{(c)}(t)\right)$, where $\bar{r}_{n,in}^{(c)}(t)\triangleq\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}[R_{n,in}^{(c)}(\tau)]$ and $f_{nc}(\cdot)$ is a typical utility function.

A. Multihop policy specification

Motivated by the downlink analysis, we introduce the virtual queues $D_n(t)$ and $Y_n^{(c)}(t)$ to handle the average power and (finite) arrival constraints, respectively, along with the processes $\gamma_n^{(c)}(t)$. Hence, queues $V_n^{(c)}(t)$, $E_n(t)$, $Y_n^{(c)}(t)$, $D_n(t)$ evolve according to the vectorized version of the corresponding downlink evolution, while for $U_n^{(c)}(t)$ it holds

$$\begin{aligned}U_n^{(c)}(t+1)\leq&\left[U_n^{(c)}(t)-\sum_b\mu_{nb}^{(c)}\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right)\right]^+ +R_{n,in}^{(c)}(t)\\&+\sum_a\mu_{an}^{(c)}\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right).\end{aligned}\tag{28}$$

Eq. (28) is an inequality rather than an equality because the actual amount of incoming traffic to node n may be less than $\sum_a\mu_{an}^{(c)}\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right)$ due to low queue occupancy of its neighbors. The following constraints also exist

$$\begin{aligned}&R_{n,in}^{(c)}(t)\leq V_n^{(c)}(t)+A_n^{(c)}(t)\quad\forall n,c,t,\\&\mu_{ab}^{(c)}\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right)\geq 0\quad\forall a,b,t,\\&\mu_{ab}^{(c)}\left(\boldsymbol{S}(t),\boldsymbol{P}(t)\right)=0\quad\forall (a,b)\notin\mathcal{L}^{(c)},\\&\boldsymbol{P}(t)\in\mathcal{P}_{\boldsymbol{S}(t)},\quad\sum_b P_{nb}(t)\leq\min\left(E_n(t),\hat{P}\right)\quad\forall n,t,\end{aligned}$$

[3] in the following, the indices n,c will range over the sets $\mathcal{N}$, $\mathcal{K}$, respectively, unless otherwise stated.

where indices a, b range over the set of incoming and outgoing neighbors of node n, respectively. The second constraint in (29) models the fact that all commodity c bits may be required to be transmitted through links belonging to a specific set $\mathcal{L}^{(c)} \subseteq \mathcal{L}$ *only* (setting $\mathcal{L}^{(c)} = \mathcal{L}\ \forall\, c$ effectively removes this constraint so that all commodities can be routed through all links).

As in the downlink case, the proposed policy is similar to that of [11] and is presented for completeness.

Network Rechargeable Adaptive Backpressure Policy (NRABP)

- $\boldsymbol{R}_{in}(t)$, $\boldsymbol{\gamma}(t)$ are selected according to the vectorized versions of (10), (11), respectively (see [16] for the exact form of the optimization problems).
- observe $\boldsymbol{S}(t)$ and select $\boldsymbol{P}(t)$ according to

$$\boldsymbol{P}(t) = \arg\max \sum_{ab} \left[W^*_{ab}(t) c_{ab}\left(\boldsymbol{S}(t), \boldsymbol{p}\right) - D_a(t) p_{ab}\right]$$
$$\text{s.t.} \quad \boldsymbol{p} \in \mathcal{P}_{\boldsymbol{S}(t)}, \quad \sum_b p_{nb} \leq \min\left(E_n(t), \hat{P}\right) \tag{29}$$

where $W^*_{ab}(t) = \left[\max_{c:(a,b)\in\mathcal{L}^{(c)}} \left\{U^{(c)}_a(t) - U^{(c)}_b(t)\right\}\right]^+$ so that, for each time slot and link, we need only select the commodity of largest differential backlog $c^*_{ab}(t) = \arg\max_{c:(a,b)\in\mathcal{L}^{(c)}} \{U^{(c)}_a(t) - U^{(c)}_b(t) \ :\ U^{(c)}_a(t) > U^{(c)}_b(t)\}$ (obviously, different links may carry different commodities. Also, it holds $W^*_{ab}(t) = 0 \Leftrightarrow c^*_{ab}(t) = \varnothing$, i. e. the bits always travel from the higher to the lower backlog node).

The computational complexity and distributed solution of the optimization problems above is briefly discussed in Section V-C. The structure of these problems leads to the following result, which is proved in [16].

Lemma 6: Application of NRABP satisfies the following conditions

- for each $c \in \mathcal{K}$, there exists a sufficiently large $\hat{F}^{(c)}$ such that $R^{(c)}_{n,in}(t) = 0$ whenever $U^{(c)}_n(t) > \hat{F}^{(c)}$.
- no transmission of commodity c data occurs on a link (a, b) (i. e. $P_{ab}(t) = 0$) if it holds $U^{(c)}_a(t) \leq U^{(c)}_b(t)$.

