Adaptive Modulation in Wireless Networks with Smoothed Flow Utility

Ekine Akuiyibo, Stephen Boyd, and Daniel O'Neill
Electrical Engineering Department
Stanford University
Stanford CA 94305-9510
Email: {ekine, boyd, dconeill}@stanford.edu.

***Abstract*—We investigate flow rate optimization on a wireless link with randomly varying channel gain using techniques from adaptive modulation and network utility maximization. We consider the problem of choosing the data flow rate to optimally trade off average transmit power and the average utility of the smoothed data flow rate. The smoothing allows us to model the demands of an application that can tolerate variations in flow over a certain time interval; we will see that this smoothing leads to a substantially different optimal data flow rate policy than without smoothing. We pose the problem as a convex stochastic control problem. For the case of a single flow, the optimal data flow rate policy can be numerically computed using stochastic dynamic programming. For the case of multiple data flows on a single link, we propose an approximate dynamic programming approach to obtain suboptimal data flow rate policies. We illustrate, through numerical examples, that these approximate policies perform very well.**

I. Introduction

In wireless communications, controlling a link's transmit rate and power depending on channel conditions is called *adaptive modulation* (AM), see, *e.g.*, [1], [2], [3]. One drawback of AM is that as a physical layer technique it has no knowledge of upper layer optimization protocols. Network utility maximization (NUM) is used for upper layer protocols, see, *e.g.*, [4], [5], [6]. In the NUM framework, performance of an upper layer protocol (*e.g.*, TCP) is determined by *utility* of flow attributes, *e.g.*, utility of link flow rate. Combining both AM and NUM techniques, *e.g.*, [7], obtains improved performance over either one technique alone.

Our approach combines both AM and NUM, but is unique in several respects: We consider the utility of smoothed flows, we consider an infinite discrete-time network, and we consider multiple flows over the same wireless link. We model the system as a convex stochastic control problem and propose a simple suboptimal flow control policy. In [8] we introduced our linear smoothing model which highlighted distinctions between our resultant multi-period problem and some single-period variants, *e.g.*, [7]. In this paper we give another interpretation of the multi-period problem and propose simple, approximate, flow control policies.

These flow policies serve as the theoretical equivalent of a network protocol. Suppose two network applications, say, video streaming and batch file transfer, represented by two separate flows, share a single wireless link. Suppose also that the file transfer application is less time-sensitive, *i.e.*, compared to video streaming, file transfer can tolerate more variations in flow over a certain time interval. Our smoothing model captures these heterogeneous flow rate demands. At each time t, our flow policy must decide for each flow how many data packets to inject into the wireless link, taking into account the different levels of smoothing, the time-varying link gains, and the power required by the link to support the data packets. We say that the flow representing file transfer has *higher* smoothing. (We review the smoothing model in the next section.)

We show that the optimal policy includes a "no-transmit" zone, *i.e.*, a region in the smoothed flow/channel gain plane, where the optimal flow rate is zero. Not suprisingly, the optimal flow policy can be roughly described as waiting until the channel gain is large, or until the smoothed flow has fallen to a low level, at which point we transmit. Roughly speaking, the higher the level of smoothing, the longer we can afford to wait for a large channel gain before transmitting. The average power required to support a given utility level decreases, sometimes dramatically, as the level of smoothing increases.

We show that the optimal policy for the case of a single flow is readily computed numerically, working from Bellman's characterization [9] of the optimal policy, and the behavior of the optimal policy is not particularly sensitive to the details of the utility functions, smoothing levels, or power functions. For the case of multiple flows, we cannot easily compute the optimal policy. For this case we propose an approximate policy, based on approximate dynamic programming (ADP), see, *e.g.*, [10], [11], [12]. By computing an upper bound, by allowing the flow control policy to use future values of channel gain, we show in numerical experiments that these suboptimal policies can perform very well.

This paper is organized as follows: Section II describes the problem setup and system performance model and Section III characterizes the optimal flow policy. Sections IV and V give our policies along with numerical examples for both the single and multi-flow cases. We conclude in Section VI.

