Energy Optimal Control for Time Varying Wireless Networks

Michael J. Neely
University of Southern California
http://www-rcf.usc.edu/∼mjneely

***Abstract*— We develop a dynamic control strategy for minimizing energy expenditure in a time varying wireless network with adaptive transmission rates. The algorithm operates without knowledge of traffic rates or channel statistics, and yields average power that is arbitrarily close to the minimum possible value achieved by an algorithm optimized with complete knowledge of future events. Proximity to this optimal solution is shown to be inversely proportional to network delay. We then present a similar algorithm that solves the related problem of maximizing network throughput subject to peak and average power constraints. The techniques used in this paper are novel and establish a foundation for stochastic network optimization.**

***Index Terms*— Wireless Networking, Stochastic Optimization, Queueing Analysis, Distributed Algorithms**

I. Introduction

Wireless systems operate over time varying channels that are influenced by random environmental conditions, wireless fading, and power allocation decisions. To improve performance and meet the ever increasing demand for high throughput and low delay, modern wireless devices are designed with channel monitoring capabilities and rate adaptive technology. Such technology is currently being implemented for cellular communication with High Data Rate (HDR) services [2], and the ability to measure and react to channel information is expected to improve significantly.[1] It is of central importance to develop control strategies that take maximum advantage of this information to improve network performance and energy efficiency.

In this paper, we develop throughput optimal control strategies that conform to peak power constraints while minimizing average power expenditure. This design goal is crucial in all modern wireless scenarios, regardless of whether transmissions take place at a basestation, a hand-held unit, or at a node within an ad-hoc sensor network. Indeed, peak power constraints are important in systems with fixed hardware saturation levels or external environment regulations, while average power levels are important to extend network lifetime in systems with limited energy resources.

Here, we consider an ad-hoc network with N nodes and L wireless links, as shown in Fig. 1. We assume a slotted structure with slots equal to 1 time unit. Packets randomly arrive to the network every timeslot and must be delivered to their destinations, perhaps by routing over multi-hop paths.

[1]Indeed, it is claimed in [3] that channel measurements can be obtained almost as often as the symbol rate of the link in certain local area wireless networks.

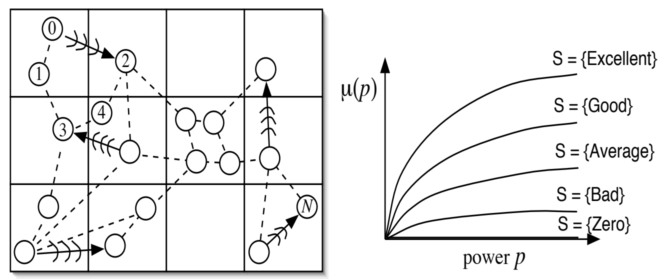

Fig. 1. A cell-partitioned wireless network, and an example set of rate-power curves for 5 different channel states.

The transmission rates of each data link are determined every timeslot by link channel conditions and network power allocation decisions according to an L-dimensional rate function $\vec{\mu}(\vec{P}(t), \vec{S}(t))$, where $\vec{P}(t)$ is a vector of power allocations and $\vec{S}(t)$ is a vector of parameters describing the current channel conditions. An example collection of rate-power curves for one data link with a discrete set of possible channel states in shown in Fig. 1. Although the network topology is assumed to remain fixed, link conditions may vary dramatically due to environmental effects, local mobility, or wireless fading.

Power vectors are restricted to a compact set Π of acceptable power allocations, so that $\vec{P}(t) \in \Pi$ for all t. All of our results hold for general rate functions $\vec{\mu}(\vec{P}, \vec{S})$ and general power sets Π. For example, the transmission rates over a particular link l can be modeled as a concave function of the signal to interference ratio at the receiver of the link, so that $\mu_l(\vec{P}, \vec{S})$ depends on the full vector of power allocations and channel states [14] [11] [16]. However, to simplify the multiple access control layer while capturing the geographic structure and interference properties of ad-hoc networks, we focus on a *cell partitioned network model*.

Under this model, the network region is divided into cells, each containing a distinct set of nodes. Specifically, we define $cell(n)$ as the cell of each node $n \in \{1, \ldots, N\}$, and define $tran(l)$ and $rec(l)$ as the transmitting and receiving nodes associated with a given wireless link $l \in \{1, \ldots, L\}$. We assume that each cell can support at most one active link transmission per timeslot, and that nodes can transmit only to other nodes in the same cell or in adjacent cells. That is, the feasible power set Π includes the constraint that if $P_l > 0$ for some link l, then $P_{\tilde{l}} = 0$ for all links $\tilde{l}$ such that $cell(tran(\tilde{l})) = cell(tran(l))$. We further assume that the transmission rate of each link depends only on the channel state and the power allocated to that link, so that $\vec{\mu}(\vec{P}, \vec{S}) = (\mu_1(P_1, S_1), \ldots, \mu_L(P_L, S_L))$. This structure arises if nodes in neighboring cells transmit over orthogonal frequency bands. In this way, if a node is transmitting then it cannot concurrently receive from nodes within the same cell, and any data it receives from adjacent cells must be on a different frequency band. It is well known that only 9 orthogonal subbands are required if the cell structure is rectilinear, and this number can be reduced to 7 if cells are arranged according to a hexagonal pattern. While the cell partitioned structure is not critical to our analysis, it simplifies exposition and allows scheduling decisions to be decoupled cell by cell. Relaxations or further restrictions on power assignment can easily be incorporated by modifying the set constraint Π or the rate function $\vec{\mu}(\vec{P}, \vec{S})$.

The goal of this paper is to develop a decentralized power allocation and routing algorithm that supports all incoming traffic while minimizing total energy. We develop a robust policy that does not require knowledge of input rates or channel probabilities yet uses a total energy that is arbitrarily close to the minimum energy expended by a system optimized with complete knowledge of future events. Distance to the minimum energy level is controlled by a parameter V effecting an explicit tradeoff in average end-to-end network delay.

Previous work in the area of power allocation for wireless systems can be categorized into static optimization solutions [4]-[11] and dynamic control algorithms [12]-[21]. In [4], a utility optimization problem is presented for a static wireless downlink, and pricing schemes are developed to enable power allocations to converge to a fair allocation vector. Linear programming, geometric programming, and other convex optimization methods are considered in [5]-[9] for routing and power allocation problems in wireless systems and sensor networks. Such techniques rely on the mathematical theory of Lagrangian duality (see, for example, [22]). This theory was applied in the landmark paper [23] to develop a mechanism for reaching an optimal static resource allocation in a non-wireless network.

We note that convex optimization approaches traditionally yield single-operating point solutions, which may not be well suited to cases when optimal networking involves *dynamic allocation* of resources. Indeed, in [11] it is shown that minimizing energy in an ad-hoc network with interference involves the computation of a *periodic schedule*, where wireless links are scheduled for transmission on a slot-by-slot basis to meet a given set of data transmission requirements. Specifically, the work in [11] treats a static network with known rate and channel parameters. Lagrangian duality is used to determine the structure of the optimal schedule by defining the solution variables as the *fractions of time* each particular resource configuration is scheduled, yielding dramatic improvements over any fixed resource allocation. A similar scheduling problem is shown to be NP-complete in [10], and a polynomial time algorithm for achieving a solution within a factor of 3 from optimality is given in [24] for a network with independent channels and transmission and reception constraints. The work in [11] [10] [24] is the most relevant to our current paper, although the same optimization techniques cannot be used because we consider a *stochastic network* with randomly varying channel states and unknown packet arrival rates. Two further differences are that we optimize multi-hop transfers over all possible decisions, and that we consider the full effects of queueing. Indeed, we show that queue information can be exploited to make energy optimal control decisions.

Prior work in the area of *stochastic* optimization and dynamic control for wireless networks considers much smaller systems with more a-priori information. In [16], an energy optimal transmission rule is developed for a static wireless downlink with one queue and with packet arrival times that are known in advance. A similar problem is formulated in [17] for a single-user satellite downlink with stochastic arrivals and channel states (with known arrival and channel probabilities), and optimal strategies are constructed with respect to energy and delay using dynamic programming. The capacity of a multi-user wireless downlink with randomly varying channels is established in [18], although the capacity achieving solution assumes perfect knowledge of channel statistics and traffic rates, and does not consider stochastic arrivals and queueing. A related downlink problem is treated in [19], where a utility optimal scheduling rule is designed and shown to conform to a simple index policy. The indices can be computed (in principle) if channel probabilities are fully known, or are approximated based on long-term measurements of channel conditions.

A wireless downlink with stochastic arrivals, peak and average power constraints, and queueing is developed in [21], and the strategy is shown to be delay-optimal under certain symmetry assumptions. This work makes use of a Lyapunov drift theory for stable scheduling, although it requires full knowledge of channel probabilities in order to meet the average energy requirement. Lyapunov theory can be used to design stabilizing power allocation and routing algorithms that do not require knowledge of arrival rates or channel statistics in cases where there are only peak power constraints on the wireless devices [14]. Historically, Lyapunov theory has been extremely useful in the development of stable queue control policies for radio networks and switching systems [12] [20] [15] [25] [13]. However, there was previously no Lyapunov method for performing system optimization (such as stabilizing a system with minimum energy). In this paper, we develop a novel Lyapunov drift technique that enables system stability and performance optimization to be achieved simultaneously. The technique bridges the gap between convex optimization theory and stochastic control problems, and establishes a new framework for *dynamic network optimization*.

For simplicity of exposition and to highlight the issues of power allocation, in the first half of this paper we consider only single-hop networks with no routing. The paper is organized as follows: In the next section we consider a motivating example of a 2-user wireless downlink. In Section III we develop an energy-minimizing control policy for one-hop networks. In Section IV we treat a related problem of maximizing throughput subject to peak and average power constraints (for cases when traffic is either supportable or insupportable). Extensions to multi-hop networks are treated in Section V, and simulations are presented in Section VI.