B. Stability and performance properties of NRABP

The derivation of a performance bound for NRABP follows the downlink procedure, i. e. we need to establish a performance upper bound for any policy and show that NRABP can approach it arbitrarily close. For the former, we use the concept of consistent flow from [11]

Definition 1: For given node set $\mathcal{N}$, commodity set $\mathcal{K}$ and link constraint sets $\mathcal{L}^{(c)}$, a (consistent) multi-commodity flow $\left\{f^{(c)}_{ab}\right\}$ is a vector that satisfies the following conditions for all $a, b \in \mathcal{N}$ and $c \in \mathcal{K}$

$$f^{(c)}_{ab} \geq 0, \quad f^{(c)}_{aa} = f^{(c)}_{dest(c),b} = 0,$$
$$(a,b) \notin \mathcal{L}^{(c)} \Rightarrow f^{(c)}_{ab} = 0, \tag{30}$$

where $dest(c)$ is the destination node for commodity c.

We also define the set

$$\mathcal{R}_\delta \triangleq \Bigg\{\boldsymbol{r} = \left(r^{(c)}_n\right) : \exists\, \pi^{\boldsymbol{s}}_{\boldsymbol{p}},\ \text{flow } \{f^{(c)}_{ab}\} \text{ s.t. } \sum_{\boldsymbol{p}} \pi^{\boldsymbol{s}}_{\boldsymbol{p}} = 1\ \forall\, \boldsymbol{s}$$
$$r^{(c)}_n \leq \sum_b f^{(c)}_{nb} - \sum_a f^{(c)}_{an}\ \forall\, n, c,$$
$$\sum_c f^{(c)}_{ab} \leq \sum_{\boldsymbol{s},\boldsymbol{p}} c_{ab}\left(\boldsymbol{s}, \boldsymbol{p}\right) \pi^{\boldsymbol{s}}_{\boldsymbol{p}} \pi_{\boldsymbol{s}}\ \ \forall\, a, b,$$
$$\sum_{\boldsymbol{s},\boldsymbol{p}} \sum_b p_{nb} \pi^{\boldsymbol{s}}_{\boldsymbol{p}} \pi_{\boldsymbol{s}} \leq (1-\delta)\bar{B}_n\ \ \forall\, n \Bigg\}, \tag{31}$$

and the quantity $g^*_\delta \triangleq \max_{\boldsymbol{r}\in\Lambda_\delta} \sum_{n,c} f_{nc}\left(r^{(c)}_n\right)$, where $\Lambda_\delta = \{\boldsymbol{r} \in \mathcal{R}_\delta : \boldsymbol{0} \leq \boldsymbol{r} \leq \boldsymbol{\lambda}\}$. Reference [11] now provides the following analogue to Lemma 1

Lemma 7: No stabilizing policy acting on a multihop network with infinite capacity batteries and an average power constraint of $\bar{\boldsymbol{B}}$ can achieve performance greater than $g^*_{avg} \triangleq g^*_0$. It also holds $\lim_{\delta\downarrow 0} g^*_\delta = g^*_0$.

In order to show that NRABP can approach g^*_{avg} arbitrarily close, we repeat the downlink procedure starting from (15) (through its obvious vectorization) which, as evidenced from the last line of (18), relies crucially on all queues $U^{(c)}_n(t)$ being finitely bounded. The latter fact cannot be proved exclusively by a verbatim repetition of the proof of Lemma 3 due to the additional term $\sum_a \mu^{(c)}_{an}\left(\boldsymbol{S}(t), \boldsymbol{P}(t)\right)$ of intra-node traffic (which does not exist in the downlink case). Hence, a new approach is needed. It is easy to see that, in order to prove the finiteness of the queues under NRABP, it suffices to prove the following important result

Lemma 8: For any policy that satisfies the conditions of Lemma 6, there exists a sequence $\{a^{(c)}_n\}_{n=1}^N$ such that it holds for all t, c

$$\max_{\substack{\mathcal{I}\subseteq\{1,\ldots,N\}\\ |\mathcal{I}|=k}} \sum_{i\in\mathcal{I}} U^{(c)}_i(t) \leq k\hat{F}^{(c)} + a^{(c)}_k. \tag{32}$$

Proof: The proof is quite long and requires some intermediate results to be established first. Due to space restrictions, it is presented in detail in [16]. ■

Since NRABP already satisfies the conditions of Lemma 6, we can combine Lemmas 6, 8 to show that $\boldsymbol{U}(t)$ is finitely bounded under NRABP. Repeating the arguments in the proof of Lemma 3 for $\boldsymbol{Y}(t)$, $D(t)$ now yields the following

Corollary 2: Under NRABP, all queues are bounded according to

$$\begin{aligned} &Y^{(c)}_n(t) \leq V f'_{nc}(0) + \hat{A}^{(c)}_n \triangleq \hat{Y}^{(c)}_n, \\ &U^{(c)}_n(t) \leq \hat{U}^{(c)}_n, \\ &D_n(t) \leq \hat{U}_n \hat{C}_n + \hat{P}_n \triangleq \hat{D}_n, \end{aligned} \tag{33}$$

where $\hat{\hat{U}}_n = \max_c \hat{U}^{(c)}_n$ and $\hat{C}_n = \max_{\boldsymbol{s}} \sum_b \frac{\partial c_{ab}}{\partial p_{ab}}\left(\boldsymbol{s}, \boldsymbol{0}\right)$.