II. System Model

A. Average smoothed flow utility

A wireless communication link supports n data flows in a channel that varies with time, which we model using discrete-time intervals $t = 0, 1, 2, \ldots$. We let $f_t \in \mathbf{R}_+^n$ be the data flow rate vector on the link, where $(f_t)_j$, $j = 1, \ldots, n$, is the jth flow's data rate at time t, and $\mathbf{R}_+$ denotes the set of positive numbers. We let $F_t = \mathbf{1}^T f_t$ denote the total flow rate over all flows, where $\mathbf{1}$ is the vector with all entries one. The flows, and the total flow rate, will depend on the random channel gain (through the flow policy, described below) and so are random variables.

We work with a *smoothed* version of the flow rates, which is meant to capture the tolerance of the applications using the data flows to time variations in data rate. This was introduced in [13] using delivery contracts, in which the utility is a function of the total flow over a given time interval. Here, we use instead, a simple first order linear smoothing. At each time t, the smoothed data flow rate vector $s_t \in \mathbf{R}_+^n$ is given by

$$s_{t+1} = \Theta s_t + (I - \Theta) f_t, \qquad t = 0, 1, \ldots,$$

where $\Theta = \mathbf{diag}(\theta)$, and $\theta_j \in [0, 1)$, $j = 1, \ldots, n$, is the smoothing parameter for the jth flow, and we take $s_0 = 0$. Thus we have $(s_t)_j = \sum_{\tau=0}^{t-1}(1-\theta_j)\theta_j^{t-1-\tau}(f_\tau)_j$.

The smoothing parameter θ_j determines the level of smoothing on flow j. Small smoothing parameter values (θ_j close to zero) correspond to light smoothing; large values (θ_j close to one) correspond to heavy smoothing. (Note that $\theta_j = 0$ means that flow j is not smoothed; we have $(s_t)_j = (f_t)_j$.) The level of smoothing can be related to the time scale over which the smoothing occurs. We define $T_j = 1/\log(1/\theta_j)$ to be the *smoothing time* associated with flow j. Roughly speaking, the smoothing time is the time interval over which the effect of a flow on the smoothed flow decays by a factor $1/e$. Light smoothing corresponds to short smoothing times, while heavy smoothing corresponds to longer smoothing times.

We associate with each smoothed flow rate $(s_t)_j$ a concave nondecreasing utility function $U_j : \mathbf{R}_+ \to \mathbf{R}$, where the utility of $(s_t)_j$ is $U_j((s_t)_j)$. The average utility derived over all flows, over all time is

$$\bar{U} = \lim_{N\to\infty} \mathbf{E}\, \frac{1}{N} \sum_{t=0}^{N-1} U(s_t),$$

where $U(s_t) = U_1((s_t)_1) + \cdots + U_n((s_t)_n)$. Here, the expectation is over the smoothed flows s_t, and we are assuming that the expectations and limit above exist.

While most of our results will hold for more general utilities, we will focus on the family of power utility functions, defined for $x > 0$ as

$$U(x) = \beta x^\alpha, \tag{1}$$

parametrized by $\alpha \in (0, 1)$ and $\beta > 0$. The parameter α sets the curvature (or risk aversion), while β sets the overall weight of the utility. (For small values of α, U approaches a log utility.)

Before proceeding, we make some general comments on our use of smoothed flows. The smoothing can be considered a type of time averaging; then we apply a concave utility function; and finally, we average this utility. The time averaging and utility function operations do not commute, except in the case when the utility is linear (or affine). Jensen's inequality tells us that average smoothed utility is greater than or equal to the average utility applied directly to the flow rates. So the time smoothing step does affect our average utility; we will see later that it has a dramatic effect on the optimal flow policy.

B. Average power

We model the wireless channel with time-varying positive gain parameters g_t, $t = 0, 1, \ldots$, which we assume are independent identically distributed (IID), with known distribution. At each time t, the gain parameter affects the power P_t required to support the total data flow rate F_t. The power P_t is given by

$$P_t = \phi(F_t, g_t),$$

where $\phi : \mathbf{R}_+ \times \mathbf{R}_{++} \to \mathbf{R}_+$ is increasing and convex in F_t for each value of g_t. (The notation $\mathbf{R}_{++}$ is the set of positive numbers.)