II. A Simple Example

To illustrate the decisions involved in energy-optimal scheduling, we consider the following example of a two-queue wireless downlink, where a single node (labeled 'node 0') transmits data to two different stations over downlink channels 1 and 2 (as in Fig. 1 in the case when only node 0 is active). The system operates in slotted time, and every slot the channel states are measured, power allocation decisions are made, and new arrivals are queued according to their destinations.

Let $U_1(t)$ and $U_2(t)$ represent the current backlog queued for transmission to destinations 1 and 2, respectively, and consider the decision of whether or not to allocate power to channel 1. Clearly no power should be allocated if $U_1(t) = 0$. When $U_1(t) > 0$, we must decide whether to allocate power on the current slot or wait for a more energy-efficient future channel state. In this example, we consider only ON/OFF power constraints and assume that either no power is allocated to any channel, or full power of 1 Watt is allocated to either channel 1 or channel 2. Link conditions for each channel 1 and 2 vary between '$Good$,' '$Medium$,' and 'Bad' states:

$$\vec{P}(t) = (P_1(t), P_2(t)) \in \Pi = \{(0,0),(1,0),(0,1)\}$$

$$S_1(t), S_2(t) \in \{G, M, B\}$$

Assume identical rate functions for $i = 1, 2$, given by:

$$\mu_i(0, S_i) = 0 \text{ units/slot} \quad \text{for all } S_i \in \{G, M, B\}$$
$$\mu_i(1, G) = 3, \mu_i(1, M) = 2, \mu_i(1, B) = 1 \text{ (units/slot)}$$

Let $A_1(t)$ and $A_2(t)$ represent the number of new data units arriving during slot t and destined for nodes 1 and 2, respectively. Queueing dynamics proceed according to the equation:

$$U_i(t+1) = \max[U_i(t) - \mu_i(P_i(t), S_i(t)), 0] + A_i(t)$$

Suppose arrivals $A_i(t)$ and channel states $S_i(t)$ for the first 9 timeslots $t \in \{0, \dots, 8\}$ are as given in Fig. 2, and consider the policy of allocating power to the channel with the largest rate-backlog product $U_i(t)\mu_i(t)$. This policy can be shown to stabilize the system whenever possible [20] [15] [13], although it is not necessarily energy-efficient. According to the figure, both queues are empty at time $t = 0$ when arrivals enter the system according to vector $(A_1(0), A_2(0)) = (3, 2)$, resulting in a backlog vector $(U_1(1), U_2(1)) = (3, 2)$ at the beginning of slot 1. Because the channel states at slot 1 are given by $(S_1(1), S_2(1)) = (G, M)$, the rate-backlog indices for channels 1 and 2 at slot 1 are given by $U_1(t)\mu_1(1, S_1(t)) = 9$, $U_2(t)\mu_2(1, S_2(t)) = 4$, so that the Max $U_i\mu_i$ policy places full power to channel 1 (as indicated by the boxed values in the figure).

Because there were no new arrivals during slot 1, the resulting backlog vector at time $t = 2$ is given by $(U_1(t), U_2(t)) = (0, 2)$, as shown in the figure. The policy proceeds by expending 1 Watt of power for time $t \in \{1, \dots, 8\}$, and the scheduling decision at slot $t = 8$ will leave the system again empty at time $t = 9$. If the same arrival and channel patterns were extended periodically every 9 timeslots, the Max $U_i(t)\mu_i(t)$ policy would allocate 1 Watt of power 8 timeslots out of every 9, yielding an average power consumption of $P_{av} = 8/9$ Watt. Similar power consumption levels are observed when the policy is simulated for random arrivals and channel states with the same steady state distributions as this example (see Section VI).

	t	0	1	2	3	4	5	6	7	8
Arrivals	$A_1(t)$	3	0	3	0	0	1	0	1	0
	$A_2(t)$	2	0	1	0	1	1	0	0	0
Channels	$S_1(t)$	G	G	M	M	G	G	M	M	G
	$S_2(t)$	M	M	B	M	B	M	B	G	B
Max $U_i\mu_i$ Policy	$U_1(t)$	0	$\boxed{3}$	0	$\boxed{3}$	$\boxed{1}$	0	1	1	$\boxed{2}$
	$U_2(t)$	0	2	$\boxed{2}$	2	2	$\boxed{3}$	$\boxed{2}$	$\boxed{1}$	0
Better Choices	$U_1(t)$	0	3	3	6	$\boxed{6}$	$\boxed{3}$	1	1	$\boxed{2}$
	$U_2(t)$	0	2	2	$\boxed{3}$	1	2	3	$\boxed{3}$	0

Fig. 2. An example set of arrivals, channel conditions, and queue backlogs for a two queue wireless downlink under two different scheduling algorithms, illustrating the power efficiency gains enabled by having full knowledge of future arrivals and channel states.

Now consider the alternate policy of waiting until slot 3 to allocate power, and then making decisions as shown in the figure. These decisions also leave the system empty at slot 9, but yield an average power expenditure of $P_{av} = 5/9$ Watt over the 9 slot interval.

The above example illustrates the energy gains available by more intelligent scheduling. In cases where power can be allocated as a continuous variable, more complex decisions are involved: Should we exploit better channel states by transmitting at higher data rates with the same power level, or by transmitting at the same data rate with reduced power? Optimizing over all of these decisions is a daunting task, and is conjectured to be impossible in [26] due to the uncertainties of future events. Remarkably, in the next section we develop a simple decision making strategy that does not require knowledge of future events, traffic rates, or channel statistics, yet yields an average power expenditure that is arbitrarily close to optimal.

III. Single Hop Networks

Consider the wireless network of Fig. 1 with N nodes and L links, where each link corresponds to a directed transmission from one node to another. Packets randomly arrive to the system and are queued according to their destinations. This is a single-hop network, and hence incoming data is associated with a particular transmission link $l \in \{1, \dots, L\}$ and is assumed to leave the network once it is transmitted. Let $A_l(t)$ represent the amount of bits arriving for transmission over link l during slot t, and let $U_l(t)$ represent the current queue backlog (or 'unfinished work') in queue l. Let $\vec{S}(t)$ and $\vec{P}(t)$ represent the L-dimensional vectors of channel states and power allocations. In vector notation, the queueing dynamics are:

$$\vec{U}(t+1) = \max[\vec{U}(t) - \vec{\mu}(\vec{P}(t), \vec{S}(t)), 0] + \vec{A}(t) \tag{1}$$

where $\vec{\mu}(\vec{P}, \vec{S})$ is the rate function associated with the given physical layer modulation and coding strategies used for wireless communication.

We assume that there are a finite number of channel state vectors $\vec{S}$, and that $\vec{\mu}(\vec{P}, \vec{S})$ is a continuous function of the power vector $\vec{P}$ for each channel state $\vec{S}$. Every timeslot a power vector $\vec{P}(t)$ is chosen in reaction to queue backlog and current channel conditions, subject to the constraint that $\vec{P}(t) \in \Pi$ for all t, where Π is a compact set of acceptable power vectors. Throughout this paper, we use these general rate functions and set constraints to present our main results. However, in all examples of distributed implementation, we assume the rate function has the structure that $\vec{\mu}(\vec{P}, \vec{S}) = (\mu_1(P_1, S_1), \ldots, \mu_L(P_L, S_L))$. Further, we assume that Π consists of all vectors $\vec{P}$ satisfying the cell-partition constraint (i.e., that if $P_l > 0$ for some link l, then $P_{\tilde{l}} = 0$ for all $\tilde{l}$ such that $cell(tran(\tilde{l})) = cell(tran(l))$) and such that each entry P_l is limited by a peak value P_{peak} according to either the continuous power constraint $0 \leq P_l \leq P_{peak}$ or the discrete ON/OFF power constraint $P_l \in \{0, P_{peak}\}$.

A. Minimum Power For Stability

Here we characterize the minimum average power required to stabilize the system. We begin with a precise definition of stability in terms of the *overflow function* $g(M)$ associated with a queue with unfinished work process $U(t)$:

$$g(M) \triangleq \limsup_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t} Pr[U(\tau) > M]$$

The function $g(M)$ represents the largest limiting fraction of time the unfinished work is above the value M.[2]

Definition 1: A queue with unfinished work process $U(t)$ is *stable* if $g(M) \to 0$ as $M \to \infty$. A *network of queues is stable* if all individual queues are stable.

Assume that inputs and channel processes are ergodic with arrival rates $\vec{\lambda} = (\lambda_l)$ and channel probabilities $\pi_{\vec{S}}$. In [1], the network capacity region Λ is defined as the closure of the set of all rate vectors stabilizable under some power allocation algorithm that conforms to the power constraint $\vec{P}(t) \in \Pi$. The following theorem specifies the minimum average power required for network stability, among the class of all algorithms with complete knowledge of future events.

