Cross-Layer Design with Adaptive Modulation: Delay, Rate, and Energy Tradeoffs

Daniel O'Neill* Andrea J. Goldsmith* and Stephen Boyd*
*Department of Electrical Engineering,
Stanford University
Stanford CA 94305
Email: {dconeill, goldsmith, boyd}@stanford.edu

***Abstract*— We present a crosslayer framework for optimizing the performance of wireless networks as measured by applications or upper layer protocols. The approach combines adaptive modulation with Network Utility Maximization. We extend the approach to find optimal source rates and transmitter power and rate policies without explicit knowledge of the distribution of channel states. These optimal power and rate policies balance delay (backlog), transmission rate and energy to maximize network performance under constraints on average transmitter power and link buffer arrival and departure rates. Explicit policies are found for single links, and algorithmic methods presented to find optimal policies for complex interfering networks.**

I. Introduction

Adaptive Modulation (AM) is a physical layer technique to improve the performance of wireless systems by adapting to channel conditions. Specifically, AM yields physical layer policies that adapt to changes in the condition of the wireless channel, generally under constraints on link or network resources. AM has most often been applied to optimize point-to-point physical layer metrics, and has generally not taken into account the requirements of applications or other upper layer protocols, creating a possible mismatch between the optimum behavior expected by an application or upper layer protocol and that supplied by the physical layer. To address these limitations, we proposed in [1] a crosslayer technique we called NUM/AM. NUM/AM yields policies that optimize network performance as measured by applications or upper layer network protocols. In this paper we extend NUM/AM to find optimal rate and power policies without explicit knowledge of the distribution of channel states. These optimal power and rate policies balance delay (backlog), transmission rate, and energy to maximize network performance under constraints on average transmitter power and link buffer arrival and departure rates.

AM has generated considerable research interest and commercial activity [2], [3], [4], [5], [6]. The fundamental concept is the real time adjustment of transmitter parameters, such as rate, power, BER, coding rate, etc., under flat fading or other channel variations while meeting an average power constraint. When spectral efficiency is the performance metric (SE/AM), then rate and power policies are greedy, taking advantage of good channel conditions and budgeting little or no transmitter power to poor channel conditions.

Network Utility Maximization, NUM, has been extensively studied in the context of wireline networks and is a rapidly expanding area of research in wireless networks [7], [8]. In NUM the goal is to maximize network performance as measured by metrics for an upper layer protocol or network application. The network is typically modeled as a collection of links, generally of deterministic error free capacity, that can carry one or more flows. Recent results suggest that the NUM formulation may not adequately model wireless networks that have randomly time varying wireless channels [9].

Conceptually, NUM and AM are symbiotic, with NUM modeling the upper layers of the protocol stack and AM the physical layer. NUM/AM exploits this relationship by combining the two frameworks into a crosslayer technique that captures the performance needs of applications and the random nature of the wireless channel. Performance is measured by average or expected utility, which is a broad and flexible metric. NUM/AM policies include physical layer AM policies as well as higher layer policies that maximize overall network performance.

In this paper we study NUM/AM for a network of buffered links under the assumption that the distribution of the channel states is unknown. Our approach directly estimates Lagrange multipliers and optimal policies. These policies adapt to changes in channel conditions and balance link delay, transmitter energy/power, transmission rates and the rate at which upper layer protocols inject packets into the network. We investigate networks of increasing complexity. For a single link, results include analytical expressions for optimal policies that explicitly trade-off delay rate and energy. For multiple interfering links, an algorithmic approach to computing optimal policies is presented.

The remainder of this paper is organized as follows: Section II describes the system model and a general class of utility functions. Section III describes the fundamental NUM/AM problem and presents a distribution free method of solution. Section IV investigates the single link case and presents optimal policies. Section V considers multiple interfering links and describes a numerical method for finding the optimal NUM/AM strategies. Conclusions and future work are presented in Section VI.