The rest of the procedure for producing a performance bound for NRABP consists of appropriate vectorizations of (15) and (17)–(24) and is omitted due to space restrictions (see [16] for details).

C. Distributed implementation of NRABP

Since $\boldsymbol{R}_{in}(t)$, $\boldsymbol{\gamma}(t)$ are selected in NRABP as the solutions to the vectorized versions of the corresponding DRABP optimization problems, they retain the properties of their DRABP counterparts, namely "bang-bang" form and separability, respectively (see also the last paragraph of Section III). Hence, $\boldsymbol{R}_{in}(t)$ is computed in closed form, while $\boldsymbol{\gamma}(t)$ is computed in closed form or numerically, through a simple one-dimensional optimization. Both computations use local information only. Thus, the computational core (in terms of algorithmic complexity) of NRABP is the resource allocation problem of (29). The possibility for a distributed solution to (29) is crucially affected by the exact form of $c_{ab}(\boldsymbol{S}(t), \boldsymbol{p})$ (which, in turn, is determined by the interference properties of the network). To the best of our knowledge, no distributed algorithm has been proposed for the general case of arbitrary interference and the search for such an algorithm is an important topic for future work. Nevertheless, there exist special cases for which efficient distributed algorithms have been developed. Two such cases are presented next.

In the first case, we assume there is no interference (say, by employing a suitable CDMA scheme) so that the link capacity depends only on the power allocated to that link and the link's channel state, i. e. $c_{ab}(\boldsymbol{S}(t), \boldsymbol{p}(t)) = a_{s_{ab}(t)} p_{ab}(t)$. We further assume that the power allocated to each link ranges over the integer set $\{0, 1, \ldots, \hat{P}\}$ (recall that $\boldsymbol{P}(t)$, $\boldsymbol{B}(t)$ are quantized), whence it follows that (29) is reduced to the solution of a bounded knapsack problem [17] for each node, where the node need only know the queue sizes of its neighbors (this information is acquired through one-hop message exchanges), while the actual knapsack computation is performed in each node separately.

In the second case, we consider the well-known "node-exclusive" interference model, where each node can transmit to at most one other node and no node can simultaneously transmit and receive. Assuming that each node can transmit only with peak power, if at all, (i. e. the allocated power ranges over the set $\{0, \hat{P}\}$) and the link capacity has the same form as in the previous paragraph, it is clear that (29) is reduced to a maximum-weight matching problem.[4] Although the optimal solution to (29) can only be obtained, to the best of our knowledge, through a standard centralized algorithm, there exists a simple distributed algorithm [18] that achieves, for each slot, a weight no smaller than $1/2$ times the optimal value, where each node i need transmit only $O(deg(i))$ control messages (where $deg(i)$ stands for the degree of node i).

Combining the latter fact with a known result on imperfect scheduling [11, Section 4.7] implies that the incorporation of this suboptimal algorithm to NRABP stabilizes the system for any rate that belongs to the set $(1/2)\Lambda_\delta$ with a corresponding reduction in the performance bound.[5] Similar suboptimal solutions, which compute a *maximal* instead of a maximum weight matching, have been proposed in [19], [20], [21] for non-rechargeable battery settings. Finally, a distributed algorithm for maximum weight matching that achieves a fraction of the stability region arbitrarily close to unity with a constant number of exchanged messages per node has been recently developed in [22], albeit for a model that ignores channel variations. Although the algorithm can, in principle, be applied to our model, it is not clear at this point whether a guaranteed performance bound can be derived in our varying channel model. Such a derivation is a possible topic for future work.

[4]the weight $\tilde{w}_{ab}(t)$ for link (a,b) is given by $\tilde{w}_{ab}(t) = \left[W^*_{ab}(t) a_{s_{ab}(t)} - D_a(t)\right] \mathbb{I}\left[E_a(t) > \hat{P}\right]$, where $\mathbb{I}[\cdot]$ denotes the indicator function of $[\cdot]$, and is queue size, channel state and battery dependent.

[5]specifically, g^*_δ in (27) should be replaced by $g^*_\delta(1/2)$, where $g^*_\delta(1/2) \triangleq \max_{\boldsymbol{r}\in(1/2)\Lambda_\delta} \sum_{n,c} f_{nc}\left(r_n^{(c)}\right)$.

VI. Numerical Results

Although theoretical performance bounds already exist in the form of (27), a numerical evaluation of the proposed policy is useful in examining its performance in the non-asymptotic regime (i. e. for moderate values of $E_{max}/\hat{P}$). Specifically, we consider a single downlink user and a time-varying channel that has 5 states with probability $\pi_{\boldsymbol{S}} = (0.045, 0.526, 0.332, 0.087, 0.01)$ and a channel gain, respectively, $a_{\boldsymbol{S}} = (1, 2, 5, 8, 10)$ bits per slot per energy quantum (i. e. for the first channel state, we must expend one energy quantum for each transmitted bit in the current slot, while, for the third state, a single quantum allows the transmission of 5 bits per slot etc.). The objective is throughput maximization subject to the instantaneous power constraint $\hat{P} = 50$ and the battery constraints. The exogenous traffic arrival is modeled either as a Poisson process or as a deterministic process with rate strictly larger than the maximum transmission rate (the latter implies that the external queue $V(t)$ will never become zero so that the node will always have available bits for transmission. This is hereafter referred to as the "infinite arrival" model), while the recharge process is a random variable in the integer set $\{0, \ldots, \hat{B}\}$ with a "triangle-on-a-pedestal" distribution (i. e. a triangular distribution where the probability masses for 0 and $\hat{B}$ are non-zero) of expected value $\bar{B} = \hat{B}/2$.