Theorem 1: (Minimum Power for Stability) If the network is stabilizable (so that $\vec{\lambda} \in \Lambda$), the minimum power required for stability is given by P_{av}^*, where P_{av}^* is the solution to the following nonlinear optimization problem (defined in terms of auxiliary probability variables $\alpha_k^{\vec{S}}$ and power vectors $\vec{P}_k^{\vec{S}}$ for all $\vec{S}$ and for $k \in \{1, \ldots, L+2\}$):

$$\begin{aligned} \text{Minimize:} \quad & P_{av} = \textstyle\sum_{\vec{S}} \pi_{\vec{S}} \sum_{k=1}^{L+2} \alpha_k^{\vec{S}} \vec{1}' \vec{P}_k^{\vec{S}} \\ \text{Subject to:} \quad & \vec{\mu}_{av} \triangleq \textstyle\sum_{\vec{S}} \pi_{\vec{S}} \sum_{k=1}^{L+2} \alpha_k^{\vec{S}} \vec{\mu}(\vec{P}_k^{\vec{S}}, \vec{S}) \geq \vec{\lambda} \\ & \vec{P}_k^{\vec{S}} \in \Pi \, , \alpha_k^{\vec{S}} \geq 0 \quad \text{for all } k, \vec{S} \\ & \textstyle\sum_{k=1}^{L+2} \alpha_k^{\vec{S}} = 1 \quad \text{for all } \vec{S} \end{aligned}$$

That is, minimum power for stability is achieved among the class of stationary policies that measure the current channel state $\vec{S}(t)$ and then randomly allocate a power vector $\vec{P}_k^{\vec{S}}$ with probability $\alpha_k^{\vec{S}}$. The value of P_{av}^* is the resulting average power of this stationary policy, and $\vec{\mu}_{av}$ is the resulting time average transmission rate vector. This is expressed in the following corollary to Theorem 1.

Corollary 1: Minimum power for stability is given by the value P_{av}^*, minimized over the class of all stationary randomized algorithms yielding:

$$\sum_l \mathbb{E}\left\{P_l(t)\right\} = P_{av}^*$$

$$\vec{\mu}_{av} \triangleq \mathbb{E}\left\{\vec{\mu}\left(\vec{P}(t), \vec{S}(t)\right)\right\} \geq \vec{\lambda} \tag{2}$$

Theorem 1 is proven via the following two claims: (Claim 1) No algorithm can achieve stability with a smaller average power P_{av}^*, and (Claim 2) any rate vector $\vec{\lambda}$ strictly interior to Λ can be stabilized with an average power that is arbitrarily close to P_{av}^*. Claim 1 is proven in Appendix A by extending the dimensionality of the system from L to $L+1$ and applying Caratheodory's Theorem [22]. Below we prove Claim 2:

Proof: (Claim 2) The network capacity region Λ is proven in [1] to consist of all rate vectors $\vec{\lambda}$ such that a stationary power allocation rule exists satisfying (2). The value of P_{av}^* is by definition the average power consumption, minimized over all such stationary rules. If $\vec{\lambda}$ is strictly interior to Λ, there exists a positive value ϵ such that $\vec{\lambda} + \vec{\epsilon} \in \Lambda$ (where $\vec{\epsilon}$ is the L-dimensional vector with all entries equal to ϵ). It follows that there exists a stationary power allocation rule satisfying:

$$\mathbb{E}\left\{\vec{\mu}\left(\vec{P}(t), \vec{S}(t)\right)\right\} \geq \vec{\lambda} + \vec{\epsilon} > \vec{\lambda}$$

and we define $P_{av}^*(\epsilon)$ as the minimum average power consumed by any such stationary policy. The time average transmission rate of each queue is strictly larger than the arrival rate, and hence the network is stable [1]. This holds for arbitrarily small values of ϵ, and by continuity it follows that $P_{av}^*(\epsilon) \to P_{av}^*$ as $\epsilon \to 0$. □

We note that the concept of randomization used in Theorem 1 is vitally important to treat the general case where the set $\{\vec{\mu}(\vec{P}, \vec{S}) \mid \vec{P} \in \Pi\}$ is not necessarily convex. Otherwise, optimality can be achieved by a strategy that allocates a fixed power vector $\vec{P}^{\vec{S}}$ whenever the channel is in state $\vec{S}$. Note that even if there are only two possible channel states for every link, the total number of channel state vectors is 2^L. Thus, while the above static optimization defines the minimum power level P_{av}^*, it is not practical to envision solving the optimization via standard techniques, even if the channel state probabilities $\pi_{\vec{S}}$ are fully known. In the next section we overcome this problem by developing a novel *stochastic optimization* technique.

B. An Energy-Optimal Control Algorithm

Here we develop a practical control algorithm that stabilizes the system and expends an average power that is arbitrarily close to the minimum power solution P_{av}^*. For simplicity of exposition, we assume the arrival vector $\vec{A}(t)$ is i.i.d. over timeslots with arrival rate $\mathbb{E}\{\vec{A}\} = \vec{\lambda}$, and that the channel

[2] The notion of $\lim\sup$ is used because the limit always exists, even for non-ergodic sequences. The $\lim\sup$ is equal to the regular limit when queues are Markovian [1], as is the case in this paper, and throughout we simplify exposition by considering only regular limits.

state vectors $\vec{S}(t)$ are i.i.d. over timeslots with channel probabilities $\pi_{\vec{S}}$.[3] The algorithm below uses an arbitrary control parameter $V > 0$ that affects a tradeoff in average queueing delay.

Energy-Efficient Control Algorithm (EECA): Every timeslot, observe the current levels of queue backlog $\vec{U}(t)$ and channel states $\vec{S}(t)$ and allocate a power vector $\vec{P}(t) = (P_1, \ldots, P_L)$ according to the following optimization:

$$\text{Maximize:} \quad \sum_{l=1}^{L} \left[2U_l(t)\mu_l(\vec{P}, \vec{S}(t)) - VP_l\right] \quad (3)$$
$$\text{Subject to:} \quad \vec{P} = (P_1, \ldots, P_L) \in \Pi$$

Distributed Implementation: For cell-partitioned networks, we have $\vec{\mu}(\vec{P}, \vec{S}) = (\mu_1(P_1, S_1), \ldots, \mu_L(P_L, S_L))$. In this case, the above optimization is implemented according to the following simple algorithm: Each node measures the channel state $S_l(t)$ for each of its own outgoing links l and computes a *quality value* Q_l, where Q_l is the maximum value of $2U_l(t)\mu_l(P_l, S_l(t)) - VP_l$ over either the continuous interval $0 \leq P_l \leq P_{peak}$ or the 2-valued set $P_l \in \{0, P_{peak}\}$. Define $\tilde{P}_l$ as the *quality maximizing power level* for link l. Define Ω_n as the set of links $l \in \{1, \ldots, L\}$ such that $tran(l) = n$. Each node n then computes l_n^* and Q_n^*, defined as follows:

$$l_n^* \triangleq \arg\max_{l \in \Omega_n} Q_l \, , \quad Q_n^* \triangleq Q_{l_n^*}$$

The value of Q_n^* is the contribution that node n brings to the summation in (3) if it is chosen for transmission. Each node then broadcasts its value of Q_n^* to all other nodes in its cell, and the node n with the largest Q_n^* is selected to transmit in that cell (ties are broken arbitrarily). Transmission takes place over link $l = l_n^*$, with power level $\tilde{P}_l$.

Example 1: Under the ON/OFF constraint $P_l \in \{0, P_{peak}\}$, the power $\tilde{P}_l$ for each link l is given by:

$$\tilde{P}_l = \begin{cases} P_{peak} & \text{if } 2U_l(t)\mu_l(P_{peak}, S_l(t)) > VP_{peak} \\ 0 & \text{else} \end{cases}$$

In this case, we see that power is allocated only when the backlog exceeds a channel state dependent threshold.

Example 2: Suppose we have a continuous constraint $0 \leq P_l \leq P_{peak}$ and that rate functions have a logarithmic profile: $\mu_l(P, S) = \log(1 + \gamma_S P)$, where γ_S is an attenuation/noise coefficient associated with channel state S. In this case, the optimal power level is a continuous function of the queue backlog. Indeed, for any link l with channel state $S_l(t) = S$ and queue backlog $U_l(t) = U$, the quality maximizer $\tilde{P}_l$ is a critical point of $2U\mu_l(P, S) - VP$ over the interval $0 \leq P \leq P_{peak}$. Differentiating with respect to power, we have:

$$\frac{d}{dP}\left[2U\mu_l(P, S) - VP\right] = \frac{2U\gamma_S}{1 + \gamma_S P} - V$$

and it easily follows that:

$$\tilde{P}_l = \min\left[\max\left[\frac{2U_l(t)}{V} - \frac{1}{\gamma_{S_l(t)}}, 0\right], P_{peak}\right] \quad \square$$

[3] We note that the i.i.d. assumptions are not necessary, and the same algorithms can be used for general ergodic arrivals and channels, resulting in modified but more involved delay expressions [1].

To evaluate the above algorithm, define A_{max}^2, μ_{max}^{out}, and B as follows:

$$A_{max}^2 \triangleq \max_n \sum_{l \in \Omega_n} \mathbb{E}\left\{A_l^2\right\}$$
$$\mu_{max}^{out} \triangleq \max_{\{n, \vec{S}, \vec{P} \in \Pi\}} \sum_{l \in \Omega_n} \mu_l(\vec{P}, \vec{S})$$
$$B \triangleq A_{max}^2 + (\mu_{max}^{out})^2 \quad (4)$$

Now assume that $\vec{\lambda}$ is *strictly interior* to the network capacity region Λ, and define the scalar value ϵ_{max} as the largest value that can be added to each component of $\vec{\lambda}$ so that the resulting vector is still within the capacity region, i.e., $(\lambda_l + \epsilon_{max}) \in \Lambda$.

Theorem 2: If $\vec{\lambda}$ is *strictly interior* to Λ, then the EECA algorithm with any $V > 0$ stabilizes the system, with a resulting average congestion $\sum_l \overline{U}_l$ given by:

$$\sum_l \overline{U}_l \triangleq \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \left[\sum_l \mathbb{E}\left\{U_l(\tau)\right\}\right] \leq \frac{BN + VNP_{peak}}{2\epsilon_{max}}$$

Furthermore, average power P_{av} is given by:

$$P_{av} \triangleq \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \left[\sum_l \mathbb{E}\left\{P_l(\tau)\right\}\right] \leq P_{av}^* + BN/V$$

where P_{av}^* is the minimum power solution of the optimization in Theorem 1.

