

Energy-Efficient Power Control in Fading Channels With Markovian Sources and QoS Constraints

Mustafa Ozmen and M. Cenk Gursoy

Abstract—In this paper, energy-efficient power adaptation policies in fading channels are analyzed when data arrivals are modeled as Markovian processes (namely, discrete Markov, Markov fluid, and discrete and fluid Markov modulated Poisson processes) and statistical quality of service (QoS) constraints are imposed on buffer overflow probabilities. In the analysis, both transmission and circuit power consumptions are considered. After formulating energy efficiency (EE) as maximum throughput normalized by the total power consumption, optimal power control policies that maximize EE are obtained for different source models. The impact of source randomness on EE is determined. Optimal power control schemes maximizing the throughput under EE or average power constraints are also investigated. With this, tradeoff between throughput and EE is studied. Finally, the analysis is extended to multichannel scenarios. Overall, the influence of source statistics, QoS constraints, and number of subchannels on the optimal power control policies, throughput, and EE performance is identified.

Index Terms—Energy efficiency, fading channels, power control, Markov arrivals, QoS constraints.

I. INTRODUCTION

ENERGY and bandwidth are two critical resources in wireless communications and it is estimated that information and communication technologies (ICT) account for 4-8% of the world-wide energy consumption [1]. Due to exponential growth in wireless systems and mobile traffic and applications, it is expected that throughput and energy demand will increase even further. In order to address these issues together with increasing energy costs and environmental concerns, energy efficiency in wireless communications has been attracting much interest [2], [3].

Recent years have also experienced an exponential growth in multimedia traffic in wireless networks. Indeed, mobile multimedia traffic now represents more than half of the entire mobile traffic [5]. For such multimedia traffic, providing certain quality-of-service (QoS) guarantees to the end-user is a critical consideration. For instance, in voice over IP (VoIP)

systems, multimedia streaming and online gaming applications, constraints on delay, packet loss, or buffer overflow probabilities are imposed so that acceptable performance and quality levels can be met for the end-users. Satisfying such QoS requirements is a challenging task especially when the source traffic and channel conditions randomly vary over time. However, if proper stochastic models for source traffic and channel are incorporated into the analysis, networks can be designed to efficiently use the limited resources while satisfying the QoS requirements of the stochastic information flows. For instance, as an accurate model for voice traffic, ON-OFF process with fixed-rate data arrivals in the ON state can be exploited. Moreover, video traffic can be modeled as autoregressive, Markovian, or Markov-modulated processes as it exhibits correlations [4].

With these motivations, in this paper, we explicitly model the information flows stochastically using Markovian models and study energy-efficient wireless transmission strategies in the presence of statistical QoS constraints imposed on buffer overflow probabilities. In particular, we identify energy-efficient power control policies in fading channels for different source arrival models. We further investigate the tradeoff between throughput and energy efficiency (EE).

A. Literature Overview

To provide statistical performance guarantees in the presence of random arrivals, Chang in [6] developed the stochastic network calculus by formulating the theory of effective bandwidth of a time-varying source. Given the source characteristics, effective bandwidth identifies the minimum constant transmission rate needed to satisfy statistical QoS requirements [7] in the form of limitations on the buffer/delay violation probabilities. Under constraints on the buffer overflow probability, Elwalid and Mitra studied the effective bandwidth of Markovian traffic sources in [11]. Furthermore, effective bandwidth formulations were provided for discrete-time Markov, Markov fluids as well as memoryless (Poisson) sources in [12].

As a dual concept to effective bandwidth, Wu and Negi defined in [10] the effective capacity, which describes the maximum constant arrival rate that a given time-varying service process can support while satisfying the statistical QoS requirements. Following this work, effective capacity of wireless communication systems has been analyzed in various contexts, recently including energy efficiency and power control (see e.g., [14]–[21]). For instance, the fundamental limits of energy efficiency in the low signal-to-noise ratio (SNR)

Manuscript received December 3, 2015; revised June 4, 2016 and September 10, 2016; accepted September 16, 2016. Date of publication October 4, 2016; date of current version December 15, 2016. This work was supported by the National Science Foundation under Grants CNS-1443966, ECCS-1443994, and CCF-1618615. The paper was presented at the IEEE Global Conference on Signal and Information Processing, Austin, TX, USA, Dec. 2013. The associate editor coordinating the review of this paper and approving it for publication was N. B. Mehta.

The authors are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: mozmen@syr.edu; mcgursoy@syr.edu).

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

Digital Object Identifier 10.1109/TCOMM.2016.2614978

regime in fading channels were determined under QoS constraints in [14]. Musavian and Le-Ngoc in [15] incorporated the circuit power consumption into their analysis of energy efficiency. Ru *et al.* in [16] investigated the minimum energy per bit and wideband slope in a hybrid cellular system. Liu in [17] considered the optimal power control to achieve the maximum energy efficiency. Helmy and Musavian in [18] considered a multichannel scenario in which they obtained the optimal power allocation for each channel to achieve the maximum global energy efficiency. Authors in [19] studied the energy-efficient design in downlink OFDMA systems. In a recent study in [22], the authors analyzed energy-efficient resource allocation strategies in MIMO-OFDM systems in the presence of random arrivals and statistical QoS requirements. In particular, they characterized the optimal energy-efficient queue-length based resource allocation policy that minimizes the total power consumption while satisfying the QoS requirements. Furthermore, in [24]–[26], power control policies were examined under QoS constraints.

B. Contributions

As exemplified above, the studies on energy efficiency and power control conducted with effective capacity formulations have mainly centered around the assumption that sources have constant arrival rates. In this paper, we take into account the stochastic nature of information flows and investigate the effect of the randomness and burstiness of the source traffic on the energy-efficient design of wireless systems. Specifically, we consider Markovian source models (namely discrete-time Markov, Markov fluid, and both discrete and fluid Markov modulated Poisson processes (MMPP)) and determine the optimal energy-efficient power allocation policies.¹ The contributions of this paper can be further detailed as follows:

- Considering two-state (ON/OFF) source models, throughput expressions are provided and subsequently energy efficiency metric is identified for discrete-time Markov, Markov fluid, and MMPP arrival models. Overall, an analytical framework is provided to study the energy efficiency of wireless transmissions in the presence of random data arrivals and statistical queueing constraints.
- After taking both the circuit and transmission power into account, optimal power allocation policies that maximize the energy efficiency are determined for different source models.
- Power control policies that maximize the throughput under either energy efficiency or power constraints are also obtained.
- In addition to single-channel systems, power allocation and control strategies that maximize the throughput in multichannel systems under energy efficiency constraints are determined.
- Via both analytical and numerical results, the impact of source randomness, channel fading, queueing constraints, and power control strategies on the energy efficiency

¹Recently, we have studied the throughput and energy efficiency with random arrivals in [20] in which no power control has been considered (i.e., transmissions occur with fixed power level) and circuit power consumption has not been addressed.

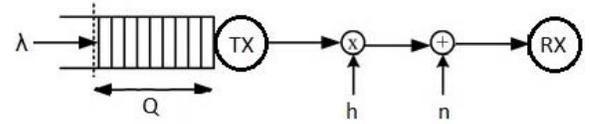


Fig. 1. System Model.

is identified. Tradeoff between energy efficiency and throughput is explored.

The remainder of the paper is organized as follows: In Section II, we describe the channel model. Section III contains the preliminaries regarding the statistical queueing constraints, effective bandwidth, effective capacity and throughput formulations. In Section IV, we study power control with the goal of maximizing the energy efficiency with Markovian source models. We analyze power control schemes that achieve the maximum throughput under energy efficiency and average power constraints in Sections V and VI, respectively. We investigate power control in multichannel communications in Section VII for Markovian arrivals. Finally, concluding remarks are given in Section VIII.

II. CHANNEL MODEL

We consider a flat-fading channel between the transmitter and receiver. The channel input-output relation can be expressed as

$$y_i = h_i x_i + n_i \text{ for } i = 1, 2, \dots \quad (1)$$

where x_i and y_i are the channel input and output, respectively, and h_i denotes the channel fading coefficient. We assume that the transmitter, equipped with perfect channel side information (CSI), performs power control. Hence, the transmit power $P(\theta, z_i)$, where $z_i = |h_i|^2$ and θ is a QoS parameter described in the following section, varies with QoS requirements and fading. Fading coefficients are assumed to be identically distributed, and the fading distribution can be arbitrary with finite variance. We consider a block-fading model and assume that the realizations of the fading coefficients stay fixed for a block of symbols and change independently for the next block. Finally, $\{n_i\}$ is a sequence of independent, zero-mean, circularly-symmetric, complex Gaussian noise components with variance $\mathbb{E}\{|n_i|^2\} = N_0$.

III. PRELIMINARIES

A. Effective Bandwidth of Markov Arrivals

As depicted in Fig. 1, we assume that the data to be sent is generated from Markovian sources and is initially stored in a buffer before transmission over the flat-fading channel. Statistical constraints are imposed on the buffer length. In particular, we assume that the buffer overflow probability satisfies

$$\lim_{q \rightarrow \infty} \frac{\log \Pr\{Q \geq q\}}{q} = -\theta \quad (2)$$

where Q denotes the stationary queue length, and θ is the decay rate of the tail distribution of the queue length. The above limiting formula implies that for large buffer threshold q , we have

$$\Pr\{Q \geq q\} \approx e^{-\theta q}. \quad (3)$$

Indeed, a closer approximation is [10]

$$\Pr\{Q \geq q\} \approx \zeta e^{-\theta q} \quad (4)$$

where $\zeta = \Pr\{Q > 0\}$ is the probability of non-empty buffer. Hence, for a sufficiently large threshold, the buffer overflow probability should decay exponentially with rate controlled by the QoS exponent θ . Note that as θ increases, stricter queueing or QoS constraints are imposed.

Conversely, for a given buffer threshold q and overflow probability limit $\varrho = \Pr\{Q \geq q\}$, the desired value of θ can be determined as

$$\theta = \frac{1}{q} \log_e \frac{\zeta}{\varrho}. \quad (5)$$

In the given setting, the delay violation probability is also characterized to decay exponentially and is approximated by [13]

$$\Pr\{D \geq d\} \approx \zeta e^{-\theta a^*(\theta)d} \quad (6)$$

where D is the queueing delay in the buffer at steady state, d is the delay threshold, and $a^*(\theta)$ is the effective bandwidth of the arrival process, described below.

We consider four types of Markovian sources, namely discrete Markov source, Markov fluid source, discrete-time Markov-modulated Poisson process (MMPP), continuous-time MMPP, and concentrate on two-state (ON-OFF) model. For these types of sources, we briefly describe below the effective bandwidth, which characterizes the minimum constant transmission (or service) rate required to support the given time-varying data arrivals while the buffer overflow probability satisfies (2).

1) *Discrete Markov Source*: Data arrivals from this source are modeled as a discrete-time Markov process with a transition probability matrix \mathbf{J} . We consider a two-state Markov chain in which r bits arrive (i.e., the arrival rate is r bits/block) in the ON state while there are no arrivals in the OFF state. With the state transition probability matrix

$$\mathbf{J} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad (7)$$

the effective bandwidth is given by [9]

$$a(\theta) = \frac{1}{\theta} \log_e \left(\frac{p_{11} + p_{22} e^{r\theta}}{2} + \frac{\sqrt{(p_{11} + p_{22} e^{r\theta})^2 - 4(p_{11} + p_{22} - 1)e^{r\theta}}}{2} \right) \quad (8)$$

where p_{11} denotes the probability of staying in the OFF state and p_{22} denotes the probability of staying in the ON state. The probabilities of transitioning from one state to a different one are therefore denoted by $p_{21} = 1 - p_{22}$ and $p_{12} = 1 - p_{11}$.