The single downlink model is sufficiently simple to allow the computation of the upper bound of Lemma 1 (which is reduced to the solution of a linear program) as well as the DP solution, via value iteration, of the underlying infinite-horizon sequential decision problem [23]. In fact, the availability of this comparative data is the main reason for simulating a single downlink scenario rather than the (more realistic) multihop network, where DP is practically inapplicable. For the application of DRABP, we select $M = 500$, $\delta = 0.01$ and simulate the system for 10^8 slots. The results are shown in Fig. 2 for the infinite arrival model and different recharge parameters, where all quantities are normalized w. r. t. the upper bound of Lemma 1, i. e. we plot the ratio (DRABP perf.)$/g^*_\delta$, where both the numerator and denominator of the ratio are implicit functions of $\bar{B}$ (the performance of DRABP also depends on E_{max} while g^*_δ is independent of E_{max}, since it refers to the infinite battery model). We also present, in Table I, the performance of DRABP in absolute terms, whence it is apparent that the achieved throughput is an increasing function of $\bar{B}$, as expected (ignore the DRABP_0 columns for now). The surprising result is the fact that DRABP performs sufficiently well, e. g. above 90% of the upper bound, for moderate $E_{max}/\hat{P}$ values, say 10 (even for $E_{max}/\hat{P} = 1$,

TABLE I
COMPARISON BETWEEN DRABP AND DRABP_0 PERFORMANCE FOR INFINITE ARRIVALS

	$B = 2.5$ ($g^*_{avg} = 21$)		$B = 5$ ($g^*_{avg} = 40.55$)		$B = 10$ ($g^*_{avg} = 65.55$)	
$E_{max}/\hat{P}$	DRABP	DRABP_0	DRABP	DRABP_0	DRABP	DRABP_0
1:1	14.7999	8.88134	28.606	17.7623	40.6234	35.5245
2:1	17.4403	8.88134	29.733	17.7623	57.1583	35.5245
5:1	19.2334	8.88135	34.4681	17.7623	59.7849	35.5245
10:1	19.5879	8.88136	36.9955	17.7623	62.4758	35.5245
20:1	19.6997	8.88138	37.3065	17.7623	63.7591	35.5246
50:1	19.7045	8.88144	37.3119	17.7624	63.779	35.5246
100:1	19.7062	8.88153	37.3132	17.7625	63.7801	35.5247

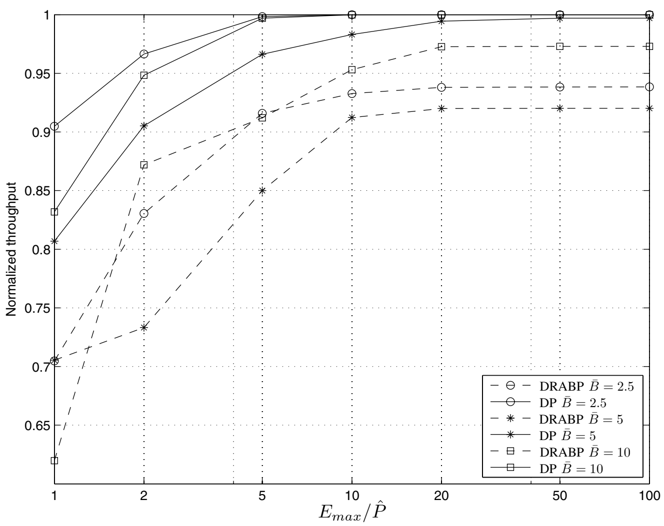


Fig. 2. Single queue DRABP and DP performance (normalized w. r. t. g^*_{avg}) comparison for infinite arrivals: $\bar{B} = 2.5, 5, 10$.

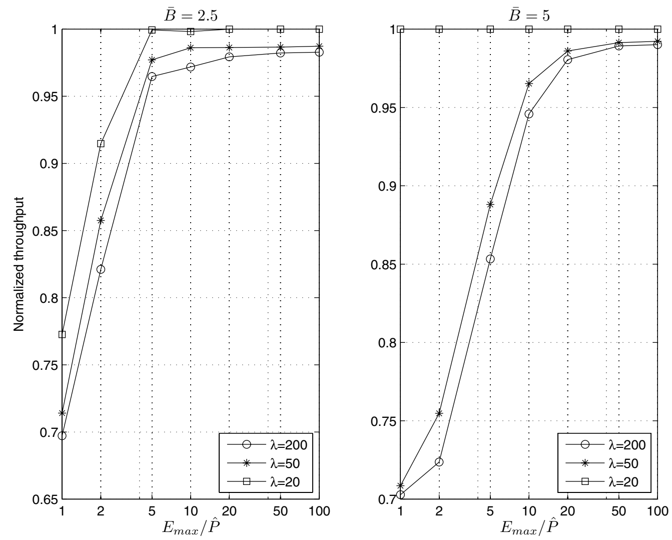


Fig. 3. Single queue DRABP performance (normalized w. r. t. g^*_{avg}) for Poisson arrivals with varying mean λ: $\bar{B} = 2.5$ and $\bar{B} = 5$.