Thus, the V parameter can be chosen so that BN/V is arbitrarily small, yielding average power that is arbitrarily close to the optimum. However, the congestion bound grows linearly with V. By Little's Theorem, average backlog is proportional to average bit delay. Hence, power can be pushed arbitrarily close to the minimum value, with a corresponding linear increase in average delay. This holds because the V parameter effectively determines the amount by which the time average transmission rate vector $\vec{\mu}_{av}$ is larger than the input rate $\vec{\lambda}$. Pushing $\vec{\mu}_{av}$ downward towards $\vec{\lambda}$ decreases average power consumption while increasing queueing delay. Theorem 2 is proved in the next subsection using a novel drift argument.

C. Performance Analysis

To prove the performance results of the previous subsection, we first establish a novel Lyapunov drift technique enabling stability and performance optimization to be achieved simultaneously. Let $\vec{U}(t)$ be a vector of queue backlogs as a function of time. To measure aggregate network congestion, define a *Lyapunov function* $L(\vec{U})$ as the sum of squares of the individual queue backlogs: $L(\vec{U}) \triangleq \sum_l U_l^2$. Let $\vec{P}(t) = (P_1(t), \ldots, P_L(t))$ represent a process of non-negative auxiliary control variables. Let $g(\vec{P})$ be any non-negative cost function of the vector $\vec{P}$, and let g^* represent a target cost value. The goal is to stabilize the $\vec{U}(t)$ process while keeping the time average cost of $g(\vec{P}(t))$ near or below the value of g^*. (Note that if $g(\vec{P}) = \sum_l P_l$, then minimizing cost corresponds to minimizing time average power).

Lemma 1: (Lyapunov Drift with Performance Optimization) If there are positive constants V, B, ϵ such that for all timeslots t the one-step Lyapunov drift satisfies:

$$\mathbb{E}\left\{L(\vec{U}(t+1)) - L(\vec{U}(t))\right\} \leq B - \epsilon \sum_l \mathbb{E}\left\{U_l(t)\right\} + Vg^* - V\mathbb{E}\left\{g(\vec{P}(t))\right\} \quad (5)$$

then the system is stable and time average backlog satisfies:

$$\overline{\sum_l U_l} \triangleq \limsup_{M\to\infty} \frac{1}{M}\sum_{\tau=0}^{M-1}\sum_l \mathbb{E}\left\{U_l(\tau)\right\} \leq \frac{B+Vg^*}{\epsilon}$$

while time average cost satisfies:

$$\overline{g} \triangleq \limsup_{M\to\infty} \frac{1}{M}\sum_{\tau=0}^{M-1}\mathbb{E}\left\{g(\vec{P}(\tau))\right\} \leq g^* + B/V$$

From the above statement, it is clear that if the V parameter can be increased while holding all other constants fixed, then the time average cost can be pushed arbitrarily near or below the target cost level g^*, with a corresponding tradeoff in average queue backlog.

Proof: The drift condition is satisfied for all timeslots t. Summing (5) over timeslots $t \in \{0, \ldots, M-1\}$ and dividing by M yields:

$$\frac{\mathbb{E}\left\{L(\vec{U}(M)) - L(\vec{U}(0))\right\}}{M} \leq B - \frac{\epsilon}{M}\sum_{\tau=0}^{M-1}\sum_l \mathbb{E}\left\{U_l(\tau)\right\} + Vg^* - \frac{V}{M}\sum_{\tau=0}^{M-1}\mathbb{E}\left\{g(\vec{P}(\tau))\right\} \quad (6)$$

By non-negativity of the Lyapunov function and of the $g(\vec{P})$ function, a simple manipulation of (6) yields:

$$\frac{1}{M}\sum_{\tau=0}^{M-1}\sum_l \mathbb{E}\left\{U_l(\tau)\right\} \leq \frac{B+Vg^*+\mathbb{E}\left\{L(\vec{U}(0))\right\}/M}{\epsilon}$$

Taking limits of the above inequality as $M \to \infty$ yields the time average backlog bound. In [1], it is shown that this time average backlog bound implies system stability.

Similarly, by again manipulating (6) we obtain:

$$\frac{1}{M}\sum_{\tau=0}^{M-1}\mathbb{E}\left\{g(\vec{P}(\tau))\right\} \leq g^* + \frac{B}{V} + \frac{\mathbb{E}\left\{L(\vec{U}(0))\right\}}{VM} \quad (7)$$

Taking limits as $M \to \infty$ yields the result. □

The art of stochastic optimal networking is designing a strategy to ensure the drift condition of Lemma 1 is satisfied. In the remainder of this section, we illustrate the technique with a constructive proof of Theorem 2. The first step is to establish a general expression for Lyapunov drift under any power allocation policy.

Lemma 2: If arrivals $\vec{A}(t)$ are i.i.d. every slot with rates $\mathbb{E}\{\vec{A}(t)\} = \vec{\lambda} = (\lambda_1, \ldots, \lambda_L)$, then the conditional Lyapunov drift under any power allocation policy satisfies:

$$\Delta(\vec{U}(t)) \triangleq \mathbb{E}\left\{L(\vec{U}(t+1) - L(\vec{U}(t)) \mid \vec{U}(t)\right\} \leq BN - 2\sum_l U_l(t)\left[\mathbb{E}\left\{\mu_l(\vec{P}(t), \vec{S}(t)) \mid \vec{U}(t)\right\} - \lambda_l\right] \quad (8)$$

where B is defined in (4). □

The lemma follows simply by squaring the dynamical queueing equation (1) and taking expectations, and is proved in Appendix B. We now massage the right hand side of (8) into a form suitable for application of Lemma 1 by adding and subtracting the same value. We have:

$$\Delta(\vec{U}(t)) \leq BN - 2\sum_l U_l(t)\left[\mathbb{E}\left\{\mu_l(\vec{P}(t), \vec{S}(t)) \mid \vec{U}(t)\right\} - \lambda_l\right] + V\sum_l \mathbb{E}\left\{P_l(t) \mid \vec{U}(t)\right\} - V\sum_l \mathbb{E}\left\{P_l(t) \mid \vec{U}(t)\right\}$$

Rearranging terms on the right hand side yields:

$$\Delta \leq BN - V\sum_l \mathbb{E}\left\{P_l(t) \mid \vec{U}(t)\right\} + 2\sum_l U_l(t)\lambda_l - \mathbb{E}\left\{\sum_l \left[2U_l(t)\mu_l(\vec{P}(t), \vec{S}(t)) - VP_l(t)\right] \mid \vec{U}(t)\right\} \quad (9)$$

The design principle behind the EECA algorithm of section III-B is now apparent: *The EECA algorithm (3) was designed to minimize the value of the final term in the above expression (9) over all possible power allocation strategies.*

Suppose now that $\vec{\lambda}$ is strictly interior to the capacity region Λ, and let ϵ be a positive value such that $\vec{\lambda} + \vec{\epsilon} \in \Lambda$. From Corollary 1, it follows that there exists a stationary randomized power allocation strategy that chooses power independent of queue backlog, and such that:

$$\sum_l \mathbb{E}\left\{P_l(t)\right\} = P_{av}(\epsilon) \quad (10)$$

$$\mathbb{E}\left\{\mu_l\left(\vec{P}(t), \vec{S}(t)\right)\right\} \geq \lambda_l + \epsilon \quad \text{(for all } l\text{)} \quad (11)$$

where $P_{av}(\epsilon)$ is the minimum power required to stabilize the data rates $\vec{\lambda} + \vec{\epsilon}$. Note that $P_{av}(\epsilon) \to P_{av}^*$ as $\epsilon \to 0$. Because this stationary rule is simply a particular power allocation strategy, the final term in (9) under the EECA algorithm is less than or equal to the resulting value under the stationary rule. However, this value in (9) under the stationary rule can be explicitly calculated using (10) and (11), and we have:

$$\Delta(\vec{U}(t)) \leq BN - V\sum_l \mathbb{E}\left\{P_l(t) \mid \vec{U}(t)\right\} + 2\sum_l U_l(t)\lambda_l - \left[2\sum_l U_l(t)(\lambda_l + \epsilon) - VP_{av}(\epsilon)\right]$$

Canceling the $U_l(t)\lambda_l$ terms in the above expression and taking expectations over the distribution of $\vec{U}(t)$ yields:

$$\mathbb{E}\left\{L(\vec{U}(t+1)) - L(\vec{U}(t))\right\} \leq BN - 2\epsilon\sum_l \mathbb{E}\left\{U_l(t)\right\} + VP_{av}(\epsilon) - V\sum_l \mathbb{E}\left\{P_l(t)\right\}$$

The above expression is in the exact form specified in Lemma 1 in the case $g(\vec{P}) = \sum_l P_l$. It follows that time average unfinished work satisfies:

$$\overline{\sum_l U_l} \leq \frac{BN + VP_{av}(\epsilon)}{2\epsilon} \leq \frac{BN + VNP_{peak}}{2\epsilon} \quad (12)$$

and time average power satisfies:

$$P_{av} = \sum_l \overline{P}_l \leq P_{av}(\epsilon) + BN/V \quad (13)$$

The performance bounds in (12) and (13) hold for any value $\epsilon > 0$ such that $\vec{\lambda} + \vec{\epsilon} \in \Lambda$. However, the particular choice of ϵ only affects the bound calculation and does not affect the EECA allocation policy or change any sample path of system dynamics. We can thus optimize the bounds in (12) and (13) separately over all possible ϵ values. The bound in (13) is clearly minimized by taking a limit as $\epsilon \to 0$, yielding: $\sum_l \overline{P}_l \leq P_{av}^* + BN/V$. Conversely, the bound in (12) is minimized by considering the largest feasible ϵ such that $\vec{\lambda} + \vec{\epsilon} \in \Lambda$ (defined as ϵ_{max}), yielding: $\overline{\sum_l U_l} \leq (BN + VNP_{peak})/(2\epsilon_{max})$. This proves Theorem 2.