2) *Markov Fluid Source*: In this case, the data arrivals are modeled as a continuous-time Markov process with a generating matrix \mathbf{G} . Hence, as opposed to the discrete-time Markov model in which state transitions can only occur in discrete time steps, state holding times are now exponentially distributed and state transitions can occur anytime. The generating matrix for the two-state case is in the form of

$$\mathbf{G} = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix} \quad (9)$$

where α and β are the transition rates from one state to another. When the arrival rates for the two-state model are r and 0 and hence we basically have ON and OFF states, the effective bandwidth expression is [12]

$$a(\theta) = \frac{1}{2\theta} \left[\theta r - (\alpha + \beta) + \sqrt{(\theta r - (\alpha + \beta))^2 + 4\alpha\theta r} \right]. \quad (10)$$

3) *Discrete-Time Markov Modulated Poisson Sources*: In this source model, the data arrival to the buffer is a Poisson process whose intensity is controlled by a discrete-time Markov chain. We again consider a two-state model in which the intensity of the Poisson arrival process is r and 0 in the ON and OFF states of the Markov chain, respectively. Therefore, the source arrival is modeled as a Markov-modulated Poisson process (MMPP). Assuming that the matrix \mathbf{J} in (7) is the transition probability matrix of the Markov chain, the effective bandwidth is given by [6]

$$a(\theta) = \frac{1}{\theta} \log_e \left(\frac{p_{11} + p_{22} e^{r(e^\theta - 1)}}{2} + \frac{\sqrt{(p_{11} + p_{22} e^{r(e^\theta - 1)})^2 - 4(p_{11} + p_{22} - 1)e^{r(e^\theta - 1)}}}{2} \right). \quad (11)$$

4) *Continuous-Time Markov Modulated Poisson Sources*: In this case, we consider that the data arrival to the buffer is again a Poisson process but now its intensity is controlled by a continuous-time Markov chain. Assuming that the intensity of the Poisson arrival process is r and 0 in the ON and OFF states of the Markov chain, respectively, and \mathbf{G} in (9) is the irreducible transition rate matrix of the Markov chain, the effective bandwidth is given by [11], [12]

$$a^*(\theta) = \frac{1}{2\theta} \left[(e^\theta - 1)r - (\alpha + \beta) \right] + \frac{1}{2\theta} \sqrt{\left[(e^\theta - 1)r - (\alpha + \beta) \right]^2 + 4\alpha(e^\theta - 1)r}. \quad (12)$$

B. Effective Capacity of Fading Channels

Effective capacity, as a dual concept to effective bandwidth, identifies the maximum constant arrival rate that can be supported by a given time-varying service process while satisfying (2). Under the block-fading assumption, the effective capacity can be expressed as [10], [14]

$$C_E(\theta) = -\frac{1}{\theta} \log_e \mathbb{E} \left\{ e^{-\theta T B \log_2(1 + P(\theta, z) \frac{z}{N_0 B})} \right\} \quad (13)$$

$$= -\frac{1}{\theta} \log_e \mathbb{E} \left\{ e^{-\theta T B \log_2(1 + \mu(\theta, z) z)} \right\} \quad (14)$$

$$= -\frac{1}{\theta} \log_e \mathbb{E} \left\{ (1 + \mu(\theta, z) z)^{-\eta} \right\} \quad (15)$$

where T is the frame/block duration over which the fading stays fixed, B is the total bandwidth, $P(\theta, z)$ is the instantaneous transmission power and we define $\mu(\theta, z) = \frac{P(\theta, z)}{N_0 B}$ (as the instantaneous SNR) and $\eta = \theta T B \log_2 e$.

In a multichannel scenario (e.g., in multicarrier models), assuming that there are K subchannels each with bandwidth $\frac{B}{K}$, the instantaneous service rate becomes

$$R(\mathbf{z}) = \sum_{k=1}^K \frac{B}{K} \log_2(1 + \mu_k(\mathbf{z})z_k). \quad (16)$$

Above, we define $\mu_k(\mathbf{z}) = \frac{P_k(\theta, \mathbf{z})}{N_0 \frac{B}{K}}$, where $P_k(\theta, \mathbf{z})$ is the instantaneous transmission power in the k^{th} subchannel, and $\mathbf{z} = [z_1, \dots, z_K]$, where $z_k = |h_k|^2$ is the magnitude-square of the fading coefficient in the k^{th} subchannel. Under the block-fading assumption, the effective capacity with K subchannels can be expressed as

$$\begin{aligned} C_E(\theta) &= -\frac{1}{\theta} \log_e \mathbb{E} \left\{ e^{-\theta T R(\mathbf{z})} \right\} \\ &= -\frac{1}{\theta} \log_e \mathbb{E} \left\{ \prod_{k=1}^K e^{-\theta T B \frac{1}{K} \log_2(1 + \mu_k(\mathbf{z})z_k)} \right\} \\ &= -\frac{1}{\theta} \log_e \mathbb{E} \left\{ \prod_{k=1}^K (1 + \mu_k(\mathbf{z})z_k)^{-\frac{\theta}{K}} \right\}. \end{aligned} \quad (17)$$

C. Throughput With Markovian Source Models

In this section, we formulate the throughput of wireless fading channels when the data arrivals are modeled by Markovian processes. As described in the previous subsection, we primarily consider two-state Markovian arrival models with ON and OFF states. For the discrete-time Markov processes or Markov fluids, the arrival rates are constants r and 0 in the ON and OFF states, respectively. On the other hand, for MMPP sources, arrival rates are random in the ON state with a Poisson distribution with intensity r (i.e., the expected number of arrivals per block or equivalently average arrival rate is r) while there are no arrivals in the OFF state. Note that if there is no OFF state, this model specializes to the pure Poisson arrival process.

We denote by P_{ON} the probability that source is in the ON state. Hence, the average arrival rate of the two-state Markovian source models simply becomes

$$r_{\text{avg}} = P_{\text{ON}} r \quad (18)$$

which is equal to the average departure rate when the queue is in steady state [8].

Now, we seek to determine the throughput by identifying the maximum average arrival rate that can be supported by the fading channel described in Section II while satisfying the statistical QoS limitations given in the form in (2). As shown in [8, Th. 2.1], (2) is satisfied, i.e., buffer overflow probability decays exponentially fast with rate controlled by the QoS exponent θ , if the effective bandwidth of the arrival process is equal to the effective capacity of the service process, i.e.,

$$a(\theta) = C_E(\theta). \quad (19)$$

Hence by solving (19),² we can determine the maximum average arrival rate $r_{\text{avg}}^*(\theta)$. In the following, we provide closed-form expressions (in terms of effective capacity and source statistics) of the maximum average arrival rates for the considered ON-OFF Markovian source models.

1) *Discrete Markov Source*: In order to express the maximum average arrival rate in terms of $C_E(\theta)$, we use the effective bandwidth expression in (8) and simplify (19) as follows:

$$\begin{aligned} & \left(p_{11} + p_{22} e^{r\theta} - 2e^{\theta C_E} \right)^2 \\ &= (p_{11} + p_{22} e^{r\theta})^2 - 4(p_{11} + p_{22} - 1)e^{r\theta}. \end{aligned} \quad (20)$$

After solving (20) for r and using (18), we obtain the maximum average arrival rate as

$$r_{\text{avg}}^*(\theta) = \frac{P_{\text{ON}}}{\theta} \left[\log_e \left(\frac{e^{2\theta C_E(\theta)} - p_{11} e^{\theta C_E(\theta)}}{(1 - p_{11} - p_{22}) + p_{22} e^{\theta C_E(\theta)}} \right) \right]. \quad (21)$$

Note that the probability of ON state is given as

$$P_{\text{ON}} = \frac{1 - p_{11}}{2 - p_{11} - p_{22}}. \quad (22)$$

For the special case in which $p_{11} = 1 - q$ and $p_{22} = q$, the expression for the average arrival rate can be simplified further as

$$r_{\text{avg}}^*(\theta) = \frac{q}{\theta} \log_e \left(\frac{e^{\theta C_E(\theta)} - (1 - q)}{q} \right). \quad (23)$$

2) *Markov Fluid Source*: For Markov fluid sources, in order to obtain the maximum average arrival rate in terms of $C_E(\theta)$, we incorporate (10) into (19) and perform straightforward simplifications on (19) to obtain

$$(\theta r - (\alpha + \beta) - 2\theta C_E)^2 = (\theta r - (\alpha + \beta))^2 + 4\alpha\theta r. \quad (24)$$

Solving the above equation, we can express the maximum average arrival rate as

$$r_{\text{avg}}^*(\text{SNR}, \theta) = P_{\text{ON}} \frac{\theta C_E(\theta) + \alpha + \beta}{\theta C_E(\theta) + \alpha} C_E(\theta). \quad (25)$$

Note that the probability of ON state is given as

$$P_{\text{ON}} = \frac{\alpha}{\alpha + \beta}. \quad (26)$$

3) *Discrete-Time Markov Modulated Poisson Process*: In order to determine the maximum average arrival rate in terms of $C_E(\theta)$, we insert the effective bandwidth expression in (11) into (19) and obtain

$$\begin{aligned} & \left(p_{11} + p_{22} e^{r(e^\theta - 1)} - 2e^{\theta C_E} \right)^2 \\ &= (p_{11} + p_{22} e^{r(e^\theta - 1)})^2 - 4(p_{11} + p_{22} - 1)e^{r(e^\theta - 1)}. \end{aligned} \quad (27)$$

²Solving (19) is ensured due to the following facts: Effective bandwidth is a monotonically increasing function of θ , ranging between the average rate of the arrival process at $\theta = 0$ and the peak rate as $\theta \rightarrow \infty$ [6], [11]. Effective capacity is, on the other hand, is a decreasing function of θ , ranging between the average transmission rate (i.e., the ergodic capacity) at $\theta = 0$ and the minimum constant transmission rate (which is often zero for typical fading distributions) as $\theta \rightarrow \infty$ [28].

After solving the above equation for r , we obtain the maximum average arrival rate as

$$r_{\text{avg}}^*(\theta) = \frac{P_{\text{ON}}}{(e^\theta - 1)} \left[\log_e \left(\frac{e^{2\theta C_E(\theta)} - p_{11} e^{\theta C_E(\theta)}}{(1 - p_{11} - p_{22}) + p_{22} e^{\theta C_E(\theta)}} \right) \right]. \quad (28)$$

4) *Continuous-Time Markov Modulated Poisson Process*: Similarly, we find the maximum average arrival rate $r_{\text{avg}}^*(\theta)$ by incorporating (12) into (19) and expressing (19) as

$$\begin{aligned} & ((e^\theta - 1)r - (\alpha + \beta) - 2\theta C_E)^2 \\ &= ((e^\theta - 1)r - (\alpha + \beta))^2 + 4\alpha(e^\theta - 1)r. \end{aligned} \quad (29)$$

We can simplify the above equality and solve for the maximum Poisson arrival intensity in the ON state to obtain the maximum average arrival rate as

$$r_{\text{avg}}^*(\theta) = P_{\text{ON}} \frac{\theta [\theta C_E(\theta) + \alpha + \beta]}{(e^\theta - 1) [\theta C_E(\theta) + \alpha]} C_E(\theta). \quad (30)$$

Having formulated the maximum average arrival rates in terms of the effective capacity and source statistics, we next identify the optimal power control policies, maximizing the energy efficiency. In order to have convex optimization problems below, we need to show that throughput $r_{\text{avg}}^*(\theta)$ is concave in $\text{SNR} = \mathbb{E}\{\mu(\theta, z)\}$. In [14, Lemma 1], it is proven that effective capacity is a concave function of SNR. In [11], it is shown that effective bandwidth of the source is strictly monotonically increasing and is also convex in source arrival rates. Therefore, inverse function of the effective bandwidth $a^{*-1}(C_E(\theta))$ exists and is a nondecreasing concave function of the effective capacity, which is concave in SNR. Using the composition properties of concave functions [27], we immediately conclude that the maximum average arrival rate

$$r_{\text{avg}}^*(\theta) = P_{\text{ON}} a^{*-1}(C_E(\theta)) \quad (31)$$

is also a concave function of SNR.