II. System Model

There are L links and N data sources in the network. A single link is modeled in Figure 1. Packets are injected into the link buffer by the upper layer protocol stack and are removed and transmitted by the wireless link. The channel is modeled by a channel state (gain) matrix $G \in \mathbf{R}^{L\times L}$, where G_{ij} is the power gain from the transmitter on link j to the receiver on link i. The vector of transmitter powers is given by $S \in \mathbf{R}^L$. For concreteness the link rate function is assumed to be of the form

$$R_l(S,G) = \log\left(1+\frac{KG_{ii}S_l}{\sum_{j\neq l} G_{lj}S_j+N}\right) \quad l = 1,\ldots,L \qquad (1)$$

where $K = -\log(BER)$ scales the received power to meet an instantaneous BER ceiling and N is receiver noise. Each transmitter has an average power budget $\bar{S}$. Time is discrete. The distribution of $G \sim p(G)$ is iid in every time period and is unknown to the network. We assume the channel state is estimated without error and is known at the set of transmitters. The system can adapt to changing channel conditions by estimating G and adapting parameters such as transmit power $S = S(G)$, transmission rate $R = R(S(G),G)$, the upper layer source rate $r = r(G)$ $r \in \mathbf{R}^N$, etc. We consider the case of continuous link rate and source rate adaptation subject to instantaneous BER constraints.

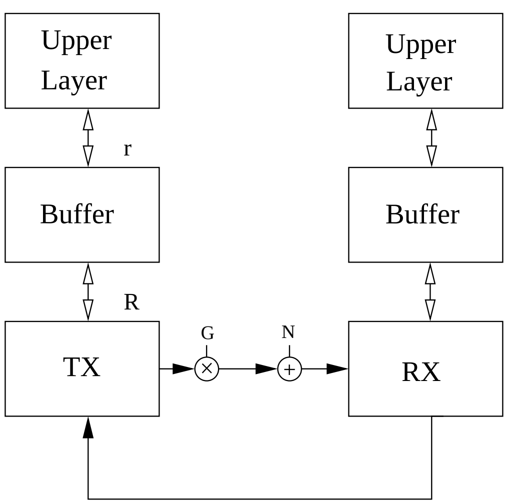

Fig. 1: System Model

A. Network Utility Maximization

The canonical NUM problem is to find the optimal source rate r that maximizes overall network utility of a network of links. The links in most prior work are assumed to have fixed, error free capacities, $\bar{R}$. In a fading context, this implies transmitter power is unconstrained. Each link is generally assumed to have an associated buffer. Formally the NUM problem can be expressed as

$$\begin{array}{ll} \underset{r\geq 0}{\text{maximize}} & \sum_n U_n(r_n) \\ \text{s.t.} & Ar \leq \bar{R} \end{array} \qquad (2)$$

where $A \in \mathbf{R}^{L\times N}$ describes the fixed typology of the network. The operation of the network is described as an iterative optimization algorithm seeking to solve this problem. The iteration index is often interpreted as time and need not correspond to physical time.

Utility functions are used as the metric of network performance [10]. Utility functions can model network protocols, applications, or user preferences. TCP in particular [11] has been modeled in this way. Each flow in a network is associated with a utility function $U(r)$. Each $U(r)$ is assumed to be continuously differentiable, nondecreasing, and strictly concave.

In this paper we consider the following general class of utility functions often used in the literature:

$$U(r) = \begin{cases} \frac{r^{1-\alpha}}{1-\alpha} & \alpha > 0 \quad \alpha \neq 1 \\ \ln(r) & \alpha = 1 \end{cases} \qquad (3)$$

The parameter α corresponds to different properties of the utility function.

III. NUM/AM

NUM/AM is a crosslayer technique that combines NUM and AM. The NUM formulation (2) is extended by formally introducing random channel (or other network component) variations and modifying the performance metrics and constraints to be averages. In this paper we consider the goal is to find adaptive rate and power policies that maximize the average utility of the network, under constraints on source and link rates and average power transmitted. By policies we mean rate and power functions that optimally adapt to changes in channel state, and we write $S(G_t)$, $r(G_t)$, $R(S(G_t),G_t)$ for the respective transmitter power, source rate and link rate policies. Formally this can be stated as

$$\begin{array}{ll} \underset{r(G_t),S(G_t)}{\text{maximize}} & \lim \frac{1}{T}\int_T \sum_l U_l(r_l(G_t))dt \\ \text{s.t.} & \lim \frac{1}{T}\int_T Ar(G_t)dt \leq \underline{\lim}\, \frac{1}{T}\int_T R(S(G_t),G_t)dt \\ & \lim \frac{1}{T}\int_T S_l(G_t)dt = \bar{S_l} \end{array} \qquad (4)$$

where A is an incidence matrix routing source flows r across links. The objective function is the time average of the instantaneous utility of the network. The utility is assumed to be a function of the source rate $r(G_t)$, the rate at which the upper layers of the protocol stack inject packets or bits into the network. The first constraint is a buffer constraint, which requires that the average arrival rate to the buffer $r(G_t)$ must be less than the departure rate from that buffer $R(S(G_t),G_t)$. The second constraint requires that averaged transmitter power $S(G_t)$ cannot exceed a maximum.