IV. Average Power Constraints

In this section we consider a related problem of maximizing network throughput subject to both peak and average power constraints. Specifically, we consider the same one-hop network of the previous section, but assume that each node $n \in \{1, \ldots, N\}$ must satisfy the average power constraint:

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \left[\sum_{l\in\Omega_n} P_l(\tau) \right] \leq P_{av}^n \quad (14)$$

Using a proof similar to that given in Theorem 1, it can be shown that the new capacity region Λ reduces to the set of all rates $\vec{\lambda}$ for which there exists a stationary randomized power allocation scheme such that (2) is satisfied, and such that the additional constraints $\mathbb{E}\left\{\sum_{l\in\Omega_n} P_l(t)\right\} \leq P_{av}^n$ are satisfied for all $n \in \{1, \ldots, N\}$. Here we consider cases where the arrival rate vector $\vec{\lambda}$ is *either inside the capacity region or outside of the capacity region*. This requires an additional set of admission control decisions to be made on top of the power allocation decisions, as only a fraction of the arriving traffic can be successfully delivered if inputs exceed capacity (see Fig. 3).

Let $R_l(t)$ represent the packets accepted into the network at queue l on timslot t (where $R_l(t) \leq A_l(t)$, that is, $R_l(t)$ is the fraction of new arrivals that are accepted on slot t, where the remaining data is dropped). Define $\overline{R}_l \triangleq \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mathbb{E}\left\{R_l(\tau)\right\}$ as the long term expected admission rate into queue l, and let $\vec{R}_{av} = (\overline{R}_1, \ldots, \overline{R}_L)$. The goal is to design a joint strategy for power allocation and admission control that satisfies all power constraints while maximizing the weighted throughput metric $\sum_l \theta_l \overline{R}_l$ (where θ_l values are arbitrary positive weights) subject to the demand requirement $\vec{R}_{av} \leq \vec{\lambda}$ and the stability requirement $\vec{R}_{av} \in \Lambda$. Define $\vec{R}^*$ as the optimal admission rate vector for this problem. This optimum could in principle be computed if the arrival rates $\vec{\lambda}$ and the capacity region Λ were known in advance. Below we design a practical algorithm that performs arbitrarily close to the utility of $\vec{R}^*$.

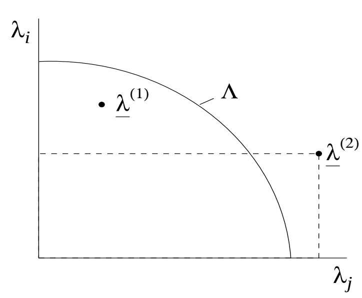

Fig. 3. A capacity region Λ (illustrated in 2 dimensions) with a rate vector $\vec{\lambda}^{(1)}$ strictly in the interior. The rate vector $\vec{\lambda}^{(2)}$ is outside of the capacity region.

A. The Virtual Power Queue

We first establish a novel mechanism for ensuring the average power constraints are met at every node. To this end, each node n maintains a *virtual power queue* with occupancy $X_n(t)$ equal to the maximum excess power expended beyond the average power constraint over any interval ending at slot t. Indeed, defining $X_n(0) = 0$, we propagate the $X_n(t)$ values as follows:

$$X_n(t+1) = \max[X_n(t) - P_{av}^n, 0] + \sum_{l\in\Omega_n} P_l(t) \quad (15)$$

Thus, the $X_n(t)$ process acts as a single server queue with constant server rate given by the average power constraint P_{av}^n, with 'arrivals' given by the total power allocated for outgoing transmissions of node n on the current timeslot. The intuition behind this construction is given by the following observation: *If a power allocation algorithm conforms to the power constraint* $\vec{P}(t) \in \Pi$ *for all* t *while stabilizing all actual queues* $U_l(t)$ *and all virtual queues* $X_n(t)$ *(for* $l \in \{1, \ldots, L\}$, $n \in \{1, \ldots, N\}$*), then the strategy also satisfies the average power constraints for each node.* This observation holds because if the excess backlog $X_n(t)$ in virtual power queue n is stabilized, it must be the case that the time average 'power arrivals' $\sum_{l\in\Omega_n} \overline{P}_l$ (corresponding to time average power expenditure in node n) is less than or equal to the 'service rate' P_{av}^n.

B. An Energy Constrained Control Algorithm (ECCA)

We use the virtual power queues in the following energy constrained control algorithm. Assume the weights θ_l are known to the controllers, and let $V > 0$ represent an arbitrary control parameter.

Admission Control: Every timeslot and for each queue l, we allow the full set of new arrivals $A_l(t)$ into the queue whenever $U_l(t) \leq V\theta_l/2$. Else, we drop all new arrivals for queue l entering on that timeslot.

Power Allocation: Allocate power $\vec{P}(t) = \vec{P}$ according to the following optimization:

$$\text{Max: } \sum_{n=1}^N \sum_{l\in\Omega_n} \left[U_l(t)\mu_l(\vec{P}, \vec{S}(t)) - X_n(t)P_l \right] \quad (16)$$

$$\text{Subject to: } \vec{P} \in \Pi$$

The virtual power queues $X_n(t)$ are then updated via (15).

Note that distributed implementation of this algorithm for the case $\vec{\mu}(\vec{P}, \vec{S}) = (\mu_1(P_1, S_1), \ldots, \mu_L(P_L, S_L))$ is similar to the implementation of EECA given in Section III-B. The only difference here is that the quality maximizing values $\tilde{P}_l$ are computed by using the value $2X_n(t)$ instead of the scalar V [compare (16) and (3)]. To simplify the analysis of the above algorithm, we additionally assume that the total arrivals to any node are bounded by a constant value A_{max} every timeslot, that is, $\sum_{l\in\Omega_n} A_l(t) \leq A_{max}$ for all nodes n. Further, assume that the rate functions are differentiable with respect to power, and define β as the maximum value of the partial derivative over all links, power levels, and channel states:

$$\beta = \max_{l,\vec{S},\vec{P}\in\Pi}\left[\frac{\partial}{\partial P_l}\mu_l(\vec{P},\vec{S})\right]$$

Theorem 3: For any input rate vector $\vec{\lambda}$, the above ECCA algorithm conforms to both peak and average power constraints and yields a performance bound of:

$$\sum_l \theta_l \overline{R}_l \geq \sum_l \theta_l R_l^* - \frac{(B+C)N}{V} \quad (17)$$

where: $C \triangleq P_{peak}^2 + \frac{1}{n}\sum_{n=1}^{N}(P_{av}^n)^2$

Further, for each queue l and slot t, backlog satsifies:

$$U_l(t) \leq U_l^{max} \triangleq \frac{V\theta_l}{2} + A_{max}$$

and for each node n and for all t, excess energy satisfies:

$$X_n(t) \leq X_n^{max} \triangleq \frac{\beta V}{2}\left(\max_{l\in\Omega_n}\{\theta_l\}\right) + \beta A_{max} + P_{peak}$$

Note that the queue backlog is bounded for every instant of time. Hence, the algorithm yields the same performance if all buffers are finite with buffer size $\mathit{Buffer} = V\theta_{max}/2 + A_{max}$. In systems with finite buffers, the parameter V could be defined according to this equation, resulting in performance:

$$\sum_l \theta_l \overline{R}_l \geq \sum_l \theta_l R_l^* - \frac{(B+C)N\theta_{max}}{2(\mathit{Buffer} - A_{max})}$$

and hence performance can be pushed arbitrarily close to optimality by increasing the buffer size. The excess energy bound is also very strong, and implies that the total energy expended by node n over any interval of size M is less than or equal to $MP_{av} + X_n^{max}$. It is remarkable that these performance guarantees do not depend on the channel statistics or arrival rates.

C. Performance Analysis

We analyze the above strategy in a manner similar to the EECA algorithm of the previous section. In particular, the $R_l(t)$ variables play the role of packet arrivals $A_l(t)$:

$$U_l(t+1) = \max[U_l(t) - \mu_l(\vec{P}(t),\vec{S}(t)), 0] + R_l(t)$$

The virtual queue backlogs $\vec{X}(t)$ evolve according to (15). Define the Lyapunov function $L(\vec{U},\vec{X}) = \sum_l U_l^2 + \sum_n X_n^2$, and define the one-step drift:

$$\Delta(\vec{U}(t),\vec{X}(t)) \triangleq$$
$$\mathbb{E}\left\{L(\vec{U}(t+1),\vec{X}(t+1)) - L(\vec{U}(t),\vec{X}(t)) \mid \vec{U}(t),\vec{X}(t)\right\}$$

To simplify formulas, below we use the shortened notation $\Delta, \mu_l, \vec{U}$, and $\vec{X}$ to represent $\Delta(\vec{U}(t))$, $\mu_l(\vec{P}(t),\vec{S}(t))$, $\vec{U}(t)$, and $\vec{X}(t)$.