While our results will hold for the more general case, we will focus on the more specific power function described here. We suppose that the signal-to-interference-and-noise ratio (SINR) of the channel is given by $g_t P_t$. (Here g_t includes the effect of time-varying channel gain, and noise.) The channel capacity is then $\mu \log(1 + g_t P_t)$, where μ is a constant; this must equal at least the total flow rate F_t, so we obtain

$$P_t = \phi(F_t, g_t) = \frac{e^{F_t/\mu} - 1}{g_t}. \tag{2}$$

The total average power is given by

$$\bar{P} = \lim_{N\to\infty} \mathbf{E}\, \frac{1}{N} \sum_{t=0}^{N-1} P_t,$$

where again, we are assuming that the expectations and limit exist. We set $\mu = 1$ for simplicity.

C. Flow rate control problem

The overall objective is to maximize a weighted difference between average utility and average power,

$$J = \bar{U} - \lambda \bar{P}, \tag{3}$$

where $\lambda \in \mathbf{R}_{++}$ is used to trade off average utility and power.

We require that the flow policy is causal, *i.e.*, when f_t is chosen, we know the previous and current values of the flows, smoothed flows, and channel gains. Standard arguments in stochastic control (see, *e.g.*, [14], [15]) can be used to conclude that, without loss of generality, we can assume that the flow control policy has the form

$$f_t = \varphi(s_t, g_t), \tag{4}$$

where $\varphi : \mathbf{R}_+^n \times \mathbf{R}_{++} \to \mathbf{R}_+^n$. In other words, the policy depends only on the current smoothed flows, and the current channel gain.

The flow rate control problem is to choose the flow rate policy φ to maximize the objective (3). This is a standard convex stochastic control problem, with linear dynamics. We let $J^\star$ be the optimal objective value, and $\varphi^\star$ be an optimal policy.

III. Optimal policy characterization

A. Static case

We first consider the special case $\Theta = 0$, where there is no smoothing. We have $s_t = f_{t-1}$, so the average smoothed utility is then the same as the average utility. In this case the stochastic control problem reduces to a single-period optimization problem. At each time t, we simply choose f_t to maximize $U(f_t) - \lambda P_t$. Several variations of the static case have been studied in the literature; see, *e.g.*, [7].

B. General case

We now consider the more general case, with smoothing. We can characterize the optimal flow rate policy $\varphi^\star$ using stochastic dynamic programming and a form of Bellman's equation [9]. The optimal flow rate policy has the form

$$\varphi^\star(z,g) = \underset{w \geq 0}{\operatorname{argmax}} \left(V(z^+) - \lambda \phi(\mathbf{1}^T w, g) \right), \tag{5}$$

where $z^+ = \Theta z + (I - \Theta) w$ and $V : \mathbf{R}_+^n \to \mathbf{R}$ is the Bellman value function. The value function (and optimal value) is characterized via the fixed point equation

$$J^\star + V = \mathcal{T} V, \tag{6}$$

where for any function $W : \mathbf{R}_+^n \to \mathbf{R}$, the Bellman operator $\mathcal{T}$ is given by

$$(\mathcal{T}W)(z) = U(z) + \mathbf{E} \left(\max_{w \geq 0} \left(W(z^+) - \lambda \phi(\mathbf{1}^T w, g) \right) \right),$$

where the expectation is over g. The fixed point equation and Bellman operator are invariant under adding a constant; that is, we have $\mathcal{T}(W + a) = \mathcal{T}W + a$, for any constant (function) a, and similarly, V satisfies the fixed point equation if and only if $V + a$ does. So without loss of generality we can assume that for some fixed smoothed flow state $z^{\rm nrm}$, $V(z^{\rm nrm}) = 0$. (If $0 \in \mathbf{dom}\, U$, a natural choice is $z^{\rm nrm} = 0$.)