Under conditions of stationarity and ergodicity we can re-express (4) as the following:

$$\begin{array}{lll} \underset{r(G),S(G)\geq 0}{\text{maximize}} & \mathbf{E}[\sum_l U_l(r_l(G))] & \\ \text{subject to} & & \\ & \mathbf{E}[S_l(G)] = \bar{S_l} & l = 1,\ldots,L \\ & \mathbf{E}[Ar] \leq \mathbf{E}[R(S(G),G)] & \end{array} \qquad (5)$$

For simplicity in this paper we assume each source traverses exactly one link, $A = I$. More complex topologies are analyzed similarly. The optimization is over the policies $r(G)$ and $S(G)$ and indirectly the link rate $R(S(G),G)$.

If we define the optimal source and link rate policies by $r^*(G)$ and $R^*(S(G),G)$, then by Jensen's inequality

$$U(\mathbf{E}[r^*(G)]) \geq \mathbf{E}[U(r^*(G))], \tag{6}$$

so the optimal source rate policy is a constant equal to $\mathbf{E}[R^*(S(G),G)]$. This is not surprising since the rate constraint effectively couples the distributions of $r(G)$ and $R(S(G),G)$ only through their first moment. We conclude that the optimal NUM/AM source rate policy is a constant rate. In the single link case when U is strictly increasing and concave, this rate is equal to the SE/AM rate.

A. Method of Solution

FROEC, Full Recourse Optimization with Expected Constraints, is used to solve (5). FROEC is an online discrete time approach to optimization. It takes as its input the sequence of channel states seen by the network and produces as its output estimates of the optimal Lagrange multipliers and optimal policy values. The time index is k, and we indicate the estimates of the optimal Lagrange multiplier λ^* by λ^k. Policy values are denoted by $r^k = r(G^k,\lambda^k)$, $S^k = S(G^k,\lambda^k)$, and $R^k = R((G^k,\lambda^k),G^k)$, e.g. S^k is the value of the power policy at channel state G^k and λ^k. Optimal policies can sometimes be expressed analytically. More generally FROEC numerically calculates the values of the optimal policies. FROEC does not assume knowledge of $p(G)$ and under suitable conditions adjusts to changes in the channels empirical distribution.

FROEC solves the dual problem to (5). The dual function is first defined as

$$g(\lambda) = \underset{r(G)\geq 0, S(G)\geq 0}{\operatorname{argmax}} L(r(G),S(G),\lambda) \tag{7}$$

where

$$\begin{aligned} L(r(G),S(G),\lambda) \quad = \quad & \mathbf{E}[U(r(G)) \\ & -\lambda_q(r(G) - R(S(G),G)) \\ & -\lambda_s(S(G) - \bar{S}))] \end{aligned} \tag{8}$$

and $\lambda = [\lambda_q^T, \lambda_s^T]^T$ is the vector of Lagrange multipliers.

The dual problem

$$\underset{\lambda \geq 0}{\text{minimize}} \quad g(\lambda). \tag{9}$$

The FROEC approach generates a sequence of stochastic subgradients to $g(\lambda)$. These in turn are used to optimize (9). The FROEC algorithm has three steps. In the first step, the channel is estimated at time k and policy values calculated:

$$[r^k,S^k] = \underset{r\geq 0, S\geq 0}{\operatorname{argmax}} \left[U(r) - \lambda_q^k(r - R(S,G^k)) - \lambda_s^k(S - \bar{S})) \right]. \tag{10}$$

The second step calculates stochastic subgradients:

$$\delta g = - \left[\begin{array}{c} (r^k - R^k) \\ (S^k - \bar{S}) \end{array} \right] \tag{11}$$

which is a vector composed of the "slack" in the constraints evaluated at the current policy estimates. In the third step, the λ^k are updated using the subgradient recursion

$$\lambda^{k+1} = \left[\lambda^k - \Delta_k \delta g \right]^+ \tag{12}$$

where $[]^+$ is the positivity operator and the step size Δ_k is a sequence of positive constants.