Lemma 3: The one-step drift satisfies:

$$\Delta \leq N(B+C) - 2\sum_l U_l \mathbb{E}\left\{\mu_l - R_l \mid \vec{U},\vec{X}\right\}$$
$$-2\sum_n X_n\left[P_{av}^n - \mathbb{E}\left\{\sum_{l\in\Omega_n} P_l \mid \vec{U},\vec{X}\right\}\right]$$

The lemma follows by summing the corresponding drift of the actual queues and virtual queues (compare with Lemma 2), and the derivation is omitted for brevity. Adding and subtracting the optimization metric $V\sum_l \theta_l \mathbb{E}\left\{R_l \mid \vec{U},\vec{X}\right\}$ to the right hand side of the drift expression and rearranging terms yields:

$$\Delta \leq N(B+C) + V\sum_l \theta_l \mathbb{E}\left\{R_l \mid \vec{U},\vec{X}\right\} - 2\sum_n X_n P_{av}^n$$
$$+\sum_l (2U_l - V\theta_l)\mathbb{E}\left\{R_l \mid \vec{U},\vec{X}\right\}$$
$$-2\sum_{n=1}^{N}\sum_{l\in\Omega_n}\mathbb{E}\left\{U_l\mu_l - X_n P_l \mid \vec{U},\vec{X}\right\}$$

The design methodology of the ECCA algorithm is now apparent: *The admission control algorithm minimizes the second to last term of the above expression over all possible admission decisions, and the power allocation algorithm minimizes the last term of the above expression over all possible power decisions.*

In particular, the optimal input rate vector $\vec{R}^* = (R_1^*, \ldots, R_L^*)$ could in principle be achieved by the simple backlog-independent admission control algorithm of including all new arrivals $A_l(t)$ for a given link l and slot t independently with probability $\alpha_l = R_l^*/\lambda_l$, yielding:

$$\mathbb{E}\left\{R_l \mid \vec{U},\vec{X}\right\} = \mathbb{E}\left\{R_l\right\} = \alpha_l \mathbb{E}\left\{A_l\right\} = R_l^* \quad (18)$$

Likewise, because $\vec{R}^* \in \Lambda$, there must exist a stationary power allocation policy that chooses power independent of backlog and yields:

$$\mathbb{E}\left\{\mu_l \mid \vec{U},\vec{X}\right\} = \mathbb{E}\left\{\mu_l\right\} = R_l^* \quad (19)$$

$$\mathbb{E}\left\{\sum_{l\in\Omega_n} P_l \mid \vec{U},\vec{X}\right\} = \mathbb{E}\left\{\sum_{l\in\Omega_n} P_l\right\} \leq P_{av}^n \quad (20)$$

Plugging in the expectations (18)-(20) of the particular backlog-independent policies into the last two terms of the above drift expression for the ECCA algorithm thus preserves the bound, and yields:

$$\Delta(\vec{U}(t),\vec{X}(t)) \leq N(B+C)$$
$$+V\textstyle\sum_l \theta_l \mathbb{E}\left\{R_l(t) \mid \vec{U},\vec{X}\right\} - V\sum_l \theta_l R_l^* \quad (21)$$

where we have canceled the common terms $\sum_l 2U_l R_l^*$ and $\sum_n X_n P_{av}^n$. Taking expectations of (21) with respect to $\vec{U},\vec{X}$ and summing from $t=0$ to $t=M-1$ yields:

$$\frac{1}{M}\sum_{\tau=0}^{M-1}\sum_l \theta_l \mathbb{E}\left\{R_l(\tau)\right\} \geq \sum_l \theta_l R_l^* - \frac{N(B+C)}{V}$$
$$-\mathbb{E}\left\{L(\vec{U}(0),\vec{X}(0))\right\}/(MV)$$

which yields (17) as $M \to \infty$.

Furthermore, the backlog bound $U_l(t) \leq U_l^{max}$ follows immediately from the definition of the ECCA admission control policy: No new arrivals are admitted if $U_l(t) > V\theta_l/2$, so that $U_l(t) \leq V\theta_l/2 + A_{max}$ for all t (where in the worst case we add an amount A_{max} when backlog is exactly at the $V\theta_l/2$ threshold).

Likewise, by definition of the ECCA power allocation algorithm, we clearly set $P_l = 0$ if the partial derivative of $U_l(t)\mu_l(\vec{P}, \vec{S}(t))$ with respect to P_l is smaller than $X_n(t)$ for all $\vec{P} \in \Pi$. The largest possible value for this derivative is $\beta \max_{l\in\Omega_n}\{U_l^{max}\}$, and hence no power is allocated to any outgoing link of node n if $X_n(t) > \beta \max_{l\in\Omega_n}\{U_l^{max}\}$. It follows that $X_n(t) \leq \beta \max_{l\in\Omega_n}\{U_l^{max}\} + P_{peak}$ always, proving Theorem 3.

V. MULTI-HOP NETWORKS

Here we consider the same network as before but assume that data can be routed over multi-hop paths to reach its destination (we assume the network is *connected*, so this is always possible, see Fig. 1). We optimize over all possible power allocation and routing algorithms. Thus, incoming data is not necessarily associated with any particular link, and so we redefine the arrival processes in terms of the origin and destination of the data: $A_n^c(t) \triangleq$ amount of data exogenously arriving to node n at slot t that is destined for node c. All data (from any source node) that is destined for a particular node $c \in \{1, \ldots, N\}$ is defined as *commodity* c data. Data is stored in each node according to its destination, and we let $U_n^c(t)$ represent the current backlog of commodity c data in node n.

Suppose power vector $\vec{P}(t)$ is allocated in slot t, so that the transmission rate over a link l is $\mu_l(\vec{P}(t), \vec{S}(t))$. A *routing decision* must be made to establish which commodity to transfer over link l. In general, multiple commodities could be transfered over the same link simultaneously,[4] and we define routing variables $\mu_l^c(t)$ as the rate allocated to commodity c data over link l during slot t. The problem is to allocate power every timeslot according to the power constraint $\vec{P}(t) \in \Pi$ and then to route data according to the *link rate constraint*:

$$\sum_{c=1}^{N} \mu_l^c(t) \leq \mu_l(\vec{P}(t), \vec{S}(t)) \tag{22}$$

Recall that Ω_n is the set of all links l such that $tran(l) = n$. Further define Θ_n as the set of all links l such that $rec(l) = n$. The resulting 1-step queueing equation for backlog $U_n^c(t)$ thus satisfies (for $c \neq n$):

$$U_n^c(t+1) \leq \max[U_n^c(t) - \sum_{l\in\Omega_n} \mu_l^c(t), 0] + A_n^c(t) + \sum_{l\in\Theta_n} \mu_l^c(t) \tag{23}$$

The above expression is an inequality rather than an equality because the incoming commodity c data to node n may be less than $\sum_{l\in\Theta_n} \mu_l^c(t)$ if the corresponding transmitting nodes have little or no data of this commodity waiting to be transfered.

In [14], the network layer capacity region Λ is defined as the closure of the set of rate matrices (λ_{nc}) that can be stably supported, considering all possible power allocation and routing strategies. There, it was shown that any rate matrix $(\lambda_{nc}) \in \Lambda$ is supportable via a randomized algorithm for choosing power allocations $\vec{P}(t)$ and routing variables $\mu_l^c(t)$. Here, we assume the rate matrix is inside the capacity region, and develop an energy efficient stabilizing algorithm. However, note that the objective of minimizing average power expenditure in a multihop network may place an unfair power burden on centrally located nodes that are used by many others. Thus, to balance power more evenly, we consider the more general objective of minimizing the time average of $\sum_n g_n(\sum_{l\in\Omega_n} P_l(t))$, where $g_n(p)$ is any convex increasing cost function of the power expended by node n.

Define $U_n^n(t) = 0$ for all t, and define:

$$A_{max}^2 \triangleq \max_n \sum_c \mathbb{E}\left\{(A_n^c)^2\right\}$$

$$\mu_{max}^{in} \triangleq \max_{\{n, \vec{S}, \vec{P}\in\Pi\}} \sum_{l\in\Theta_n} \mu_l(\vec{P}, \vec{S})$$

$$D \triangleq (A_{max} + \mu_{max}^{in})^2 + (\mu_{max}^{out})^2$$

Let $L(\underline{U}) = \sum_{n,c}(U_n^c)^2$. The one-step drift $\Delta(\underline{U})$ for any policy is found by squaring the dynamical equation (23) as in Lemma 2, and is given in [14] as follows: If arrivals and channel states are i.i.d. over timeslots, then

Lemma 4: $\Delta(\underline{U}(t)) \leq DN + 2\sum_{n,c} U_n^c(t)\lambda_{nc}$

$$-V\sum_n \mathbb{E}\left\{g_n(\sum_{l\in\Omega_n} P_l(t)) \mid \underline{U}(t)\right\}$$

$$-\sum_n \mathbb{E}\Big\{\sum_{l\in\Omega_n}\sum_c 2\mu_l^c(t)(U_{tran(l)}^c(t) - U_{rec(l)}^c(t)) - Vg_n(\sum_{l\in\Omega_n} P_l(t)) \Big| \underline{U}(t)\Big\}$$

The above drift expression for multi-hop networks is the same as that given in [14], with the exception that we have added and subtracted the optimization metric $V\sum_n \mathbb{E}\left\{g_n(\sum_{l\in\Omega_n} P_l(t)) \mid \underline{U}(t)\right\}$. Minimizing the last term in the above drift expression over all power allocations satisfying $\vec{P} \in \Pi$ and all routing strategies satisfying (22) leads to the following *multi-hop EECA algorithm*:

1) For all links l, find the commodity $c_l^*(t)$ such that:

$$c_l^*(t) = \arg\max_c \left\{U_{tran(l)}^c(t) - U_{rec(l)}^c(t)\right\}$$

and define:

$$W_l^*(t) = \max[U_{tran(l)}^{c_l^*}(t) - U_{rec(l)}^{c_l^*}(t), 0]$$

2) *Power Allocation:* Choose a power vector $\vec{P}(t) \in \Pi$ that maximizes:

$$\sum_n \left[\sum_{l\in\Omega_n} 2\mu_l(\vec{P}, \vec{S}(t))W_l^* - Vg_n(\sum_{l\in\Omega_n} P_l)\right] \tag{24}$$

3) *Routing:* Over link l, transmit commodity c_l^* in the amount of $\min[U_l^c(t), \mu_l(\vec{P}(t), \vec{S}(t))]$.