The optimal value function V is difficult to find analytically. We solve for V numerically using value iteration [9], [14]. Value iteration repeatedly applies $\mathcal{T}$ until we find $J^\star$ and V that satisfy (6). We take $V^{(0)} = 0$, and repeat the following iteration, for $k = 0, 1, \ldots$

1) $\tilde{V}^{(k)} = \mathcal{T} V^{(k)}$. (Apply Bellman operator.)
2) $J^{(k)} = \tilde{V}^{(k)}(z^{\rm nrm})$. (Estimate optimal value.)
3) $V^{(k+1)} = \tilde{V}^{(k)} - J^{(k)}$. (Normalize.)

Value iteration in average cost problems is not guaranteed to converge. For average cost problems with compact support, convergence is proven in [16]. Working with finite spaces we assume therefore that if $J^{(k)}$ and $V^{(k)}$ converge, they converge to $J^\star$ and V, respectively.

C. No transmit region

From the form of the optimal policy, we see that $\varphi(z,g) = 0$ if and only if $w = 0$ is optimal for the (convex) problem

$$\begin{array}{ll} \text{maximize} & V(z^+) - \lambda \phi(\mathbf{1}^T w, g) \\ \text{subject to} & w \geq 0, \end{array}$$

with variable $w \in \mathbf{R}^n$. This is the case if and only if

$$(I - \Theta) \nabla V(\Theta z) + \lambda \phi'(0, g) \mathbf{1} \leq 0,$$

where ϕ' is the derivative of ϕ with respect to its first argument. (See, *e.g.*, [17, p142].) We can rewrite this as

$$\frac{\partial V}{\partial z_i}(\Theta z) \leq \frac{\lambda \phi'(0,g)}{1 - \theta_i}, \quad i = 1, \ldots, n.$$

Using the specific power function (2) associated with the log capacity formula, we obtain

$$\nabla V(\Theta z) \leq \frac{\lambda}{g} \left(\frac{1}{1 - \theta_1}, \ldots, \frac{1}{1 - \theta_n} \right)$$

as the necessary and sufficient condition under which $\phi^\star(z,g) = 0$. Since ∇V is decreasing, the above equation is roughly explained as follows: do not transmit if the channel is bad (g small) *or* the smoothed flow is high (z large).

IV. Single flow case

A. Optimal policy

In the case of a single flow, $n = 1$, we can carry out value iteration numerically, by choosing a grid of z values, and computing the expectation numerically, after discretizing the values of g.

B. Power law approximate policy

We replace the value function (in the above optimal flow policy expression) with a simple analytic approximation to get the approximate policy

$$\hat{\varphi}(z,g) = \underset{w \geq 0}{\operatorname{argmax}} \left(\hat{V}(z^+) - \lambda \phi(w, g) \right), \tag{7}$$

where $\hat{V}$ is an approximation of the value function. Since V is increasing, concave and satisfies $V(0) = 0$, it is reasonable to fit it with a power law function as well. We let

$$\hat{V}(z) = \tilde{\beta} z^{\tilde{\alpha}},$$

with parameters $\tilde{\beta} > 0$ and $\tilde{\alpha} \in (0, 1)$ to be estimated. (For example, for discretized values of z, we find the Chebyshev fit of $\hat{V}$ to V.)

We let $\hat{J}$ be the objective value obtained using our power law approximate suboptimal policy $\hat{\varphi}$. (Recall that $J^\star$ is the optimal objective value associated with $\varphi^\star$.) To determine how well our power law policy performs, we run both controllers and compare objective values. Table I shows $\hat{J}$ and $J^\star$ for different θ (η is the percent error in objective value). Clearly $\hat{\varphi}$ is a good approximation for $\varphi^\star$.

θ	0.37	0.90	0.98
$J^\star$	0.26	0.32	0.34
$\hat{J}$	0.25	0.31	0.32
η	4%	3.2%	6.3%

TABLE I
Comparing the optimal objective value $J^\star$ with $\hat{J}$.