DRABP performs above 70% of the optimal DP value). Similar behavior is observed for Poisson (finite) arrivals, as witnessed in Fig. 3. No DP results are available for the Poisson arrivals since that would require the inclusion of the queue size into the state space, which would grow the latter into intractable proportions.

We also consider an alternative policy motivated by the following observation. Since any policy acting on the rechargeable battery satisfies (7) by default, one might be tempted to remove the virtual queue $D(t)$ from our model (since this queue was originally introduced in [11] to handle the average power constraint, which is already satisfied in our case), effectively setting $D(t) = 0 \quad \forall\, t$ in (12) , and apply the resulting policy (referred to as DRABP_0) to the rechargeable battery. For the special case of a single downlink, this naive policy is a greedy policy under which the controller always decides to transmit with maximum available power (i. e. it holds $P(t) = \min(\hat{P}, E(t))$). Although this is clearly a valid policy that satisfies all constraints, its performance is expected to be poor since the condition $D(t) = 0$ also removes the "power instability" of the joint $E(t)$, $D(t)$ queues, which is the crucial component of asymptotic optimality. As a result, the naive policy will drain the battery w. p. 1 (provided that it holds $\bar{B} < \hat{P}$) and force the battery level to remain in the region $[0 \quad \hat{P}]$ for a long time interval. Indeed, simulating DRABP_0 with the same parameters as DRABP for infinite arrivals produces the results shown in the corresponding columns of Table I, where all quantities are absolute (i. e. no normalizations) and the numbers in parentheses in the first row denote the performance upper bounds for each case. Clearly, $D(t)$ plays a more important role in DRABP than just satisfying an average power constraint.

We finally evaluate DRABP in a multiple downlink scenario (i. e. $L > 1$). The enormous size of the underlying state space prohibits a DP solution, so we need to use a different method as a benchmark against DRABP. To this end, we exploit the known fact that the ABP-based policy of [11] satisfying an average power constraint of $\bar{B}$ achieves the upper bound g^*_{avg} as $M \to \infty$. Hence, we propose to simulate the latter policy (hereafter referred to as ABP_avg) with a large M and use its simulation-derived performance as an upper bound for DRABP. We also model different channel statistics by introducing three different processes S_A, S_B, S_C with corresponding channel gains[6], respectively, $a_{S_A} = (1, 2, 5, 8, 10)$, $a_{S_B} = (2, 4, 8, 12, 20)$ and $a_{S_C} = (3, 5, 9, 12, 25)$ bps per energy quantum. For notational convenience, we describe each process through its channel gains so that we write $S_A = \{1, 2, 5, 8, 10\}$, $S_B = \{2, 4, 8, 12, 20\}$, $S_C = \{3, 5, 9, 12, 25\}$.

[6]the probabilities of the corresponding channel gains are chosen as $\pi_{S_A} = (0.4105, 0.1528, 0.3847, 0.052)$, $\pi_{S_B} = (0.2761, 0.044, 0.367, 0.2928, 0.021)$, $\pi_{S_C} = (0.416, 0.1247, 0.1176, 0.2672, 0.0745)$.

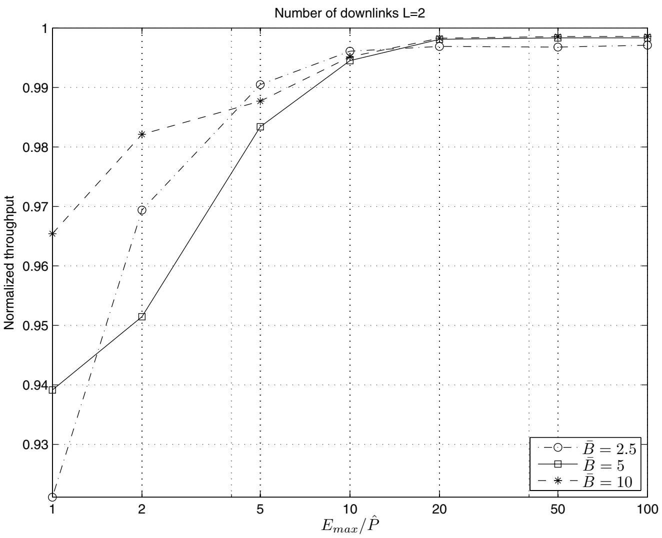


Fig. 4. DRABP performance for Poisson arrivals: Scenario 1.