B. Convergence

The convergence properties of (12) depends on the sequence $\{\Delta_k\}$. When $\Delta_k = \Delta$, the estimated Lagrange multiplier probabilistically converges to a region centered around the optimal value [12]. If we define $e(k) = ||\lambda^k - \lambda^*||$ then

$$P[e(k) \geq \varepsilon | \lambda^0] \leq A_1(\Delta) + A_2(\lambda^0) \exp(-h(\Delta)k) \tag{13}$$

where λ^0 is the initial guess for λ, and $A_1 \to 0$, $h(\Delta) \to 0$, as $\Delta \downarrow 0$. In steady state, the fixed step size approach only approximately meets the constraints, but the approximation can be made very tight for small enough Δ.

IV. NUM/AM Single Link Case

In this section we consider a single link in order to gain intuition about optimal policies. The multiple interfering link case will be described in Section V. In the single link case we can solve for analytical policy estimates. The resulting policies make optimal trade-offs between link rate, transmitter energy/power, and delay. The policies are optimal in the sense that the system converges to a small region near its optimal operating point (13).

Equation (12) can be rewritten as

$$\begin{aligned} \lambda_q^{k+1} \quad &= \quad [\lambda_q^k + \Delta_k \left(r(G^k,\lambda^k) - R(S(G^k,\lambda^k),G^k) \right]^+ \\ &= \quad [\lambda_q^0 + \textstyle\sum_{l=1}^k \delta_l T^l \left(r(G^l,\lambda^l) - R(S(G^l,\lambda^l),G^l) \right)]^+ \\ &= \quad [\lambda_q^0 + \textstyle\sum_{l=1}^k \delta_l \left(A^l - D^l \right)]^+ \end{aligned} \tag{14}$$

where we interpret A^l as the packet workload injected into the buffer and D^l as the packet workload transmitted by the link, T^l as the time duration of the l^{th} time period, λ_q^0 as the initial value of the the estimated Lagrange multiplier and $\delta_l = \frac{\Delta_l}{T^l}$. When $T^l = T$ and the step size is fixed $\Delta_k = \Delta$, then λ_q^k can be interpreted as proportional to the queue length of the buffer, where λ_q^0 is the initial backlog. A packet arriving at the buffer will be delayed by the packet workload in front of it. With this fixed step size, the queue length converges to a region centered on the optimal queue length which is proportional to λ^*, the optimum Lagrange multiplier. As channel samples vary and the system responds, the queue lengths will randomly drift within this region.

Similarly (12) can be rewritten as

$$\begin{aligned} \lambda_s^{k+1} \quad &= \quad \left[\lambda_s^k + \Delta_k(S(G_k) - \bar{S}) \right]^+ \\ &= \quad \left[\lambda_s^0 + \textstyle\sum_{l=1}^k \delta_k(E(G^l) - \bar{E}) \right]^+ \end{aligned} \tag{15}$$

where $E(G^k)$ is the energy used by the transmitter at time k when the channel is in state G^k, and $\bar{E}$ is the average energy spent by the transmitter per transmission. When $\Delta_k = \Delta$, the Lagrange multiplier λ_s^k is proportional to the energy spent by the transmitter up to time k. If on average the transmitter has exceeded its energy budget λ_s^k will be large and conversely. With a fixed step size λ_s^k will converge to a region centered about the optimal value of the Lagrange multiplier λ_s^*, and will drift within this region as channel conditions vary.

The optimal policies are computed using (10). In this form the estimated Lagrange multipliers λ_q^k and λ_s^k can be interpreted as the relative cost of allowing the queue length to increase or of exceeding the link energy budget in period $k+1$, given channel samples $\{G^l\}_{l=1}^k$. As the queue length grows, the trade-off between utility and and net arrivals in the next period changes, and it takes more marginal utility to offset any increase in queue length (delay). Similarly as λ_s^k grows the cost of exceeding the average energy per transmission budget will increase, changing the cost of energy as measured in utility. Since λ^k is a random process these costs will be different at different times or for different channel realizations. However, they will (probabilistically) remain within a region centered on their optimal values.