The analytic expression for $\hat{V}$ allows us to obtain explicit optimality conditions for (7). Substituting $\hat{V}$ from above and ϕ from §II-B into (7), we get the optimality conditions

$$w \geq 0, \quad \tilde{\beta}(1-\theta)\tilde{\alpha}(z^+)^{\tilde{\alpha}-1} - (\lambda/g)e^{(w)} = 0.$$

The first condition ensures nonnegativity. The second condition gives a unique solution only if $w > 0$. (We can find an analytical expression for w using the Lambert function [18].)

C. Numerical example

We illustrate the single-flow power law approximate policy by examining two flows: Flow 1, with light smoothing ($T = 1; \theta = 0.37$), and flow 2, with heavy smoothing ($T = 50; \theta = 0.98$). We use the utility function $U(z) = z^{1/2}$, where $\alpha = 1/2$, and $\beta = 1$. The channel gains g_t, $t = 0, 1, 2, \ldots$, are chosen to be IID from an exponential distribution with mean $\mathbf{E}\, g_t = 1$. We consider the case $\lambda = 1$.

The approximate value functions for flows 1 and 2 are $\hat{V}_1(z) = 1.7\ z^{0.60}$, and $\hat{V}_2(z) = 42.7\ z^{0.74}$. Figure 1 shows both flow policies, and as we see the policies are quite different. As expected, the heavily smoothed flow transmits less often and has a larger no-transmit region.

We also compare the average power used by both flows, for the same average utility. Choosing $\lambda = 0.29$ for flow 1 and $\lambda = 0.35$ for flow 2 gives $\bar{U} = 0.7$ for each policy. Figure 2 shows sample power trajectories (the dotted line indicates the average power required in each case). Flow 1 requires average power $\bar{P} = 0.93$, and flow 2 requires average power $\bar{P} = 0.7$. This is a 20% average power savings. (Note that a static flow policy, *i.e.*, $\theta = 0$, would be very similar to the flow 1 policy; the more lightly smoothed flow policy.)

V. A SUBOPTIMAL POLICY FOR THE MULTIPLE FLOW CASE

A. Approximate dynamic programming (ADP) policy

In this section we describe a suboptimal policy that can be used in the multiple flow case. Our proposed policy has the same form as the optimal policy, with the true value function V replaced with an approximation $V^{\rm adp}$:

$$\varphi^{\rm adp}(z, g) = \underset{w \geq 0}{\rm argmax} \left(V^{\rm adp}(z^+) - \lambda\phi(\mathbf{1}^T w, g) \right). \quad (8)$$

A policy obtained by replacing V with an approximation is called an approximate dynamic programming (ADP) policy [12].

We construct $V^{\rm adp}$ in a simple way. Let $\hat{V}_j : \mathbf{R}_+ \to \mathbf{R}$ denote the power law value function for the associated single flow problem with only the jth flow. (This can obtained numerically as described above.) We then take

$$V^{\rm adp}(z) = \hat{V}_1(z_1) + \cdots + \hat{V}_n(z_n),$$

where, $\hat{V}_j$, $j = 1, \ldots, n$, are called basis functions. Solving (8) is equivalent to solving the optimization problem

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^n \hat{V}_j(z_j^+) - \lambda\phi(F, g) \\ \text{subject to} & F = \mathbf{1}^T f, \quad f_j \geq 0, \quad j = 1, \ldots, n, \end{array} \quad (9)$$

with variables f_j, $j = 1, \ldots, n$ and F. Our ADP policy therefore is reduced to a resource allocation problem, which has several solution methods in the optimization literature. The waterfilling solution method is summarized as follows: At each time t, we solve

$$\sum_{j=1}^n \hat{V}_j(z_j^+) - \lambda\phi(F, g) + \nu(F - \mathbf{1}^T f),$$

with variables f_j, F and Lagrange multiplier ν. Maximixing over f_j and F separately and for a fixed ν we have

$$f_j = \underset{w \geq 0}{\rm argmax} \left(\hat{V}_j(z_j^+) - \nu w \right), \quad F = \log(\nu g/\lambda),$$

for $j = 1, \ldots, n$, adjusting ν, say, by bisection, until $\mathbf{1}^T f = F$. We note that for fixed ν, solving these equations above is similar to solving (7).