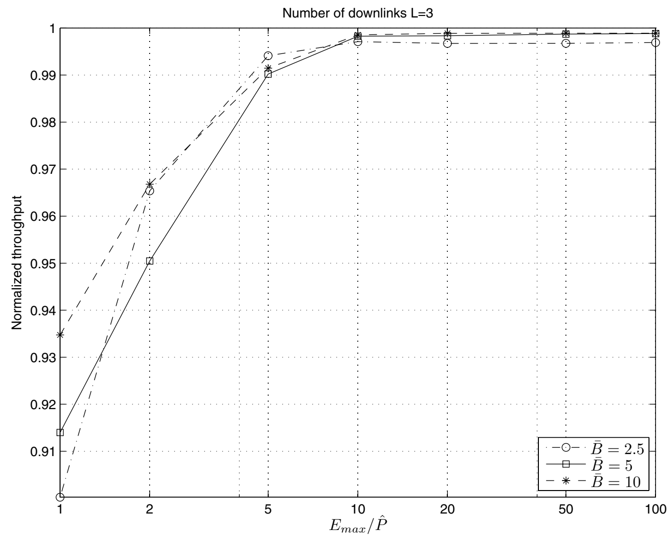


Fig. 5. DRABP performance for Poisson arrivals: Scenario 2.

We use the proportionally-fair utility $f(x) = \ln(1 + x)$ of [15], select $M = 5000$, $\delta = 0.01$ and simulate the DRABP and ABP_avg policies for 10^8 time slots, where the values of $\bar{B}$ and $E_{max}/\hat{P}$ are chosen as for the single downlink. The following scenarios are considered

1) $L = 2$, $\boldsymbol{S} = S_A \times S_B$ (i. e. the channel in the first/second link evolves according to process S_A/S_B, respectively) and the exogenous traffic is Poisson with rate $\boldsymbol{\lambda} = (60, 30)$.
2) $L = 3$, $\boldsymbol{S} = S_A \times S_B \times S_C$ and the exogenous Poisson rate is $\boldsymbol{\lambda} = (60, 30, 10)$.
3) $L = 5$, $\boldsymbol{S} = S_A \times S_A \times S_B \times S_B \times S_C$ and the Poisson rate is $\boldsymbol{\lambda} = (60, 50, 30, 20, 10)$.

The previous scenarios generated a large amount of simulation results, so we only present some typical cases in Figs. 4–6 where, as previously mentioned, DRABP performance is normalized w. r. t. the simulated performance of ABP_avg. A casual inspection of the Figures reveals that DRABP continues to perform very well even for moderate values of $E_{max}/\hat{P}$.

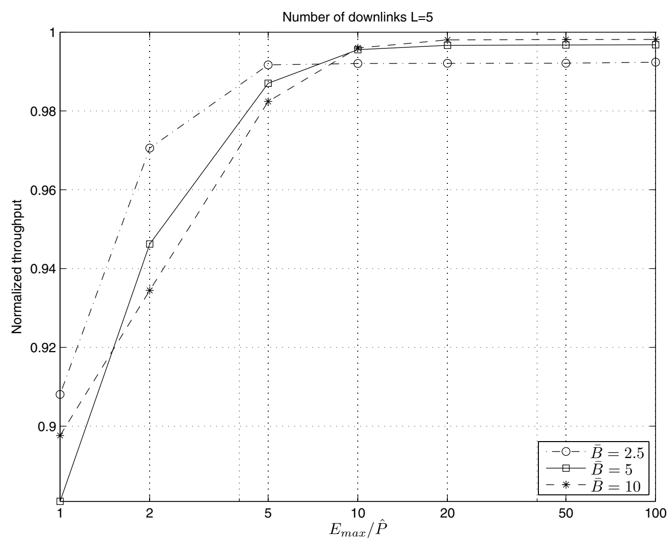


Fig. 6. DRABP performance for Poisson arrivals: Scenario 3.

VII. Conclusions

This paper presented an online adaptive policy for stabilization and optimal control of wireless networks operating with rechargeable batteries. Using a Lyapunov drift argument and modifying the framework of [11] in subtle but non-trivial ways, a performance bound was provided that ensures asymptotic optimality as $E_{max}/\hat{P} \to \infty$. The policy requires only current channel and recharged energy information and is particularly suitable for satellite or sensor networks, which typically operate under high $E_{max}/\hat{P}$ ratios. Although the analysis was presented for iid processes, the same policy can be applied in the case of general processes leading to similar performance bounds. In the future, we plan to investigate the existence of distributed optimal solutions (or suboptimal solutions with guaranteed bounds, in the spirit of [22], [24]) to the core resource allocation problems in (12), (29) under general interference patterns.