IV. ENERGY-EFFICIENT POWER CONTROL

In this paper, we employ rate per unit energy (in bits/joule) as the performance metric of energy efficiency. In our setup, we define energy efficiency (EE) as

$$\text{EE} = \frac{r_{\text{avg}}^*(\theta)}{(\frac{1}{\epsilon} \mathbb{E}\{P(\theta, z)\} + P_c) / N_0 B} = \frac{r_{\text{avg}}^*(\theta)}{(\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c)} \quad (32)$$

where P_c is the circuit power and ϵ is the efficiency of the power amplifier, and $\mu_c = P_c / N_0 B$. Normalization with the noise power $N_0 B$ in the denominator above is performed in order to express EE in terms of the instantaneous SNR $\mu(\theta, z)$, and to perform optimization over $\mu(\theta, z)$ and have simplifications in the expressions. Furthermore, to be used in subsequent formulations, we define function $g(\theta)$ as

$$g(\theta) = \mathbb{E}\{[1 + \mu(\theta, z)z]^{-\eta}\} \quad (33)$$

where again $\eta = \theta T B \log_2 e$.

After formulating the energy efficiency, we can express the optimally energy-efficient power control problem as³

$$\max_{\mu(\theta, z)} \frac{r_{\text{avg}}^*(\theta)}{(\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c)}. \quad (34)$$

Next, we will address special cases of this optimization problem by considering specific arrival models and incorporating the corresponding average arrival rate expressions.

A. Discrete Markov Source

In this section, we consider ON-OFF discrete Markov arrival models and determine the optimal power adaptation strategy that maximizes the energy efficiency. After inserting the maximum average arrival rate expression in (21) into the optimization problem in (34) and simplifying the expressions by eliminating the constant terms, we can formulate the optimal power allocation problem as

$$\mu^*(\theta, z) = \arg \max_{\mu(\theta, z)} \frac{\log_e \left(\frac{1 - p_{11} g(\theta)}{(1 - p_{11} - p_{22}) g^2(\theta) + p_{22} g(\theta)} \right)}{\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c} \quad (35)$$

where the function $g(\cdot)$ is defined in (33). Note that any function that can be expressed as the ratio of a convex function over a concave one is quasiconvex [27, Example 3.38] and the negative of a quasiconvex function is quasiconcave. Hence, the objective function in (35), being a concave function divided by an affine function of power allocation, is a quasiconcave function of the instantaneous SNR $\mu(\theta, z)$. By introducing an additional variable $\psi = \frac{1}{\mathbb{E}\{\frac{1}{\epsilon} \mu(\theta, z)\} + \mu_c}$, the problem can be transformed into

$$\min_{\mu(\theta, z) \geq 0} -\psi \log_e \left(\frac{1 - p_{11} g(\theta)}{(1 - p_{11} - p_{22}) g^2(\theta) + p_{22} g(\theta)} \right) \quad (36)$$

$$\text{subject to } \psi \left(\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c \right) = 1. \quad (37)$$

The problem in (36) is a convex optimization problem. Therefore, we can use the convex optimization tools and determine the sufficient and necessary Karush-Kuhn-Tucker (KKT) conditions. By denoting the Lagrange multiplier by λ , we form the Lagrangian as

$$\begin{aligned} \mathcal{L}(\mu(\theta, z), \psi, \lambda) &= -\psi \log_e \left(\frac{1 - p_{11} g(\theta)}{(1 - p_{11} - p_{22}) g^2(\theta) + p_{22} g(\theta)} \right) \\ &+ \lambda \left[\psi \left(\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c \right) - 1 \right]. \end{aligned} \quad (38)$$

Now, the KKT conditions are given in (39)–(41), as shown at the top of the next page. Note that (39) is due to the constraint in (37). (40) and (41) are obtained by taking the

³Since the theory of effective bandwidth and effective capacity makes use of tools from large deviations and characterizes the performance in the large-queue-length regime, we consider a saturated buffer in our analysis, and the optimal power control policies are obtained under the assumption that there is always data to transmit from the buffer.

$$\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1, \quad (39)$$

$$-\psi \eta z [1 + \mu(\theta, z)z]^{-\eta-1} \left(\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)} \right) + \frac{\lambda \psi}{\epsilon} = 0, \quad (40)$$

$$-\log_e \left(\frac{1-p_{11}g(\theta)}{(1-p_{11}-p_{22})g^2(\theta)+p_{22}g(\theta)} \right) + \lambda \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0. \quad (41)$$

derivative of the Lagrangian in (38) with respect to $\mu(\theta, z)$ and ψ , respectively. After simplifying (40), we obtain

$$z [1 + \mu(\theta, z)z]^{-\eta-1} = \frac{\lambda/(\eta\epsilon)}{\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)}}. \quad (42)$$

By solving (42) for $\mu(\theta, z)$, the optimal power allocation is found as

$$\mu^*(\theta, z) = \left[\frac{1}{v \frac{1}{1+\eta} z^{\frac{\eta}{1+\eta}} - \frac{1}{z}} \right]^+ \quad (43)$$

where

$$v = \frac{\lambda/(\eta\epsilon)}{\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)}} \quad (44)$$

and $[c]^+ = \max(c, 0)$. We notice that $\mu^*(\theta, z) = 0$ when $z \leq v$. Hence, v can be regarded as the fading gain threshold for transmission. When we consider the special case of constant-rate arrivals (i.e., when we have $P_{\text{ON}} = 1$), the above equation for v specializes to the corresponding one in [15]. Note that the expression for v in (44) depends on the Lagrange multiplier λ (and hence v can also be considered as a scaled Lagrange multiplier). By combining (41) with (44), we obtain (45), as shown at the bottom of this page. Equation (45), which does not depend on λ , can be used to determine v by incorporating the source statistics⁴ and computing $\mathbb{E} \{ \mu(\theta, z) \}$ and $g(\theta)$. For instance, in the case of Rayleigh fading, the fading power is exponentially distributed with density function $f_z(z) = e^{-z}$, and by using the expression for $\mu(\theta, z)$ in (43), these key expectations can be determined in closed-form as follows:

$$\begin{aligned} \mathbb{E} \{ \mu(\theta, z) \} &= \int_v^\infty \left[\frac{1}{v \frac{1}{1+\eta} z^{\frac{\eta}{1+\eta}} - \frac{1}{z}} \right] e^{-z} dz \\ &= \left(\frac{1}{v} \right)^{\frac{1}{1+\eta}} \int_v^\infty z^{-\frac{\eta}{1+\eta}} e^{-z} dz - \int_v^\infty \frac{e^{-z}}{z} dz \\ &= \left(\frac{1}{v} \right)^{\frac{1}{1+\eta}} \Gamma \left(\frac{1}{1+\eta}, v \right) + \text{Ei}(-v), \end{aligned} \quad (46)$$

⁴It is interesting to note that the optimal power control $\mu^*(\theta, z)$ depends on the source statistics (e.g., transition probabilities p_{11} and p_{22}) only through v .

$$\begin{aligned} g(\theta) &= \mathbb{E} \{ [1 + \mu(\theta, z)z]^{-\eta} \} \\ &= \int_v^\infty \left(\frac{z}{v} \right)^{-\frac{\eta}{1+\eta}} e^{-z} dz + \int_0^v e^{-z} dz \\ &= v^{\frac{\eta}{1+\eta}} \Gamma \left(\frac{1}{1+\eta}, v \right) + 1 - e^{-v}. \end{aligned} \quad (47)$$

Above, $\Gamma(s, w) = \int_w^\infty \tau^{s-1} e^{-\tau} d\tau$ is the upper incomplete gamma function and $\text{Ei}(w) = -\int_{-w}^\infty \frac{e^{-\tau}}{\tau} d\tau$ is the exponential integral.

When $p_{11} = 1-s$, $p_{22} = s$, we have a memoryless discrete source and power allocation problem becomes

$$\begin{aligned} \min_{\mu(\theta, z) \geq 0} \quad & -\psi \log_e \left(\frac{\frac{1}{g(\theta)} - (1-s)}{s} \right) \\ \text{subject to} \quad & \psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1. \end{aligned} \quad (48)$$

Thus, Lagrangian function transforms into

$$\begin{aligned} \mathcal{L}(\mu(\theta, z), \psi, \lambda) &= -\psi \log_e \left(\frac{\frac{1}{g(\theta)} - (1-s)}{s} \right) \\ &+ \lambda \left[\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) - 1 \right]. \end{aligned} \quad (50)$$

Now, the KKT conditions become

$$\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1, \quad (51)$$

$$\psi \frac{-\eta z [1 + \mu(\theta, z)z]^{-\eta-1}}{g(\theta) - (1-s)(g(\theta))^2} + \frac{\lambda \psi}{\epsilon} = 0, \quad (52)$$

$$-\log_e \left(\frac{\frac{1}{g(\theta)} - (1-s)}{s} \right) + \lambda \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0. \quad (53)$$

Power allocation policy formula is still as in (43) but now v is determined from

$$\begin{aligned} & - \left[g(\theta) - (1-s)(g(\theta))^2 \right] \log_e \left(\frac{\frac{1}{g(\theta)} - (1-s)}{s} \right) \\ & + v^* \eta \epsilon \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0. \end{aligned} \quad (54)$$

Next, we provide numerical results for the general case of discrete Markov source with memory. For the numerical

$$-\frac{1}{\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))((1-p_{11}-p_{22})g(\theta)+p_{22})} + \frac{1}{g(\theta)}}} \log_e \left(\frac{1-p_{11}g(\theta)}{(1-p_{11}-p_{22})g^2(\theta)+p_{22}g(\theta)} \right) + v\eta\epsilon \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0 \quad (45)$$

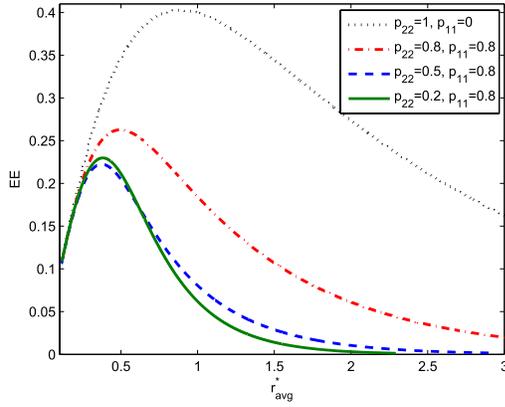


Fig. 2. Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $\eta = 1$, $P_c = 1$, $N_0 = 1$, $\epsilon = 1$.