Distributed Implementation: Given a cell partitioned network with backlog values $U_a^c(t)$ for all neighbor nodes a, the distributed method for allocating power and choosing which node transmits in every cell is similar to the implementation of EECA in Section III-B, with the exception that the quality maximizer values $\tilde{P}_l$ now maximize

[4] We find that the capacity achieving solution needs only route a single commodity over any link during a timeslot.

$\left[2\mu_l(P_l, S_l(t))W_l^* - Vg_{tran(l)}(P_l)\right]$, representing the contribution to (24) if link l is chosen for transmission (recall that a given node n may activate only one outgoing link $l \in \Omega_n$ during a timeslot).

To find the backlog values of neighbors, note that for rectilinear networks there are at most 10 queues that change their backlog values during a timeslot in any given cell. This is because the transmitting node may transmit to another node in the same cell (increasing the queue level of the transmitted commodity in the receiving node, and decreasing it in the transmitting node), and there are at most 8 other data receptions in the same cell (due to potential tranmissions from the 8 adjacent cells). Knowledge of backlog levels in neighboring nodes can thus be maintained by broadcasting the backlog changes to all nodes in the same cell and in adjacent cells. Each update requires a triplet of information: (n, c, δ), where n is the node, c is the commodity that was changed, and δ is the amount of the change. Thus, the bandwidth of the broadcast control channel must be sufficient to support the transmission of up to 10 update triplets per cell per timeslot.

Theorem 4: If the rate matrix (λ_{nc}) is interior to the capacity region Λ, then the above multihop EECA algorithm for routing and power allocation stabilizes the network and yields a time average congestion bound of:

$$\sum_{nc} \overline{U}_n^c \leq \frac{DN + V\sum_n g_n(P_{peak})}{2\epsilon_{max}}$$

(where ϵ_{max} is the largest ϵ such that $(\lambda_{nc} + \epsilon) \in \Lambda$).

Further, the time average cost satisfies:

$$\sum_n \overline{g}_n \triangleq \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\left[\sum_n g_n(\sum_{l\in\Omega_n} P_l(\tau))\right] \leq g^* + \frac{DN}{V}$$

where g^* represents the minimum time average cost of any stabilizing policy.

The proof is similar to the proof of Theorem 2, and so we present only an outline: The dynamic algorithm minimizes the final term of Lemma 4 over all policies. In [14] it is shown that if there is an ϵ such that $(\lambda_{nc} + \epsilon) \in \Lambda$, then a single stationary power allocation and routing strategy can be developed to satisfy:

$$\sum_l \sum_c \mathbb{E}\left\{\mu_l^c(t)\right\}(U_{tran(l)}^c - U_{rec(l)}^c) = \sum_{n,c} U_n^c(\lambda_{nc} + \epsilon)$$

for all non-negative U_n^c values. Further, the stationary policy also satisfies $\sum_n V\mathbb{E}\left\{g(\sum_{l\in\Omega_n} P_l(t))\right\} = g^*(\epsilon)$, where $g^*(\epsilon)$ is the minimum cost for stabilizing rates $(\lambda_{nc}+\epsilon)$ and satisfies $g^*(\epsilon) \to g^*$ as $\epsilon \to 0$. Plugging these particular policies into the last term of the drift expression in Lemma 4 thus preserves the bound and yields:

$$\Delta(\underline{U}(t)) \leq DN - 2\sum_{n,c} U_n^c(t)\epsilon$$
$$-V\sum_n \mathbb{E}\left\{g_n(\sum_{l\in\Omega_n} P_l(t)) \mid \underline{U}(t)\right\} + Vg^*$$

which yields the result upon application of Lemma 1. □

We note that the multi-hop EECA algorithm delivers all data to its destination *without knowing the network topology*. The algorithm effectively accomplishes this by expending initial energy transmitting data to neighbors in order to learn efficient routes, which emerge from backlog information.

VI. SIMULATIONS

For brevity, we present simulation results only for the simple two-queue downlink example of Section II. Packets arrive to the system according to Poisson processes with rates $\lambda_1 = 8/9, \lambda_2 = 5/9$, which are the same as the empirical rates obtained by averaging over the first 9 timeslots of the example in Fig. 2. Channel states arise as i.i.d. vectors $(S_1(t), S_2(t))$ every slot. The probability of each vector state is matched to the empirical occurrence frequency in the example, so that $Pr[(G, M)] = 3/9, Pr[(M, B)] = 2/9, Pr[(M, M)] = 1/9$, etc. We first simulate the policy of serving the queue with the largest rate-backlog index $U_i(t)\mu_i(t)$, a strategy that stabilizes the system whenever possible but does not necessarily make energy efficient decisions [20] [15] [13]. The simulation was run for 10 million timeslots. The resulting average power is $P_{av} = 0.898$ Watts, and the resulting time average backlog is 2.50 packets.

Next, we consider the EECA algorithm, where power allocation decisions are determined by the solution of the optimization problem (3). First note that $A_{max}^2 = \sum_{i=1}^2 \lambda_i^2 + \lambda_i = 2.54$, $\mu_{max}^{out} = 3$, and hence from (4) we have $B = 11.54$. It follows from Theorem 2 that the resulting average power differs from optimality by no more than $11.54/V$, where V is the control parameter of the algorithm (note that $N = 1$ in this example). Furthermore, it can be shown that $\epsilon_{max} = 0.489$ for this example, and hence by Theorem 2 we know the average backlog in the system satisfies the following inequality:

$$\overline{U}_1 + \overline{U}_2 \leq \frac{11.54 + V}{0.978}$$

By Little's Theorem, dividing both sides of the above inequality by $(\lambda_1 + \lambda_2)$ yields an upper bound on average delay.

We simulated the EECA algorithm for 20 different values of the control parameter V, ranging from 1 to 10^4. Each simulation was run for 10 million timeslots. In Fig. 4 the resulting average power is plotted against the time average backlog. The corresponding upper bound is also shown in the figure. We find that average power decreases to its minimum value of 0.518 Watts as the control parameter V is increased,

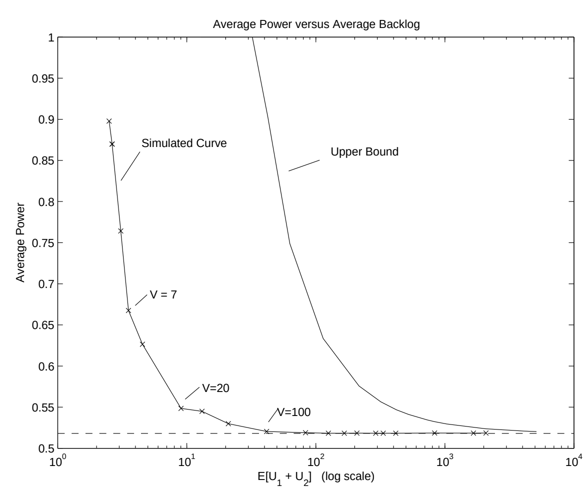

Fig. 4. Average power versus average backlog for a two queue downlink under the EECA algorithm.

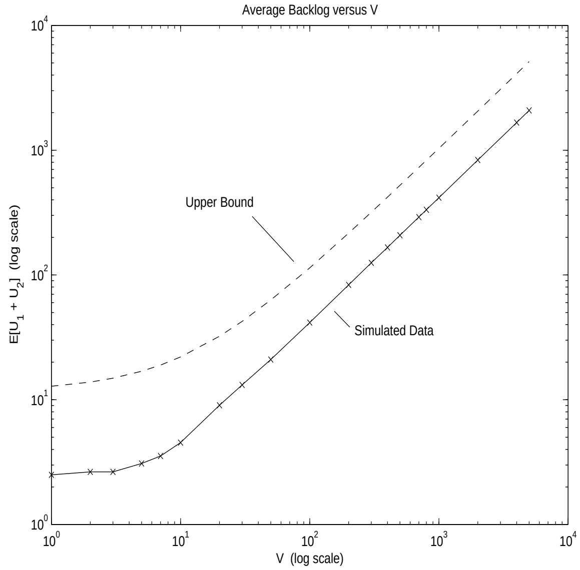

Fig. 5. Average backlog versus the V parameter from 10 million iterations of the EECA algorithm for a two queue downlink. The analytical upper bound is also plotted.

with a corresponding tradeoff in average delay. In Fig. 5 we plot average backlog versus the V parameter together with the backlog bound, illustrating that average delay grows linearly in V, as suggested by the performance bound. As a point of reference, we note that at $V = 50$, the average power is 0.53 Watts and the average sum backlog is 21.0 packets.

VII. Conclusions

We have developed energy-efficient control strategies with performance that can be pushed arbitrarily close to optimal, with a corresponding tradeoff in average network delay. Our algorithms adapt to local link conditions without requiring knowledge of traffic rates, channel statistics, or global network topology. For simplicity of exposition, channels were modeled as being independent from slot to slot. However, the algorithms can be shown to yield similar results for more general channel processes, and are robust to situations when channel statistics or traffic loadings change over time [1]. The analysis presented here uses a new Lyapunov drift technique enabling stability and performance optimization to be achieved simultaneously. This research creates a general framework for designing practical control algorithms that are provably optimal.

Appendix A — Minimum Power for Stability

Here we prove Claim 1 of Theorem 1: *Consider any allocation rule for choosing $\vec{P}(t)$ subject to $\vec{P}(t) \in \Pi$, perhaps one that uses full knowledge of future arrivals and channel states. If the rule stabilizes the system, then:*

$$\underline{P}_{av} \triangleq \liminf_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\left[\sum_{l} P_l(\tau)\right] \geq P_{av}^* \quad (25)$$

where P_{av}^* is the minimum power obtained from the optimization in Theorem 1.