Appendix

A. Proof of Lemma 5

We assume w. l. o. g. that the lost recharge energy quanta of Fig. 1 enter a virtual queue $E_{over}(t)$ of zero service rate, so that $E_{over}(t)$ is a non-decreasing function of time. Viewing the queues $E(t)$, $D(t)$, $E_{over}(t)$ of Fig. 1 as a single compound queue, this compound is unstable since it has an instantaneous arrival rate of $B(t)$ and a corresponding service rate of at most $(1-\delta)B(t)$. However, queues $E(t)$, $D(t)$ are finite, the former by definition and the latter due to Lemma 3, which implies that $\limsup_{t\to\infty} E_{over}(t) = \infty$ w. p. 1. Since $E_{over}(t)$ is increased *only* when $E(t)$ overflows, it follows that, under DRABP, the queue $E(t)$ overflows infinitely often. Define $\tilde{\tau} \triangleq \inf\{t : E(t) = E_{max}\}$ as the first time the battery hits E_{max} under DRABP. Since $\tilde{\tau}$ is finite w. p. 1, a standard probabilistic argument shows that $\lim_{t\to\infty} \Pr(E(t) < \hat{P}, \tilde{\tau} > t) = 0$, so that it holds

$$\begin{aligned}&\limsup_{t\to\infty}\frac{1}{t}\sum_{j=0}^{t-1}\Pr\left(E(j)<\hat{P}\right)\\&\leq\limsup_{t\to\infty}\frac{1}{t}\sum_{j=0}^{t-1}\Pr\left(E(j)<\hat{P},\tilde{\tau}\leq j\right).\end{aligned}\tag{34}$$

Hence, the lemma is proved if the following is true for all j

$$\begin{aligned}\Pr\left(E(j)<\hat{P},\tilde{\tau}\leq j\right)&\stackrel{?}{\leq}\Pr\left(\sum_{k=\tau_j}^{j-1}B(k)\leq\frac{\hat{P}+\hat{D}-E_{max}}{\delta}\right)\\&\stackrel{?}{\leq}\Pr\left(\sum_{k=0}^{\sigma-1}B(k)\leq\frac{\hat{P}+\hat{D}-E_{max}}{\delta}\right),\end{aligned}\tag{35}$$

where equality follows from stationarity of $B(t)$ and $\tau_j\triangleq\sup_{s\leq j}\{s:E(s)=E_{max}\}$ is the last time up to (and including) j when the battery hit E_{max}.

If $\tilde{\tau}=j$, then $\tau_j=j$, $E(j)=E_{max}$ and $\{E(j)<\hat{P},\tilde{\tau}\leq j\}=\varnothing$, so that (35) holds trivially. Consider now the event $\{E(j)<\hat{P},\tilde{\tau}<j\}$. It holds

$$\left\{E(j)<\hat{P},\tilde{\tau}<j\right\}\subseteq\left\{j-\tau_j\geq\left\lceil\frac{E_{max}-\hat{P}}{\hat{P}}\right\rceil\triangleq\sigma\right\},\tag{36}$$

since the fastest way the battery can drop from E_{max} (its value at τ_j) to $E(j)$ is by transmitting with peak power and receiving zero recharge for all intermediate slots. Examining the combined queue $E(t)$, $D(t)$ in the interval $[\tau_j\quad j]$ yields

$$\begin{aligned}&E(j)+D(j)-E(\tau_j)-D(\tau_j)\geq\delta\sum_{k=\tau_j}^{j-1}B(k)\\&\Rightarrow E(j)+\hat{D}-E_{max}\geq\delta\sum_{k=\tau_j}^{j-1}B(k),\end{aligned}\tag{37}$$

where we exploited the boundedness of $D(t)$. Note that (37) holds for any j such that $E(j)<E_{max}$ and $\tilde{\tau}<j$. Hence, the following inclusion is true

$$\begin{aligned}\{E(j)<\hat{P},\tilde{\tau}<j\}&\subseteq\left\{\hat{P}+\hat{D}-E_{max}\geq\delta\sum_{k=\tau_j}^{j-1}B(k)\right\}\\&\subseteq\left\{\hat{P}+\hat{D}-E_{max}\geq\delta\sum_{k=\tau_j}^{\tau_j+\sigma-1}B(k)\right\},\end{aligned}\tag{38}$$

where the second inclusion is due to (36). Taking probabilities in (38) and using the stationarity of $B(t)$ produces (35). Since the RHS of (35) is independent of j (and therefore invariant under a j summation), the proof is complete.

References

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirici, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, Mar. 2002.

[2] X. Jiang, J. Polastre, and D. Culler, "Perpetual environmentally powered sensor networks," in *Proc. 4th International Symposium on Information Processing in Sensor Networks (IPSN)*, 2005, pp. 463–468.

[3] A. Fu, E. Modiano, and J. Tsitsiklis, "Optimal energy allocation and admission control for communication satellites," *IEEE/ACM Trans. Networking*, vol. 11, no. 3, p. 488, June 2003.

[4] L. Lin, N. Shroff, and R. Srikant, "Asymptotically optimal power-aware routing for multihop wireless networks with renewable energy sources," in *Proc. IEEE INFOCOM*, 2005.

[5] M. Adamou and S. Sarkar, "A framework for optimal battery management for wireless nodes," in *Proc. IEEE INFOCOM*, 2002.

[6] V. Borkar, A. Kherani, and B. Prabhu, "Closed and open loop optimal control of buffer and energy of a wireless device," in *Proc. 3rd Intl. Symposium on Modeling and Optimization in Mobile Ad Hoc and Wireless Networks (WiOpt)*, 2005.

[7] K. Kar, A. Krishnamurthy, and N. Jaggi, "Dynamic node activation in networks of rechargeable sensors," *IEEE/ACM Trans. Networking*, vol. 14, no. 1, pp. 15–26, Feb. 2006.