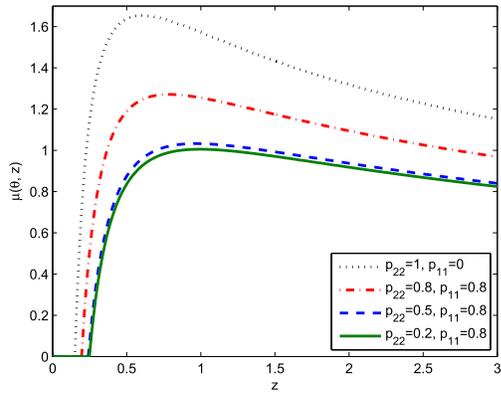


Fig. 3. Power control function $\mu(\theta, z)$ vs. z when $\eta = 1$, $P_c = 1$, $N_0 = 1$ and $\epsilon = 1$.

analysis, we set the values of the parameters as $P_c = 1$, $N_0 = 1$, $\eta = 1$, $\epsilon = 1$. In Fig. 2, we plot energy efficiency (EE) vs. maximum average arrival rate r_{avg}^* with varying source parameters. Note that when $p_{22} = 1$ and $p_{11} = 0$ (and hence $P_{\text{ON}} = 1$), we have a source with constant arrival rate. Indeed, the best performance is achieved in this case and maximal EE value (or equivalently the peak of the EE curve) is the largest. We further notice in the figure that initially the maximal EE values diminish and are achieved at a lower value of r_{avg}^* when p_{22} and consequently P_{ON} decrease and therefore the source burstiness increases. However, interestingly when p_{22} is diminished from 0.5 to 0.2, maximal EE value slightly increases even though P_{ON} is smaller when $p_{22} = 0.2$. This is due to the fact that P_{ON} is not the only criterion to indicate the burstiness of the system. In fact, as we have shown in [20], a measure of burstiness at low SNRs is $\frac{(1-p_{22})(p_{11}+p_{22})}{(1-p_{11})(2-p_{11}-p_{22})}$ whose greater values imply a more bursty source. Indeed, this expression assumes a larger value when $p_{22} = 0.5$. On the other hand, as SNR increases and higher average arrival rates are supported, P_{ON} becomes a more relevant indicator of burstiness and the source with $p_{22} = 0.2$ starts leading to lower EE values, following a crossover between the two curves.

In Fig. 3, we plot the optimal power control policy that maximizes the energy efficiency as a function of the

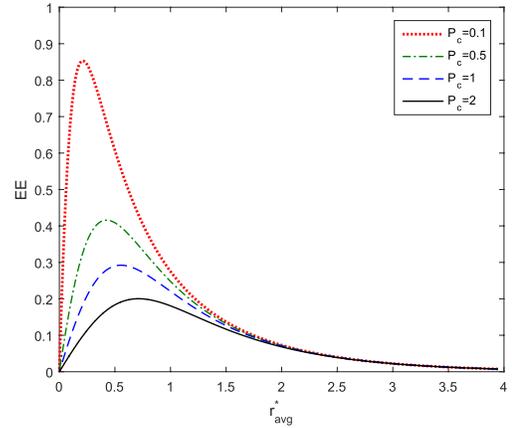


Fig. 4. Energy efficiency EE vs. maximum average arrival rate r_{avg}^* for different values of the circuit power P_c when $\eta = 1$, $p_{11} = p_{22} = 0.5$, $N_0 = 1$, $\epsilon = 1$.

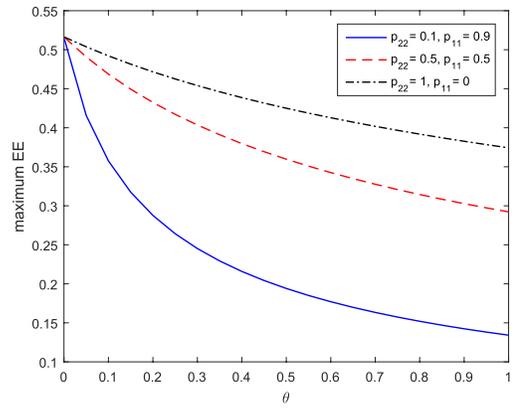


Fig. 5. Maximum EE vs. QoS exponent θ when $P_c = 1$, $N_0 = 1$, $\epsilon = 1$.

instantaneous fading power values, z . We note in all cases that no power is allocated for transmission if z is below a threshold (i.e., ν). Power level initially increases as z increases above the threshold and then starts diminishing as z further grows. Hence, power control is essentially a combination of waterfilling policy (for low values of z) and channel inversion policy (for large values of z). We also observe that more power is consumed (and consequently average power consumption is larger) for a less bursty source at the maximal EE point.

In Fig. 4, we plot the EE vs. r_{avg}^* curves for different values of the circuit power P_c . We readily notice that as P_c diminishes, a higher level of EE is achieved at a lower value of r_{avg}^* . Indeed, if circuit power is not taken into account (i.e., if we set $P_c = 0$), then maximum EE is achieved asymptotically as r_{avg}^* and hence SNR approach zero [20]. Hence, circuit power has significant impact on the performance.

In Fig. 5, we plot the maximum EE as a function of the QoS exponent θ for discrete Markov sources with different source statistics. We note in all cases that EE diminishes with increasing θ . Hence, more stringent buffer/delay constraints is detrimental to EE. Also, similarly as before, the highest levels of EE are attained when the arrival rate is constant (i.e., when $p_{22} = 1$ and $p_{11} = 0$), and the EE diminishes as the sources become more bursty.

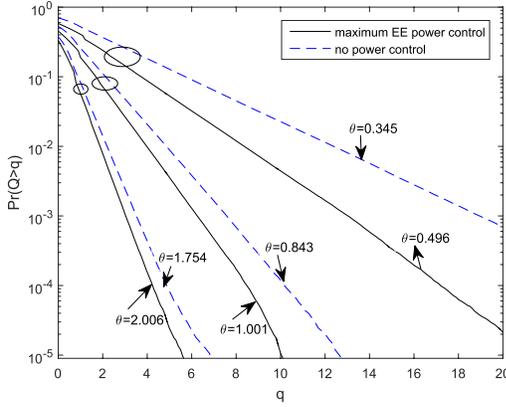


Fig. 6. Buffer overflow probability $\Pr\{Q \geq q\}$ vs. buffer threshold q for cases with optimal EE power control and fixed power when $P_c = 1$, $N_0 = 1$, $\epsilon = 1$. Discrete Markov source with $p_{11} = p_{22} = 0.75$.

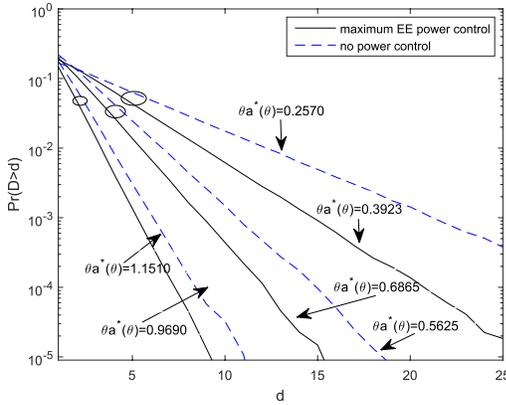


Fig. 7. Delay violation probability $\Pr\{D \geq d\}$ vs. delay threshold d for cases with optimal EE power control and fixed power when $P_c = 1$, $N_0 = 1$, $\epsilon = 1$. Discrete Markov source with $p_{11} = p_{22} = 0.75$.

Finally, in Figs. 6 and 7, we display simulation results.⁵ In particular, in Fig. 6, we have the buffer overflow probabilities $\Pr\{Q \geq q\}$ plotted as a function of the buffer threshold q with both optimal power control and no power control (i.e., with fixed transmission power). Note that we plot the buffer overflow probabilities in logarithmic scale. Note further from the approximation in (4) for large q that

$$\log \Pr\{Q \geq q\} \approx -\theta q + \log \zeta. \quad (55)$$

Hence, the logarithm of the overflow probability is expected to decay linearly in θ for large q . Indeed, we observe this linear decay already even for rather small values of q . Moreover, the simulations show excellent agreement with the theoretical analysis. Solid curves are for the case with optimal power

⁵We conduct the simulations as follows. We initially fix the value of the QoS exponent θ (e.g., $\theta = 2, 1$, or 0.5 in the figure) to provide a certain level of statistical QoS guarantee. Then, using the theoretical results from our analysis, we determine the EE-maximizing optimal power control and the value of r_{avg}^* at which EE is maximized. Subsequently, we generate random arrivals according to the discrete Markov process with average rate r_{avg}^* . We simulate the service process by generating random channel fading and using the optimal power control. Then, we have kept track of the buffer state among the arrivals and departures, and evaluated the frequency of exceeding a given threshold q to determine the values of overflow probabilities. Considering the same r_{avg}^* and the same average power and hence the same EE level, we have repeated the simulations with no power control.

control. We note that the simulated curves lead to simulated θ values of 2.006, 1.001, and 0.496 as indicated in the figure when we set $\theta = 2, 1$, and 0.5 , respectively, at the beginning of the simulations. Hence, the buffer overflow probabilities decay exponentially at the predicted rates even for very small values of q . In Fig. 6, the dashed curves next to the solid ones are the corresponding overflow probabilities at the same r_{avg}^* and EE levels but when no power control is employed. We immediately recognize that we have smaller values of θ in such cases (i.e., $\theta = 1.754, 0.843$ and 0.345 as opposed to having $\theta = 2.006, 1.001$ and 0.496 , respectively, in the power control cases) meaning that for the same threshold q , the buffer overflow probabilities are higher when transmission power is fixed. Hence, the same EE can be attained but at the cost of having more frequent buffer overflows. Conversely, we can also say that for the same overflow probability, a higher EE is achieved when power control is adopted. These observations further demonstrate the benefits of power control in practical settings.

In Fig. 7, we plot the delay violation probability $\Pr\{D \geq d\}$ in logarithmic scale as a function of the delay threshold d again from the simulations. Note from the approximation in (6) that

$$\log \Pr\{D \geq d\} \approx -\theta a^*(\theta) d + \log \zeta. \quad (56)$$

Hence, the logarithm of the delay violation probability is expected to decrease linearly in d with slope $-\theta a^*(\theta)$ where $a^*(\theta)$ is the effective bandwidth of the arrival process. We essentially have similar observations as in Fig. 6. Specifically, we again have excellent agreements with theory (e.g., the theoretical values in the power control cases are $\theta a^*(\theta) = 0.3927, 0.6944$, and 1.2022 while the corresponding simulated values are $\theta a^*(\theta) = 0.3923, 0.6865$, and 1.1510 , respectively), and having no power control increases the frequency of delay violations at a given delay threshold.