To prove (25), we first establish some convenient notation. For each $\vec{S}$, define $T_{\vec{S}}(M)$ as the set of timeslots $t \in \{0, \ldots, M\}$ during which the channel state vector is equal to $\vec{S}$, and let $||T_{\vec{S}}(M)||$ represent the total number of such slots. Define the conditional empirical average of transmission rate and power consumption as follows:

$$\left(\vec{\mu}_{av}^{\vec{S}}(M); P_{av}^{\vec{S}}(M)\right) \triangleq \sum_{\tau \in T_{\vec{S}}(M)} \frac{\left(\vec{\mu}(\vec{P}(\tau), \vec{S}); \vec{1}'\vec{P}(\tau)\right)}{||T_{\vec{S}}(M)||}$$

Lemma 5: For every M, there exist probabilities $\alpha_k^{\vec{S}}[M]$ and power vectors $\vec{P}_k^{\vec{S}}[M] \in \Pi$ such that:

$$\vec{\mu}_{av}^{\vec{S}}(M) = \sum_{k=1}^{L+2} \alpha_k^{\vec{S}}(M)\vec{\mu}\left(\vec{P}_k^{\vec{S}}(M), \vec{S}\right) \quad (26)$$

$$P_{av}^{\vec{S}}(M) = \sum_{k=1}^{L+2} \alpha_k^{\vec{S}}(M)\vec{1}'\vec{P}_k^{\vec{S}}(M) \quad (27)$$

Proof: Define $\vec{\Phi}^{\vec{S}}(\vec{P}) \triangleq \left(\vec{\mu}(\vec{P}, \vec{S}); \vec{1}'\vec{P}\right)$ as a function mapping the L dimensional power vector into $L+1$ dimensional space. Then $\frac{1}{||T_{\vec{S}}(M)||}\sum_{\tau \in T_{\vec{S}}(M)} \vec{\Phi}(\vec{P}(\tau))$ is a convex combination of points in the image of the $L+1$ dimensional function $\vec{\Phi}^{\vec{S}}(\vec{P})$ (for $\vec{P} \in \Pi$), and is therefore (by Caratheodory's theorem [22]) expressible by a convex combination of at most $L+2$ elements of the image. □

Now define:

$$(\vec{\mu}_{av}(M); P_{av}(M)) \triangleq \sum_{\vec{S}} \frac{||T_{\vec{S}}(M)||}{M}\left(\vec{\mu}_{av}^{\vec{S}}(M); P_{av}^{\vec{S}}(M)\right)$$

For brevity, we outline the rest of the proof: For each M, the number of $\alpha_k^{\vec{S}}(M)$ and $\vec{P}_k^{\vec{S}}(M)$ values is at most $(L+2)Card(\{\vec{S}\})$ (where $Card(\{\vec{S}\})$ represents the number of possible channel state vectors). By compactness, we can thus find an appropriate subsequence of integers M_k such that $M_k \to \infty$ and such that there exist limiting probabilities $\alpha_k^{\vec{S}}$ and power levels $\vec{P}_k^{\vec{S}} \in \Pi$ satisfying:

$$P_k^{\vec{S}}(M_k) \to P_k^{\vec{S}} \;,\; \alpha_k^{\vec{S}}(M_k) \to \alpha_k^{\vec{S}} \;,\; P_{av}(M_k) \to \underline{P}_{av}$$

Because channel states are ergodic, we have $\frac{||T_{\vec{S}}(M_k)||}{M_k} \to \pi_{\vec{S}}$. Using these limits with Lemma 5 shows that $\vec{\mu}_{av}(M_k)$ converges to a vector $\vec{\mu}_{av}$ satisfying the same inequalities as those in the optimization problem of Theorem 1. Furthermore, because the system is stable we must have $\vec{\mu}_{av} \geq \vec{\lambda}$ (a necessary condition for stability, see [1]). It follows that $\underline{P}_{av}$ satisfies the feasibility constraints for the optimization problem of Theorem 1, and therefore $\underline{P}_{av} \geq P_{av}^*$.

Appendix B — The Drift Expression

Here we prove the drift expression of Lemma 2: Suppose arrivals $A_l(t)$ are i.i.d. every slot with rate $\mathbb{E}\{A_l(t)\} = \lambda_l$. For each queue l, consider the evolution equation $U_l(t+1) = \max[U_l(t) - \mu_l(\vec{P}(t), \vec{S}(t)), 0] + A_l(t)$ from (1). By squaring this equation and noting that $(\max[x, 0])^2 \leq x^2$, we obtain:

$$(U_l(t+1))^2 \leq (U_l(t))^2 + \mu_l^2 - 2U_l(t)(\mu_l - A_l) + A_l^2$$

where we have simplified the notation by writing μ_l and A_l in place of $\mu_l(\vec{P}(t), \vec{S}(t))$ and $A_l(t)$. Taking conditional expectations and summing over all l yields:

$$\Delta(\vec{U}(t)) \leq \sum_l \mathbb{E}\left\{\mu_l^2 + A_l^2 \mid \vec{U}(t)\right\} \\ -2\sum_l U_l(t)\left(\mathbb{E}\left\{\mu_l \mid \vec{U}(t)\right\} - \lambda_l\right)$$

Noting that the first term on the right hand side of the above expression is bounded by $N(\mu_{max}^{out})^2 + NA_{max}^2$ proves the result. □

REFERENCES

[1] M. J. Neely. *Dynamic Power Allocation and Routing for Satellite and Wireless Networks with Time Varying Channels*. PhD thesis, Massachusetts Institute of Technology, LIDS, 2003.

[2] A. Jalali, R. Padovani, and R. Pankaj. Data throughput of cdma-hdr a high efficiency data rate personal communication wireless system. *IEEE Vehicular Technology Conference*, May 2000.

[3] L. Tsaur and D. C. Lee. Closed-loop architecture and protocols for rapid dynamic spreading gain adaptation in cdma networks. *IEEE Proc. of INFOCOM*, 2004.

[4] J. W. Lee, R. R. Mazumdar, and N. B. Shroff. Downlink power allocation for multi-class cdma wireless networks. *IEEE Proceedings of INFOCOM*, 2002.

[5] P. Marbach and R. Berry. Downlink resource allocation and pricing for wireless networks. *IEEE Proc. of INFOCOM*, 2002.

[6] L. Xiao, M. Johansson, and S. Boyd. Simultaneous routing and resource allocation for wireless networks. *Proc. of the 39th Annual Allerton Conf. on Comm., Control, Comput.*, Oct. 2001.

[7] D. Julian, M. Chiang, D. O'Neill, and S. Boyd. Qos and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks. *Proc. INFOCOM*, 2002.

[8] B. Krishnamachari and F. Ordonez. Analysis of energy-efficient, fair routing in wireless sensor networks through non-linear optimization. *IEEE Vehicular Technology Conference*, Oct. 2003.

[9] P. Giaccone, B. Prabhakar, and D. Shah. Energy constrained wireless switching. *Proc. of Allerton Conference on Communications, Control, and Computing*, Oct. 2003.

[10] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas. A framework for cross-layer design of energy-efficient communication with qos provisioning in multi-hop wireless networks. *INFOCOM*, 2004.

[11] R. Cruz and A. Santhanam. Optimal routing, link scheduling, and power control in multi-hop wireless networks. *IEEE Proceedings of INFOCOM*, April 2003.

[12] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transacations on Automatic Control*, Vol. 37, no. 12, Dec. 1992.

[13] M. J. Neely, E. Modiano, and C. E. Rohrs. Power allocation and routing in multi-beam satellites with time varying channels. *IEEE Transactions on Networking*, Feb. 2003.

[14] M. J. Neely, E. Modiano, and C. E Rohrs. Dynamic power allocation and routing for time varying wireless networks. *IEEE Journal on Selected Areas in Communications*, January 2005.

[15] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, and P. Whiting. Providing quality of service over a shared wireless link. *IEEE Communications Magazine*, 2001.

[16] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal. Energy-efficient packet transmission over a wireless link. *IEEE/ACM Trans. Networking*, vol. 10, pp. 487-499, Aug. 2002.

[17] A. Fu, E. Modiano, and J. Tsitsiklis. Optimal energy allocation for delay-constrained data transmission over a time-varying channel. *IEEE Proceedings of INFOCOM*, 2003.

[18] L. Li and A. Goldsmith. Capacity and optimal resource allocation for fading broadcast channels: Part i: Ergodic capacity. *IEEE Trans. Inform. Theory*, ppl. 1083-1102, March 2001.

[19] X. Liu, E. K. P. Chong, and N. B. Shroff. A framework for opportunistic scheduling in wireless networks. *Computer Networks*, vol. 41, no. 4, pp. 451-474, March 2003.

[20] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Trans. on Information Theory*, vol. 39, pp. 466-478, March 1993.

[21] E. M. Yeh and A. S. Cohen. Throughput optimal power and rate control for multiaccess and broadcast communications. *Proc. of the International Symposium on Information Theory*, June 2004.

[22] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar. *Convex Analysis and Optimization*. Boston: Athena Scientific, 2003.

[23] F.P. Kelly, A.Maulloo, and D. Tan. Rate control for communication networks: Shadow prices, proportional fairness, and stability. *Journ. of the Operational Res. Society*, 49, p.237-252, 1998.

[24] R. Bhatia and M. Kodialam. On power efficient communication over multi-hop wireless networks: Joint routing, scheduling, and power control. *IEEE Proceedings of INFOCOM*, March 2004.

[25] N. McKeown, V. Anantharam, and J. Walrand. Achieving 100% throughput in an input-queued switch. *Proc. INFOCOM*, 1996.

[26] C. Schurgers and M.B. Srivastava. Energy efficient routing in wireless sensor networks. *MILCOM*, October 2001.