[8] N. Jaggi, A. Krishnamurthy, and K. Kar, "Utility maximizing node activation policies in networks of partially rechargeable sensors," in *Proc. 39th Annual Conference on Information Sciences and Systems (CISS)*, 2005.

[9] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," in *Proc. 5th Intl. Symposium on Modeling and Optimization in Mobile Ad Hoc and Wireless Networks (WiOpt)*, 2007.

[10] N. Jaggi, "Node activation policies for energy-efficient coverage in rechargeable sensor networks," Ph.D. dissertation, Rensselaer Polytechnic Institue, 2007.

[11] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends in Networking*, vol. 1, no. 1, 2006.

[12] M. J. Neely, "Dynamic power allocation and routing for satellite and wireless networks with time varying channels," Ph.D. dissertation, MIT, 2003.

[13] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, pp. 1936–1949, Dec. 1992.

[14] M. J. Neely, "Energy optimal control for time varying wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 2915–2934, July 2006.

[15] M. Neely, E. Modiano, and C. Li, "Fairness and optimal stochastic control for heterogeneous networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 2, pp. 396–409, Apr. 2008.

[16] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Asymptotically optimal control of wireless networks with rechargeable batteries." [Online]. Available: http://users.auth.gr/∼leonid/public/TechReports/tecreport_rabp.pdf

[17] H. Kellerer, U. Pferschy, and D. Pisinger, *Knapsack Problems*. Springer, 2004.

[18] J. Hoepman, "Simple distributed weighted matchings." [Online]. Available: http://arxiv.org/abs/cs/0410047

[19] X. Lin and N. Shroff, "The impact of imperfect scheduling on cross-layer rate control in multihop wireless networks," in *Proc. IEEE INFOCOM*, vol. 3, 2005, pp. 1804–1814.

[20] P. Chaporkar, K. Kar, X. Luo, and S. Sarkar, "Throughput and fairness guarantees through maximal scheduling in wireless networks," *IEEE Trans. Inf. Theory*, vol. 54, no. 2, pp. 572–594, Feb. 2008.

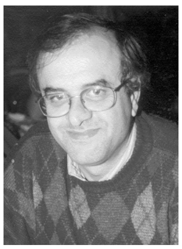
[21] X. Wu, R. Srikant, and J. Perkins, "Scheduling efficiency of distributed greedy scheduling algorithms in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, pp. 595–605, 2007.

[22] S. Sanghavi, L. Bui, and R. Srikant, "Distributed link scheduling with constant overhead," in *Proc. SIGMETRICS*, 2007, pp. 313–324.

[23] D. Bertsekas, *Dynamic Programming and Optimal Control: Vol. 2*. Athena Scientific, 1995.

[24] Y. Yi and M. Chiang, "Wireless scheduling algorithms with $O(1)$ overhead for M-hop interference model," in *Proc. IEEE International Conference on Communications (ICC)*, 2008, pp. 3105–3109.

Marios Gatzianas received the Diploma and PhD degree in Electrical and Computer Engineering from Aristotle University, Thessaloniki, Greece in 2000 and 2009, respectively, and the M.S. degree in EE from Arizona State University in 2003. His research interests include optimization and control of wireless ad-hoc networks, asymptotic analysis of algorithms, and distributed systems.



Leonidas Georgiadis (S'76-M'78-SM'96) received the Diploma degree in Electrical Engineering from Aristotle University, Thessaloniki, Greece, in 1979, and his M.S. and Ph.D. degrees both in electrical engineering from the University of Connecticut, in 1981 and 1986 respectively. From 1981 to 1983 he was with the Greek army. From 1986 to 1987 he was Research Assistant Professor at the University of Virginia, Charlottesville. In 1987 he joined IBM T. J. Watson Research Center, Yorktown Heights as a Research Staff Member. Since October 1995, he has been with the Telecommunications Department of Aristotle University, Thessaloniki, Greece. His interests are in the area of wireless networks, high speed networks, distributed systems, routing, scheduling, congestion control, modeling and performance analysis. In 1992 he received the IBM Outstanding Innovation Award for his work on Goal-Oriented Workload Management for Multi-class systems.

Leandros Tassiulas (S'89-M'91-SM'05-F'07) obtained the Diploma in Electrical Engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece in 1987, and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland, College Park in 1989 and 1991 respectively. He is Professor in the Dept. of Computer and Telecommunications Engineering, University of Thessaly, since 2002. He has held positions as Assistant Professor at Polytechnic University New York (1991-95), Assistant and Associate Professor University of Maryland College Park (1995-2001) and Professor University of Ioannina Greece (1999-2001). His research interests are in the field of computer and communication networks with emphasis on fundamental mathematical models, architectures and protocols of wireless systems, sensor networks, high-speed internet and satellite communications. Dr. Tassiulas is a Fellow of IEEE. He received a National Science Foundation (NSF) Research Initiation Award in 1992, an NSF CAREER Award in 1995 and Office of Naval Research, Young Investigator Award in 1997 and a Bodosaki Foundation award in 1999. He also received the INFOCOM 1994 best paper award and the INFOCOM 2007 achievement award.