B. Markov Fluid Source

Now, we consider the optimal power control with Markov fluid sources. By using the maximum average arrival rate expression in (25) in the objective function, eliminating the constant P_{ON} , and using the definition of $g(\theta)$ in (33), we can recast the optimal power control problem in (34) as

$$\mu^*(\theta, z) = \arg \max_{\mu(\theta, z)} \frac{-\frac{\alpha + \beta - \log_e g(\theta)}{\alpha - \log_e g(\theta)} \log_e g(\theta)}{\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c}. \quad (57)$$

Again, by introducing the additional variable $\psi = \frac{1}{\mathbb{E}\{\frac{1}{\epsilon} \mu(\theta, z)\} + \mu_c}$, the problem can be transformed into

$$\min_{\mu(\theta, z) \geq 0} \psi \frac{\alpha + \beta - \log_e g(\theta)}{\alpha - \log_e g(\theta)} \log_e g(\theta) \quad (58)$$

$$\text{subject to } \psi \left(\frac{1}{\epsilon} \mathbb{E}\{\mu(\theta, z)\} + \mu_c \right) = 1. \quad (59)$$

By employing convex optimization tools, we can determine the sufficient and necessary KKT conditions. First, the Lagrangian

function is given as

$$\begin{aligned} \mathcal{L}(\mu(\theta, z), \psi, \lambda) = & \psi \frac{\alpha + \beta - \log_e g(\theta)}{\alpha - \log_e g(\theta)} \log_e g(\theta) \\ & + \lambda \left[\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) - 1 \right]. \end{aligned} \quad (60)$$

The KKT conditions are given in (61)–(63), as shown at the bottom of this page. Similarly as for the discrete Markov source, (61) is due to the constraint in (59). (62) and (63) are obtained by taking the derivative of the Lagrangian in (60) with respect to $\mu(\theta, z)$ and ψ , respectively. After simplifying (62), we obtain

$$z [1 + \mu(\theta, z)]^{-\eta-1} = \frac{\lambda / (\eta \epsilon)}{\frac{\alpha \beta}{g(\theta) (\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)}}. \quad (64)$$

Due to similarities between (42) and (64), the optimal power control function is obtained to be in the same form as for discrete Markov sources and is given by

$$\mu(\theta, z) = \left[\frac{1}{v \frac{1}{1+\eta} z^{\frac{\eta}{1+\eta}} - \frac{1}{z}} \right]^+ \quad (65)$$

but now with

$$v = \frac{\lambda / (\eta \epsilon)}{\frac{\alpha \beta}{g(\theta) (\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)}}. \quad (66)$$

Now, we can combine (63) with (66) to obtain

$$\begin{aligned} \frac{g(\theta)}{\frac{\alpha \beta}{(\alpha - \log_e g(\theta))^2} + 1} \frac{\alpha + \beta - \log_e g(\theta)}{\alpha - \log_e g(\theta)} \log_e g(\theta) \\ + v^* \eta \epsilon \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0, \end{aligned} \quad (67)$$

which can further be used to numerically evaluate v .

In Fig. 8, we plot the EE vs. maximum average arrival rate r_{avg}^* curve for different Markov fluid sources. As expected, the source with $\beta = 0$, being the constant arrival source, has the best performance in terms of energy efficiency. As α reduces, the source becomes more bursty and the performance degrades.

In Fig. 9, we again plot the EE vs. r_{avg}^* curves for a Markov fluid source with transition rates $\alpha = 2$ and $\beta = 8$, considering the optimal power control, suboptimal water-filling power control, and constant-power transmissions. As expected, optimal power control leads to the maximum EE and outperforms the other two schemes uniformly over the entire range. Water-filling power control results in the second-highest EE level. At the same time, it is interesting to observe that transmission

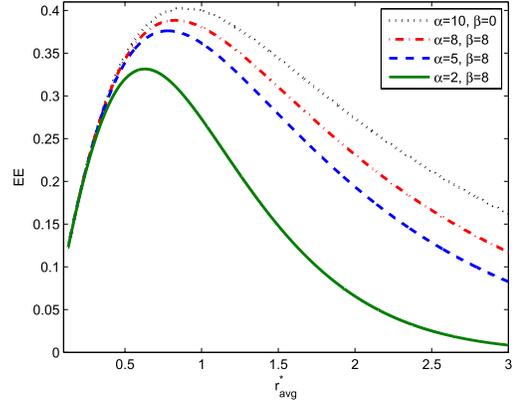


Fig. 8. Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $\eta = 1$, $P_c = 1$, $N_0 = 1$, $\epsilon = 1$.

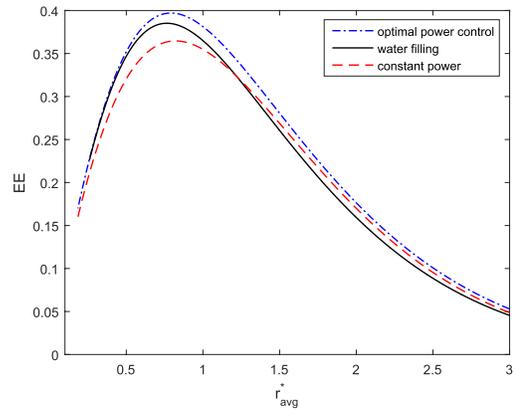


Fig. 9. Energy efficiency EE vs. maximum average arrival rate r_{avg}^* with different power control schemes when $\eta = 0.5$, $\alpha = 2$, $\beta = 8$, $P_c = N_0 = 1 = \epsilon = 1$.

with constant power starts performing better than that with the water-filling policy as r_{avg}^* increases. This observation highlights the importance of identifying the optimal power control since water-filling takes into account neither source randomness nor QoS constraints.

C. Markov-Modulated Poisson Processes

The throughput expressions for discrete-time and continuous-time Markov-modulated Poisson sources have similarities to those for discrete-time Markov and Markov fluid sources, respectively. Particularly, (28) is obtained by scaling (21) with $\frac{\theta}{e^{\theta}-1}$. The same observation holds regarding the comparison between (25) and (30). These scaling differences do not alter the optimal power control problem. Therefore, the optimal power control policies for

$$\psi \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 1, \quad (61)$$

$$-\psi \eta z [1 + \mu(\theta, z)]^{-\eta-1} \left(\frac{\alpha \beta}{g(\theta) (\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)} \right) + \frac{\lambda \psi}{\epsilon} = 0, \quad (62)$$

$$\frac{\alpha + \beta - \log_e g(\theta)}{\alpha - \log_e g(\theta)} \log_e g(\theta) + \lambda \left(\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right) = 0. \quad (63)$$

the discrete-time and continuous-time MMPP sources are the same as for the cases of discrete and fluid Markov sources, respectively.

V. OPTIMAL POWER CONTROL WITH EE CONSTRAINTS

As noticed in the previous section, when the primary goal is the maximization of the energy efficiency, small throughput values can be attained especially if the source is bursty. On the other hand, in certain wireless systems, the goal is to maximize the throughput while being cognizant of the energy efficiency requirements. Motivated by such systems, we in this section assume that there is a minimum energy efficiency constraint on the system and we seek to find the optimal power allocation scheme to maximize the throughput. The optimal power allocation problem is formulated as

$$\max_{\mu(\theta, z) \geq 0} r_{\text{avg}}^*(\theta) \quad (68)$$

$$\text{subject to } \frac{r_{\text{avg}}^*(\theta)}{\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c} \geq \zeta_{\min}(\theta) \quad (69)$$

where ζ_{\min} represents the minimum required EE level. This optimization problem also enables us to characterize the tradeoff between the throughput and energy efficiency.

Note that the constraint can also be expressed as

$$-r_{\text{avg}}^*(\theta) + \zeta_{\min}(\theta) \left[\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right] \leq 0. \quad (70)$$

Again, we first demonstrate that the power allocation problem is convex and hence we can use convex optimization tools to solve the problem. As discussed at the end of Section III-C, the objective function $r_{\text{avg}}^*(\theta)$ in (68) is a concave function of $\mu(\theta, z)$. It can be easily seen that the constraint in (70) is a convex function of $\mu(\theta, z)$ as it is the summation of a negative concave function and an affine function. Hence, the Lagrangian can be expressed as

$$\begin{aligned} \mathcal{L}(\mu(\theta, z), \lambda) \\ = r_{\text{avg}}^*(\theta, z) - \lambda \left\{ -r_{\text{avg}}^*(\theta, z) + \zeta_{\min}(\theta) \left[\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right] \right\}. \end{aligned} \quad (71)$$

A. Discrete Markov Source

The Lagrangian for discrete Markov source is simplified to

$$\begin{aligned} \mathcal{L}(\mu(\theta, z), \lambda) \\ = (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \log_e \left(\frac{1 - p_{11}g(\theta)}{(1 - p_{11} - p_{22})g^2(\theta) + p_{22}g(\theta)} \right) \\ - \lambda \zeta_{\min}(\theta) \left[\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right]. \end{aligned} \quad (72)$$

After taking the first derivative, we obtain (73), given on the bottom of this page. (73) can further be expressed as

$$z [1 + \mu(\theta, z)z]^{-\eta-1} = \nu, \quad (74)$$

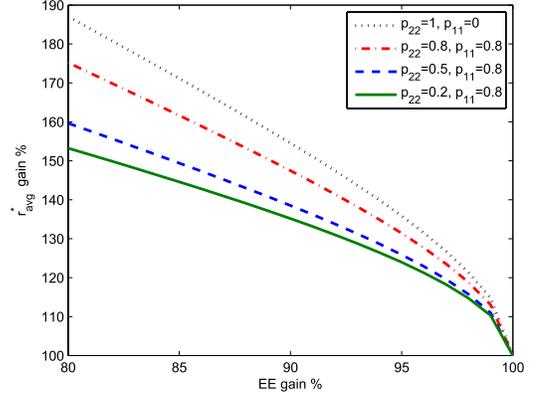


Fig. 10. Throughput r_{avg}^* gain % vs. EE gain %. $\theta = 1$. (Discrete Markov Source).

where we define ν as

$$\nu = \frac{\lambda \theta \zeta_{\min}(\theta)}{\eta \epsilon (1 + \lambda) P_{\text{ON}} \left(\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))(1-p_{11}-p_{22})g(\theta)+p_{22}} + \frac{1}{g(\theta)} \right)} \quad (75)$$

Using (74), we can derive the power allocation formula as

$$\mu(\theta, z) = \left[\frac{1}{\nu^{\frac{1}{1+\eta}} z^{\frac{\eta}{1+\eta}}} - \frac{1}{z} \right]^+. \quad (76)$$

We see that the power control formula for rate maximization under EE constraints is similar to that for maximizing EE. The key distinction lies in the formulation for ν in (75) which is different from (44).

In Fig. 10, we address the tradeoff between throughput and energy efficiency by solving the power control problem and determining the maximum throughput level under different energy efficiency constraints. More specifically, we plot the percentage gain in throughput by backing off from the maximal energy efficiency point, which is represented by the 100% EE gain.⁶ The figure shows us that decreasing the energy efficiency leads to significant improvement in throughput. For instance, 20% reduction from the maximal energy efficiency point results in 50 to 90% gain in throughput depending on the source characteristics. Even 1% decrease in energy efficiency generates about 10% gain on the throughput. We also note that the largest gain is realized in the case of constant arrival rate, and increasing burstiness reduces the throughput gain.

⁶Therefore, we can formulate EE gain percentage as $\frac{EE}{EE_{\text{max}}} \times 100\%$ where EE_{max} is the maximum energy efficiency. Similarly, throughput gain percentage is defined as $\frac{r_{\text{avg}}^*}{r_{\text{avg}, EE_{\text{max}}}} \times 100\%$ where $r_{\text{avg}, EE_{\text{max}}}$ is the average arrival rate at the maximum EE point.

$$(1 + \lambda) \frac{P_{\text{ON}}}{\theta} \eta z [1 + \mu(\theta, z)z]^{-\eta-1} \left(\frac{(1 - p_{11})(1 - p_{22})}{(1 - p_{11}g(\theta))(1 - p_{11} - p_{22})g(\theta) + p_{22}} + \frac{1}{g(\theta)} \right) - \lambda \zeta_{\min}(\theta) \frac{1}{\epsilon} = 0. \quad (73)$$

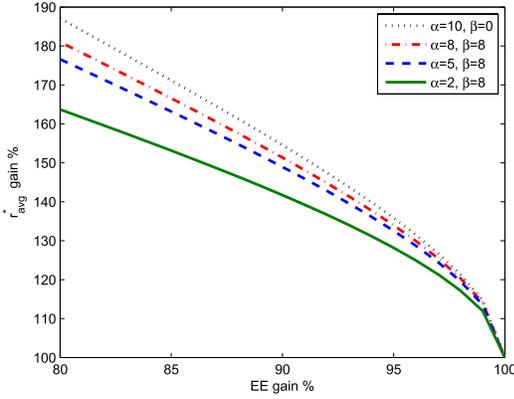


Fig. 11. Throughput r_{avg}^* gain % vs. EE gain %. $\theta = 1$. (Markov Fluid Source).