B. Prescient upper bound

In single flow case we measured policy performance by comparing $\hat{J}$ to $J^\star$. We cannot easily compute $J^\star$ in the multi-flow case so we compute instead a prescient upper bound on $J^\star$. To obtain our prescient upper bound, we relax the causality requirement imposed earlier on the flow policy in (4) and assume complete knowledge of the channel gains for all t. For each realization of channel gains, we must solve the optimization problem

$$\begin{array}{ll} \text{maximize} & \sum_{\tau=0}^{N-1} \mathbf{E} \left(\sum_{j=1}^n U_j(s_\tau)_j - \lambda\phi(\mathbf{1}^T f_\tau, g_\tau) \right) \\ \text{subject to} & s_{\tau+1} = \Theta s_\tau + (I - \Theta) f_\tau \\ & F_\tau = \mathbf{1}^T f_\tau \quad f_\tau \geq 0, \quad \tau = 0, 1, \ldots, N-1. \end{array}$$

The optimization variables are $s_1, \ldots, s_{N-1}$, $f_0, \ldots, f_{N-1}$, and s_0, λ, U, ϕ, $g_0, \ldots, g_{N-1}$, are problem data. The optimal value of the optimization problem above is a random variable. Let $J^{\rm pre}$ denote our prescient upper bound on $J^\star$, which we obtain via Monte Carlo simulation: We take N large and solve the problem for multiple, independent realizations of channel gains. The mean optimum value is our prescient upper bound.

C. Numerical example

In this section we compare our ADP suboptimal policy to the above heuristic policy using a simple numerical example with $n = 2$ flows. Let $J^{\rm adp}$ be the objective value we find using our ADP suboptimal policy.

We construct this multiple flow example from the problem instance in IV-C, where here, flow 1 (flow with light smoothing) and flow 2 (flow with heavy smoothing) share the single link. We keep the same assumptions on the g_t, *i.e.*, IID, exponentially distributed, $\mathbf{E}\, g_t = 1$. We use the utility $U_{\alpha\beta} = z^{1/2}$ and we set $\lambda = 1$.

For this example we generated 1000 realizations, each with sequences of length $N = 1000$. We found $J^{\rm adp} = 0.47$, and $J^{\rm pre} = 0.5$. Our ADP suboptimal policy obtains an objective value that is no more than 0.03-suboptimal.

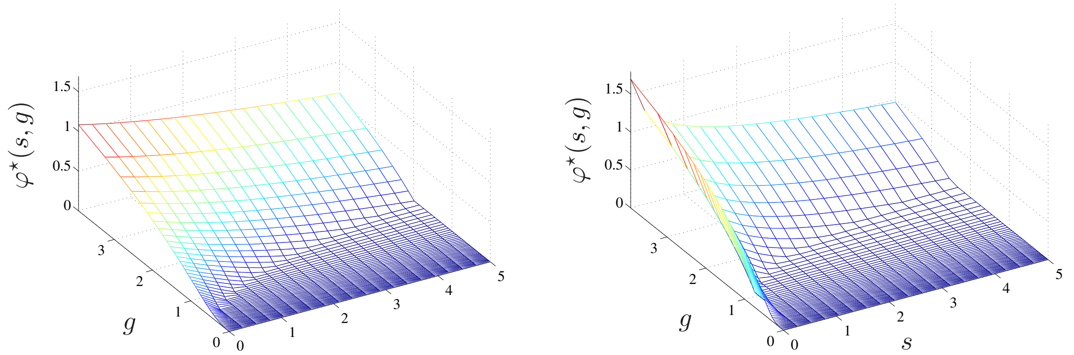

Fig. 1. Power law approximate policies for flow 1 (left) and flow 2 (right).