B. Markov Fluid Source

For Markov fluid source, Lagrangian function is given by

$$\begin{aligned} \mathcal{L}(\mu(\theta, z), \lambda) = & (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \frac{\alpha + \beta - \log_e g(\theta)}{\alpha - \log_e g(\theta)} \log_e g(\theta) \\ & - \lambda \zeta_{\min}(\theta) \left[\frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c \right], \end{aligned} \quad (77)$$

and the optimal power control policy has the same form as in (76) with ν defined as

$$\nu = \frac{\lambda \theta \zeta_{\min}(\theta)}{\eta \epsilon (1 + \lambda) P_{\text{ON}} \left(\frac{\alpha \beta}{g(\theta) (\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)} \right)}. \quad (78)$$

In Fig. 11, we again demonstrate the tradeoff between energy efficiency and throughput when the source is modeled as a Markov fluid. We immediately observe that having a small reduction in the energy efficiency results in substantial gain in the throughput. On the other hand, the percentage of the gain decreases as the burstiness of the source increases (i.e. by decreasing α while β is fixed).

C. Markov Modulated Poisson Processes

The Lagrangian $\mathcal{L}(\mu(\theta, z), \lambda)$ and the scaled Lagrange multiplier ν for the discrete-time and continuous-time Markov-modulated Poisson sources can be immediately obtained by replacing θ with $(e^\theta - 1)$ in the corresponding expressions for discrete Markov and Markov fluid sources, respectively.

VI. OPTIMAL POWER CONTROL WITH AVERAGE POWER CONSTRAINTS

In this section, we consider a setting in which throughput maximization is the sole concern of the wireless system, and we study the optimal power control strategy that maximizes

the throughput under an average power constraint. The optimization problem is formulated as

$$\max_{P(\theta, z) \geq 0} r_{\text{avg}}^*(\theta, z) \quad (79)$$

$$\text{subject to } \frac{1}{\epsilon} \mathbb{E} \{ P(\theta, z) \} + P_c \leq \bar{P}. \quad (80)$$

Note that the optimization problems studied in previous sections have no explicit average power constraints. However, implicitly average power constraints are imposed through the energy efficiency requirements due to the fact that energy efficiency eventually starts diminishing with increasing average transmit power level. However, explicit average power constraints can be addressed without much difficulty as we demonstrate in this section.

The optimization problem in this section is again convex. Normalizing all the terms in the constraint in (80) with the noise power $N_0 B$ and denoting the average SNR = $\frac{\bar{P}}{N_0 B}$, the Lagrangian function can be written as

$$\mathcal{L}(\mu(\theta, z), \lambda) = r_{\text{avg}}^*(\theta, z) - \lambda \left\{ \frac{1}{\epsilon} \mathbb{E} \{ \mu(\theta, z) \} + \mu_c - \text{SNR} \right\}. \quad (81)$$

In the following analysis, we obtain the power allocation function for different source models using a similar approach as in previous sections. Specifically, we initially evaluate the first derivative of the Lagrangian function with respect to $\mu(\theta, z)$ and make it equal to 0. For all sources, the optimal power control is in the same form as in (76) with different ν expressions which we describe below for each source.

The first derivative of the Lagrangian for the discrete Markov source is obtained in (82), as shown at the bottom of this page.

Using (82), ν is derived as

$$\nu = \frac{\lambda \theta}{\eta \epsilon P_{\text{ON}} \left(\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))(1-p_{11}-p_{22}g(\theta)+p_{22})} + \frac{1}{g(\theta)} \right)}. \quad (83)$$

We plot the throughput vs. average power curve in Fig. 12 where we take into account different discrete Markov sources. As noted before, the best performance is realized for the case of constant arrival rates (i.e., when $p_{22} = 1$ and $p_{11} = 0$), and the throughput degrades with increased burstiness. Comparing the performances with source models with parameters $p_{22} = 0.5$, $p_{11} = 0.8$ and $p_{22} = 0.2$, $p_{11} = 0.8$, we observe that ON probability, P_{ON} , becomes a dominant factor on performance as average power increases. Source with smaller P_{ON} (i.e., the one with transition probabilities $p_{22} = 0.2$, $p_{11} = 0.8$) has lower performance. On the other hand, when the average power is relatively low, this source outperforms the one with parameters $p_{22} = 0.5$, $p_{11} = 0.8$ since the metric $\frac{(1-p_{22})(p_{11}+p_{22})}{(1-p_{11})(2-p_{11}-p_{22})}$ is a more critical burstiness factor at low SNR values (as also discussed in Section IV-A). Indeed, in this case, the transition probabilities $p_{22} = 0.5$, $p_{11} = 0.8$

$$\frac{P_{\text{ON}}}{\theta} \eta z [1 + \mu(\theta, z)]^{-\eta-1} \left(\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))(1-p_{11}-p_{22}g(\theta)+p_{22})} + \frac{1}{g(\theta)} \right) - \frac{\lambda}{\epsilon} = 0 \quad (82)$$

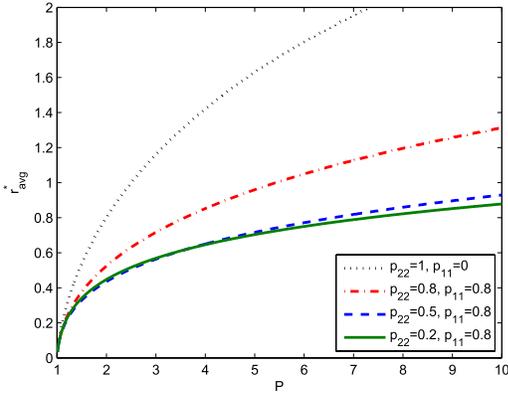


Fig. 12. Maximum average arrival rate r_{avg}^* vs. average power \bar{P} when $\theta = 1$. (Discrete Markov Source).

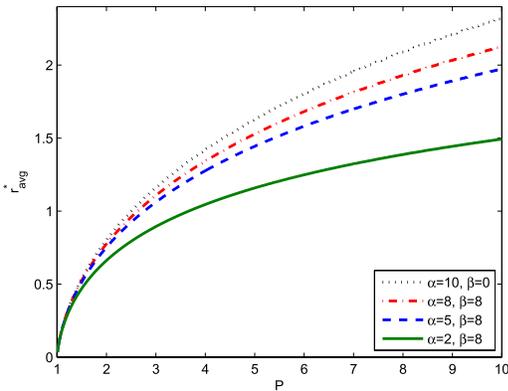


Fig. 13. Maximum average arrival rate r_{avg}^* vs. average power \bar{P} when $\theta = 1$. (Markov Fluid Source).

result in a larger value for $\frac{(1-p_{22})(p_{11}+p_{22})}{(1-p_{11})(2-p_{11}-p_{22})}$ indicating a more bursty source in the low-SNR regime.

For the Markov fluid case, we obtain (84), given on the bottom of this page as the first derivative of the Lagrangian with respect to $\mu(\theta, z)$. The parameter ν that we use in power allocation formula is given by

$$\nu = \frac{\lambda\theta}{\eta\epsilon P_{\text{ON}} \left(\frac{\alpha\beta}{g(\theta)(\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)} \right)}. \quad (85)$$

For Markov fluid source we demonstrate the throughput as a function of average power in Fig. 13. Similarly as before, burstiness hurts the performance.

For discrete-time and continuous-time MMPP, the first derivatives of the Lagrangian functions with respect to $\mu(\theta, z)$ are given, respectively, by (82) and (84) with $\frac{P_{\text{ON}}}{\theta}$ replaced by $\frac{P_{\text{ON}}}{(e^\theta - 1)}$, and with the corresponding threshold parameters ν given, respectively, by (83) and (85) when $\lambda\theta$ is replaced by $\lambda(e^\theta - 1)$.

VII. OPTIMAL POWER CONTROL IN MULTICHANNEL SYSTEMS

In this section, motivated by the fact that multicarrier channels employing orthogonal frequency division multiplexing (OFDM) can be regarded as multichannel systems, we extend our power control analysis to multichannel communication links. Note that in this case, the instantaneous transmission rates and effective capacity are given by (16) and (17), respectively.

In order to keep the analysis concise in this section, we only consider the problem of finding the optimal power allocation scheme that maximizes the throughput under a minimum energy efficiency constraint for discrete Markov and Markov fluid sources. The optimal power allocation problem can be expressed as the following convex optimization problem:

$$\max_{\mu(\theta, z) \geq 0} r_{\text{avg}}^*(\theta, z) \quad (86)$$

$$\text{subject to } \frac{r_{\text{avg}}^*(\theta, z)}{\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\theta, \mathbf{z}) \} + \mu_c} \geq \zeta_{\min}(\theta) \quad (87)$$

where we define $\mu(\theta, \mathbf{z}) = [\mu_1(\theta, \mathbf{z}), \dots, \mu_K(\theta, \mathbf{z})]$, $\mu_k(\theta, \mathbf{z}) = \frac{P_k(\theta, \mathbf{z})}{N_0 \frac{B}{K}}$ and $\mu_c = \frac{P_c}{N_0 \frac{B}{K}}$. ζ_{\min} is the minimum required energy efficiency level.

We can further rewrite the constraint as

$$-r_{\text{avg}}^*(\theta, z) + \zeta_{\min}(\theta) \left[\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\theta, \mathbf{z}) \} + \mu_c \right] \leq 0. \quad (88)$$

Now, the Lagrangian becomes

$$\begin{aligned} \mathcal{L}(\mu(\theta, \mathbf{z}), \lambda) = & r_{\text{avg}}^*(\theta, z) \\ & - \lambda \left\{ -r_{\text{avg}}^*(\theta, z) \right. \\ & \left. + \zeta_{\min}(\theta) \left[\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\theta, \mathbf{z}) \} + \mu_c \right] \right\} \\ & - \sum_{k=1}^K \lambda_k \mu_k(\theta, \mathbf{z}). \end{aligned} \quad (89)$$

To determine the optimal power control policy, we have to consider the solution of

$$\frac{\partial \mathcal{L}(\mu(\theta, \mathbf{z}), \lambda)}{\partial \mu(\theta, \mathbf{z})} = 0. \quad (90)$$

If we have $\mu_i(\theta, \mathbf{z}) > 0$ for $i \in \mathcal{N}_0 = \{1, \dots, K\}$, complementary slackness dictates that the corresponding Lagrangian multiplier λ_i is zero [27]. In the rest of the analysis we exploit this property.