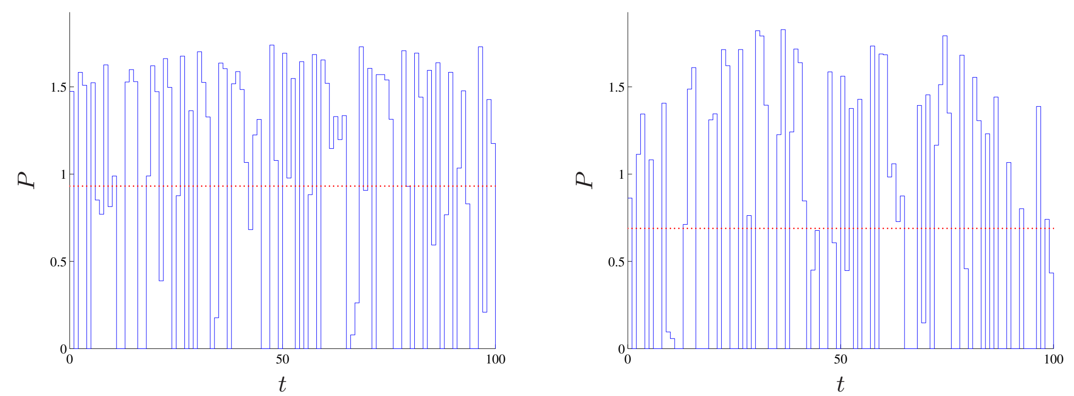

Fig. 2. Sample power trajectories for flow 1 (left) and flow 2 (right).

VI. Conclusion

We extend previous work to include new analytical power law approximate flow control policies. In the case of a single flow, we show that our power law approximate policies give near optimal performance and we show that the flow policy is dramatically different depending on the level of smoothing. We also show that smoothing results in significant power savings for applications that can wait for favorable channel conditions. In the case of multiple flows we propose a simple suboptimal ADP flow control policy. We show that this ADP policy can perform very well compared to a traditional heuristic upper bound.

References

[1] J. F. Hayes, "Adaptive feedback communications," *IEEE Transactions on Communication Technology*, vol. 16, no. 1, pp. 29–34, 1968.

[2] A. J. Goldsmith and S. G. Chua, "Variable rate variable power MQAM for fading channels," *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218–1230, 1997.

[3] X. Qiu and K. Chawla, "On the performance of adaptive modulation in cellular systems," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 884–895, 1999.

[4] F. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1997.

[5] S. H. Low and D. E. Lapsley, "Optimization flow control I: Basic algorithms and convergence," *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, pp. 861–874, Dec. 1999.

[6] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, pp. 255–312, 2007.

[7] D. O'Neill, A. J. Goldsmith, and S. Boyd, "Optimizing adaptive modulation in wireless networks via utility maximization," in *Proceedings of International Conference on Communications*, 2008.

[8] D. O'Neill, E. Akuiyibo, S. Boyd, and A. J. Goldsmith, "Optimizing adaptive modulation in wireless networks via multi-period network utility maximization," in *To appear proceedings of International Conference on Communications*, 2010.

[9] R. Bellman, *Dynamic Programming*. Courier Dover Publications, 1957.

[10] P. Schweitzer and A. Seidmann, "Generalized polynomial approximations in markovian decision processes," *Journal of Mathematical Analysis and Applications*, vol. 110, pp. 568–582, 1985.

[11] D. Bertsekas and J. Tsitsiklis, *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[12] W. Powell, *Approximate dynamic programming: solving the curses of dimensionality*. John Wiley & Sons, Inc., 2007.

[13] N. Trichakis, A. Zymnis, and S. Boyd, "Dynamic network utility maximization with delivery contracts," in *Proceedings IFAC World Congress*, 2008, pp. 2907–2912.

[14] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 2005.

[15] S. Ross, *Introduction to Stochastic Dynamic Programming: Probability and Mathematical*. Academic Press, Inc. Orlando, FL, USA, 1983.

[16] C. Derman, *Finite state Markovian deceision processes*. Academic Press, 1970.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[18] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the lambert w function," *Advances in Computational Mathematics*, vol. 5, pp. 329–359, 1996.