A. Discrete Markov Source

First, let us define

$$g(\theta, \mathbf{z}) = \mathbb{E} \left\{ \prod_{k=1}^K [1 + \mu_k(\mathbf{z})z_k]^{-\frac{\eta}{K}} \right\} \quad (91)$$

$$\frac{P_{\text{ON}}}{\theta} \eta z [1 + \mu(\theta, z)z]^{-\eta-1} \left(\frac{\alpha\beta}{g(\theta)(\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)} \right) - \frac{\lambda}{\epsilon} = 0. \quad (84)$$

where $\eta = \theta TB \log_2 e$. The Lagrangian for the discrete Markov source is expressed in (92), as shown at the bottom of this page. Initially, we assume that we utilize all of the subchannels for transmission. Then, we can immediately state that $\lambda_i = 0$ for all i and derive the optimality equations as in (93), as shown at the bottom of this page by, calculating the derivatives with respect to μ_i for $i \in \mathcal{N}_0$.

We simplify (93) as

$$\nu = z_k [1 + \mu_k(\mathbf{z})z_k]^{-\frac{\eta}{K}-1} \prod_{i \neq k} [1 + \mu_i(\mathbf{z})z_i]^{-\frac{\eta}{K}}, \quad k \in \mathcal{N}_0 \quad (94)$$

where ν is a scaled Lagrangian multiplier

$$\nu = \frac{\lambda K \zeta_{\min}(\theta) \log_e 2}{\epsilon(1 + \lambda) P_{\text{ON}} T B \left(\frac{(1-p_{11})(1-p_{22})}{(1-p_{11}g(\theta))(1-p_{11}-p_{22})g(\theta)+p_{22}} + \frac{1}{g(\theta)} \right)}. \quad (95)$$

By solving equations in (94), the optimal power allocation can be written as

$$\mu_k(\theta, \mathbf{z}) = \frac{1}{\nu^{\frac{1}{1+\eta}} \prod_{i \in \mathcal{N}_0} z_i^{\frac{\eta}{K(1+\eta)}}} - \frac{1}{z_k}, \quad k \in \mathcal{N}_0. \quad (96)$$

Now we define \mathcal{N}_1 as

$$\mathcal{N}_1 = \left\{ k \in \mathcal{N}_0 \mid \frac{1}{\nu^{\frac{1}{1+\eta}} \prod_{i \in \mathcal{N}_0} z_i^{\frac{\eta}{K(1+\eta)}}} - \frac{1}{z_k} > 0 \right\} \quad (97)$$

If $\mathcal{N}_1 = \mathcal{N}_0$ holds, then (96) is the optimal solution, otherwise we need to apply a recursive strategy which we describe as an algorithm in Table I.

Remark 1: In the algorithm, basically, we first employ the formula in (96) for the power allocation. Then, if all power levels are above zero, we stop the algorithm. Otherwise, for subchannels with power levels less than zero, we do not allocate any power and we cease using these subchannels in the algorithm.

Fig. 14 depicts the energy efficiency as a function of the maximum average arrival rate (or equivalently throughput) with varying source characteristics and the number of subchannels. The random arrivals are modeled as a discrete Markov process. When we have a higher ON probability or more subchannels, the system has better performance in terms of energy efficiency as the maximum EE point is the highest out of all scenarios when $P_{\text{ON}} = 1$ and $K = 4$. The source burstiness is in general an important factor in the presence of

TABLE I
THE OPTIMAL POWER CONTROL ALGORITHM THAT MAXIMIZES THROUGHPUT GIVEN A MINIMUM EE CONSTRAINT

Algorithm 1 The optimal power control algorithm that maximizes throughput given a minimum EE constraint

- 1: Given \mathcal{N}_1 compute $N_1 = |\mathcal{N}_1|$;
- 2: Initialize $k = 1$;
- 3: **while** $\mathcal{N}_k \neq \mathcal{N}_{k-1}$ **do**
- 4: $\mathcal{N}_{k+1} = \left\{ n \in \mathcal{N}_k \mid \frac{1}{\nu^{\frac{1}{K+N_k\eta}} \prod_{i \in \mathcal{N}_k} z_i^{\frac{\eta}{K+N_k\eta}}} - \frac{1}{z_n} > 0 \right\}$;
- 5: $N_{k+1} = |\mathcal{N}_{k+1}|$;
- 6: $k = k + 1$;
- 7: Define $\mathcal{N}^* = \mathcal{N}_k$ and $K^* = |\mathcal{N}_k|$;
- 8: $\mu_n(\theta, \mathbf{z}) = \begin{cases} \frac{1}{\nu^{\frac{1}{K+K^*\eta}} \prod_{i \in \mathcal{N}^*} z_i^{\frac{\eta}{K+K^*\eta}}} - \frac{1}{z_k} & \text{for } n \in \mathcal{N}^*; \\ 0 & \text{otherwise.} \end{cases}$

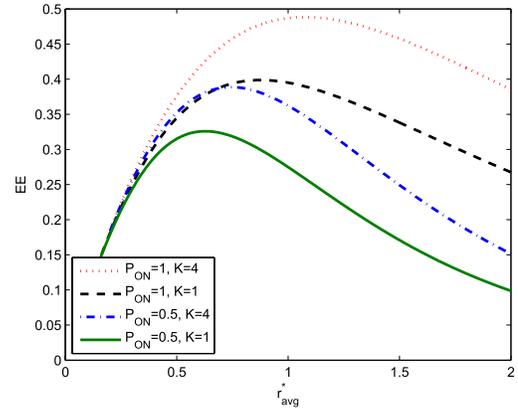


Fig. 14. Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\eta = 1$.

QoS constraints because increasing burstiness by decreasing the ON probability makes the system more susceptible to buffer overflows. To avoid higher buffer overflow probabilities, the system supports smaller throughput with the same energy budget. As noted above, having more subchannels improves the energy efficiency. Essentially, when we use a single channel, the random variations in the wireless channel, which can be detrimental in the presence of buffer overflow constraints, have a more significant impact.

In Fig. 15, we analyze the tradeoff between throughput and energy efficiency. Similarly as in Fig. 10, we describe the maximum EE point (i.e., the peak of the bell-shaped EE curves in Fig. 14) as 100% on the x -axis and decrease the energy efficiency while computing the gain in the throughput.

$$\mathcal{L}(\mu(\theta, \mathbf{z}), \lambda) = (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \log_e \left(\frac{1 - p_{11}g(\theta)}{(1 - p_{11} - p_{22})g^2(\theta) + p_{22}g(\theta)} \right) - \lambda \zeta_{\min}(\theta) \left[\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\mathbf{z}) \} + \mu_c \right] - \sum_{k=1}^K \lambda_k \mu_k(\mathbf{z}). \quad (92)$$

$$(1 + \lambda) \frac{P_{\text{ON}}}{\theta} \frac{\eta z_k}{K} [1 + \mu_k(\mathbf{z})z_k]^{-\frac{\eta}{K}-1} \prod_{i \neq k} [1 + \mu_i(\mathbf{z})z_i]^{-\frac{\eta}{K}} \left(\frac{(1 - p_{11})(1 - p_{22})}{(1 - p_{11}g(\theta))(1 - p_{11} - p_{22})g(\theta) + p_{22}} + \frac{1}{g(\theta)} \right) - \lambda \zeta_{\min}(\theta) \frac{1}{\epsilon} = 0 \quad (93)$$

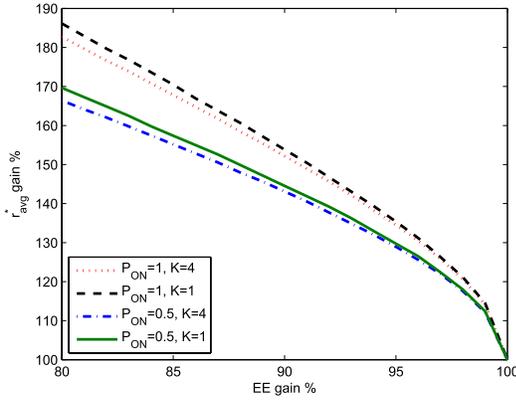


Fig. 15. Throughput r_{avg}^* gain % vs. EE gain % when $N_0 = 1$, $\eta = 1$.

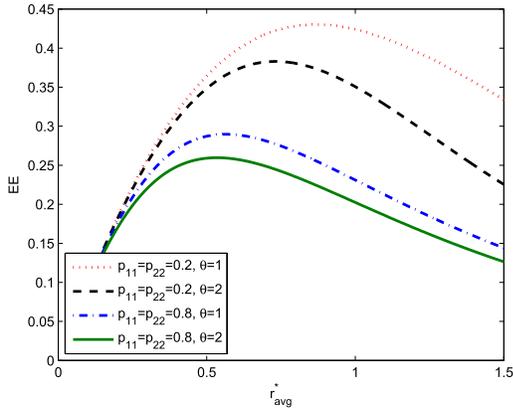


Fig. 16. Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $K = 4$.

This figure shows how much throughput can be improved by sacrificing from the maximum energy efficiency. The observation from Fig. 15 is that throughput improvement for less bursty sources is higher. Also, with smaller number of subchannels, we observe a larger improvement.

In Fig. 16 where we again depict the energy efficiency as a function of the throughput, we investigate how the system performs under different levels of burstiness and QoS requirements. We notice that burstiness does not depend on P_{ON} only. Although systems with $p_{11} = p_{22} = 0.2$ and $p_{11} = p_{22} = 0.8$ have the same $P_{\text{ON}} = 0.5$, they perform differently. Sources with higher transition probabilities from one state to a different state (i.e., higher p_{12} and p_{21}) exhibit reduced burstiness, and hence we have better performance with the source having $p_{11} = p_{22} = 0.2$. Additionally, more stringent QoS constraints (i.e. higher values of θ) clearly

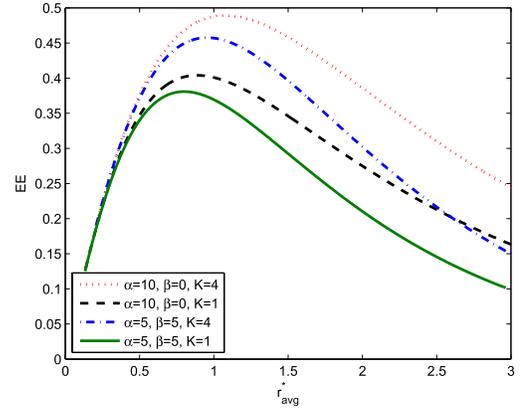


Fig. 17. Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\eta = 1$.

reduce the energy efficiency of the system no matter what the source characteristics are.

B. Markov Fluid Source

The Lagrangian function of the Markov fluid source is given at the bottom of this page in (98).

Again, we initially assume that we utilize all subchannels for transmission. Then, we can immediately state that $\lambda_i = 0$ for all i and derive the following optimality equation in (99), shown at the bottom of this page, by calculating the derivatives with respect to μ_i for $i \in \mathcal{N}_0$.

We can simplify (99) as

$$\nu = z_k [1 + \mu_k(\mathbf{z})z_k]^{-\frac{\eta}{K}-1} \prod_{i \neq k} [1 + \mu_i(\mathbf{z})z_i]^{-\frac{\eta}{K}}, \quad k \in \mathcal{N}_0 \quad (100)$$

where ν is defined as

$$\nu = \frac{\lambda K \zeta_{\min}(\theta) \log_e 2}{\epsilon(1 + \lambda) P_{\text{ON}} T B \left(\frac{\alpha\beta}{g(\theta)(\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)} \right)}. \quad (101)$$

Note that, the formulation of ν in (100) is exactly the same as in (94). Thus, the optimal power control for the case with the Markov fluid source follows from (96) and algorithm from Table I.

In Fig. 17, we plot the energy efficiency vs. maximum average arrival rate curves for Markov fluid arrivals. We immediately observe that having more subchannels again improves the performance in terms of energy efficiency. Also, the maximum energy efficiency is achieved at a larger throughput level when P_{ON} or K (number of subchannels) increases. Additionally, we notice that at high SNR levels (equivalently for large r_{avg}^* values), increased burstiness can offset improvements due to

$$\mathcal{L}(\mu(\theta, \mathbf{z}), \lambda) = (1 + \lambda) \frac{P_{\text{ON}}}{\theta} \frac{\alpha + \beta - \log_e g(\theta)}{\alpha - \log_e g(\theta)} \log_e g(\theta) - \lambda \zeta_{\min}(\theta) \left[\frac{1}{\epsilon} \sum_{k=1}^K \mathbb{E} \{ \mu_k(\theta, \mathbf{z}) \} + \mu_c \right] - \sum_{k=1}^K \lambda_k \mu_k(\theta, \mathbf{z}) \quad (98)$$

$$(1 + \lambda) \frac{P_{\text{ON}}}{\theta} \frac{\eta z_k}{K} [1 + \mu_k(\mathbf{z})z_k]^{-\frac{\eta}{K}-1} \prod_{i \neq k} [1 + \mu_i(\mathbf{z})z_i]^{-\frac{\eta}{K}} \left(\frac{\alpha\beta}{g(\theta)(\alpha - \log_e g(\theta))^2} + \frac{1}{g(\theta)} \right) - \lambda \zeta_{\min}(\theta) \frac{1}{\epsilon} = 0. \quad (99)$$

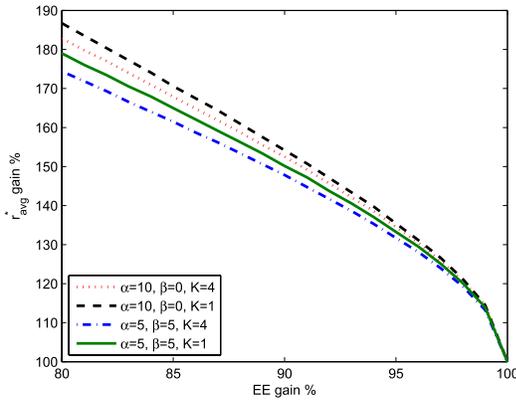


Fig. 18. Throughput r_{avg}^* gain % vs. EE gain % when $N_0 = 1$, $\eta = 1$.

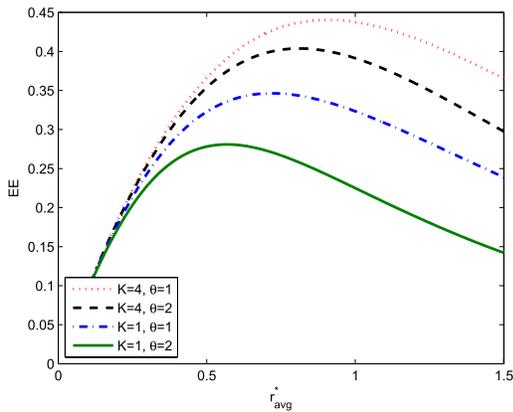


Fig. 19. Energy efficiency EE vs. maximum average arrival rate r_{avg}^* when $N_0 = 1$, $\alpha = \beta = 5$.

the increased number of subchannels, as evidenced by the crossover between the dashed curve (for which $\alpha = 10$, $\beta = 0$ and hence the arrival rate is constant, and $K = 1$) and dot-dashed curve (for which $\alpha = 5$, $\beta = 5$, $K = 4$).

For Markov fluid sources, we analyze the energy efficiency and throughput tradeoff in Fig. 18. From Fig. 17, we observe the steep loss in energy efficiency for more bursty sources. This observation is further reflected in Fig. 18 as the throughput gain is lower for more bursty sources when we have the same percentage of sacrifice from the energy efficiency. Overall, we also note that instead of working at the optimal energy efficiency point, if we reduce the energy efficiency by about 20%, we can obtain gains, reaching up to almost twice the throughput levels.

Finally, in Fig. 19 we plot the energy efficiency curves with varying number of subchannels and QoS constraints for the Markov fluid source. Again, our previous observations are verified as increasing the number of subchannels K or decreasing the value of QoS exponent θ enhances the energy efficiency.

VIII. CONCLUSION

In this paper, we have investigated energy-efficient power control strategies in fading channels when the data arrivals are modeled as Markovian processes, statistical QoS constraints

are imposed, and circuit power consumption is taken into account. Considered arrival models include discrete Markov, Markov fluid, and MMPP sources. First, we have determined the optimal power control policies that maximize the energy efficiency, which is defined as the maximum throughput normalized by the total power consumption (in bits per joule), by solving convex optimization problems. We have demonstrated that increased source burstiness can dramatically lower the energy efficiency, and source statistics alters the energy-efficiency-maximizing power control scheme primarily through the threshold parameter ν .

We have also identified the optimal power control that maximizes the throughput under energy efficiency constraints for different source models. We have analyzed the tradeoff between throughput and energy efficiency and noted that backing off even by a small percentage from the maximum energy efficiency point can lead to substantial gains in throughput with higher gains if the source is less bursty. To complete the framework, we have analyzed throughput-maximizing power control under average power constraints.

Finally, we have extended our analysis to multichannel systems and determined the optimal power control to maximize the throughput under energy efficiency constraints. We have shown that increasing the number of subchannels improves the energy efficiency.

REFERENCES

- [1] M. Peterson. (Sep. 2013). *ICT at 10% of Global Electricity Consumption*. [Online]. Available: <http://www.vertatique.com/ict-10-global-energy-consumption>
- [2] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tut.*, vol. 15, no. 1, pp. 167–178, 1st Quart., 2013.
- [3] G. Li *et al.*, "Energy efficient wireless communications: Tutorial, survey, and open issues," *IEEE Wireless Commun.*, vol. 18, no. 6, pp. 28–35, Dec. 2011.
- [4] S. Tanwir and H. Perros, "A survey of VBR video traffic models," *IEEE Commun. Surveys Tut.*, vol. 15, no. 4, pp. 1778–1802, 4th Quart., 2013.
- [5] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update (2015–2020)*, accessed on Oct. 12, 2016. [Online]. Available: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf>
- [6] C.-S. Chang, "Stability, queue length, and delay of deterministic and stochastic queuing networks," *IEEE Trans. Autom. Control*, vol. 39, no. 5, pp. 913–931, May 1994.
- [7] C.-S. Chang and J. A. Thomas, "Effective bandwidth in high-speed digital networks," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 6, pp. 1091–1100, Aug. 1995.
- [8] C.-S. Chang and T. Zajic, "Effective bandwidths of departure processes from queues with time varying capacities," in *Proc. IEEE 14th Annu. Joint Conf. Comput. Commun. Soc. Bringing Inf. People (INFOCOM)*, Apr. 1995, pp. 1001–1009.
- [9] C.-S. Chang, *Performance Guarantees in Communication Networks*. Berlin, Germany: Springer-Verlag, 2000.
- [10] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 630–643, Jul. 2003.
- [11] A. I. Elwalid and D. Mitra, "Effective bandwidth of general Markovian traffic sources and admission control of high speed networks," *IEEE ACM Trans. Netw.*, vol. 1, no. 3, pp. 329–343, Jun. 1993.
- [12] G. Kesidis, J. Walrand, and C.-S. Chang, "Effective bandwidths for multiclass Markov fluids and other ATM sources," *IEEE ACM Trans. Netw.*, vol. 1, no. 4, pp. 424–428, Aug. 1993.
- [13] Q. Du and X. Zhang, "Statistical QoS provisionings for wireless unicast/multicast of Multi-Layer video streams," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 420–433, Apr. 2010.

- [14] M. C. Gursoy, D. Qiao, and S. Velipasalar, "Analysis of energy efficiency in fading channels under QoS constraints," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4252–4263, Aug. 2009.
- [15] L. Musavian and T. Le-Ngoc, "Energy-efficient power allocation for delay-constrained systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 3554–3559.
- [16] G. Ru, H. Li, L. Liu, Z. Hu, and Y. Gan, "Energy efficiency of hybrid cellular with heterogeneous QoS provisions," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 1003–1006, Jun. 2014.
- [17] L. Liu, Y. Yang, J.-F. Chamberland, and J. Zhang, "Energy-efficient power allocation for delay-sensitive multimedia traffic over wireless systems," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2038–2047, Jun. 2014.
- [18] A. Helmy, L. Musavian, and T. Le-Ngoc, "Energy-efficient power adaptation over a frequency-selective fading channel with delay and power constraints," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4529–4541, Sep. 2013.
- [19] C. Xiong, G. Y. Li, Y. Liu, Y. Chen, and S. Xu, "Energy-efficient design for downlink OFDMA with delay-sensitive traffic," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 3085–3095, Jun. 2013.
- [20] M. Ozmen and M. C. Gursoy, "Wireless throughput and energy efficiency with random arrivals and statistical queueing constraints," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1375–1395, Mar. 2016.
- [21] L. Musavian and Q. Ni, "Effective capacity maximization with statistical delay and effective energy efficiency requirements," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3824–3835, Jul. 2015.
- [22] C. She, C. Yang, and L. Liu, "Energy-efficient resource allocation for MIMO-OFDM systems serving random sources with statistical QoS requirement," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4125–4141, Nov. 2015.
- [23] L. Liu, P. Parag, J. Tang, W.-Y. Chen, and J.-F. Chamberland, "Resource allocation and quality of service evaluation for wireless communication systems using fluid models," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1767–1777, May 2007.
- [24] J. Tang and X. Zhang, "Quality-of-service driven power and rate adaptation for multichannel communications over wireless links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4349–4360, Dec. 2007.
- [25] W. Cheng, X. Zhang, and H. Zhang, "Heterogeneous statistical QoS provisioning over 5G wireless full-duplex networks," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr./May 2015, pp. 55–63.
- [26] X. Zhang and J. Tang, "Power-delay tradeoff over wireless networks," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3673–3684, Sep. 2013.
- [27] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.
- [28] D. Qiao, M. C. Gursoy, and S. Velipasalar, "Effective capacity of two-hop wireless communication systems," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 873–885, Feb. 2013.



Mustafa Ozmen received the B.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2011. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering and Computer Science, Syracuse University. His research interests include wireless communications, information theory, quality-of-service provisioning, queueing analysis, random arrivals, and energy-efficient transmission techniques. He has received the ISIT Student Travel Awards in 2012 and 2015.



M. Cenk Gursoy received the B.S. degree (Hons.) in electrical and electronics engineering from Boğaziçi University, Istanbul, Turkey, in 1999, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 2004. In 2000, he was with Lucent Technologies, Holmdel, NJ, where he conducted performance analysis of DSL modems. From 2004 to 2011, he was a faculty member with the Department of Electrical Engineering, University of Nebraska–Lincoln (UNL). He is currently an Associate Professor with the Department of Electrical Engineering and Computer Science, Syracuse University.

His research interests are in the general areas of wireless communications, information theory, communication networks, and signal processing. He is a member of the Editorial Boards of the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS—Series on Green Communications and Networking, and *Physical Communication* (Elsevier). He was a recipient of the Gordon Wu Graduate Fellowship from Princeton University from 1999 to 2003. He received an NSF Career Award in 2006, the *EURASIP Journal of Wireless Communications and Networking* Best Paper Award, the UNL College Distinguished Teaching Award, and the Maude Hammond Fling Faculty Research Fellowship. He served as the Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2010 to 2015 and the IEEE COMMUNICATIONS LETTERS from 2012 to 2014.