The power policy is a function of both λ_q^k and λ_s^k and is given by

$$S(G^k,\lambda^k) = \begin{cases} \left(\frac{\lambda_q^k}{\lambda_s^k} - \frac{N}{G^k}\right) & \frac{\lambda_s^k}{\lambda_q^k}N < G^k \quad \lambda_q, \lambda_s^k > 0 \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

Equation (10) captures the trade-off between queue backlog, transmitter energy and channel state. The ratio $\frac{\lambda_q^k}{\lambda_s^k}$ measures the relative cost of queue backlog to energy spent at time k. We term this ratio energy normalized backlog and it is the estimated energy cost per bit to transmit data at time $k+1$.

The energy policy is positive only if channel conditions exceed a threshold determined by the energy normalized backlog and receiver noise. This threshold varies with time as the system samples the channel and the energy normalized backlog is updated. If the channel state were to encounter a period of deep fading below the initial threshold, the transmitter would not initially transmit data. Over time the queue backlog would grow and the average spent energy would decrease (since the link isn't transmitting), causing the power normalized backlog to grow and for the threshold to decline. Eventually the channel state would exceed the threshold and the link would begin transmitting. Through this type of mechanism $S(G^k,\lambda^k)$ balances the queue backlog, channel state and transmitter energy.

The rate policy is

$$R(G^k,\lambda^k) = \begin{cases} \log\left(1 - K + K\frac{\lambda_q^k}{\lambda_s^k}\frac{G^k}{N}\right) & \frac{\lambda_s^k}{\lambda_q^k}N < G^k \quad \lambda_s^k > 0 \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

This policy increases rates as energy normalized backlog grows or channel conditions improve, matching intuition. The threshold ensures that rates are positive.

It is informative to compare the NUM/AM and SE/AM power policies. The functional form of the two policies are similar with the fixed optimal Lagrange multiplier λ_s^* replacing the energy normalized backlog $\frac{\lambda_q^k}{\lambda_s^k}$ in the SE/AM case. Both frameworks cease transmission if channel conditions are poor enough. The difference is that NUM/AM adjusts its threshold as the energy normalized backlog varies, while the SE/AM threshold is fixed, matching intuition. Under NUM/AM the link will transmit if the queue backlog is large enough, irrespective of the current channel condition. SE/AM's threshold on the other hand is memoryless, with transmission occurring only if channel conditions are adequate.

The source rate policy $r(G^k)$ is effectively coupled to the the link rate R only through the backlog of the link buffer. From equation (7) it can be seen that, given the backlog λ_q^k, the source rate rate can be determined independently from the remainder of the system:

$$r(G^k,\lambda_q^k) = r(\lambda_q^k) = [\dot{U}]^{-1}(\lambda_q^k) \tag{18}$$

where $[\dot{U}]^{-1}$ is the inverse function of the derivative of the utility function. As the backlog varies the source rate will also vary. Thus, under NUM/AM the best source rate is independent of the channel.

A. Single Link Simulations

In this section we describe a single link simulation. The simulation is over 200 discrete time periods. We consider MQAM modulation with $\bar{S} = 0$ dB, $N = -20$ dB, and $\mathbf{E}[G] = 0$ dB with iid Rayleigh fading in each time interval. The utility function has parameter $\alpha = 0.5$. Figure 2 shows

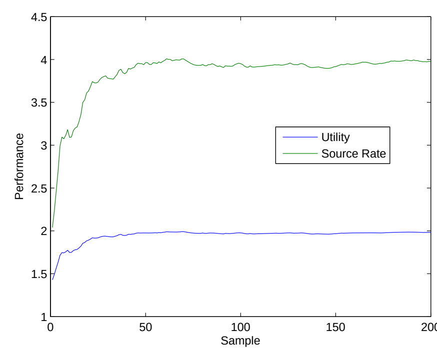

Fig. 2: Link Performance

the running average utility and source rate for the link. Running averages are used to approximate the expectation operation in (5) and to smooth the data for interpretation. The vertical axis is measured in utility and bits/sec for the utility function and source rate respectively. Both the utility and source rate improve as the number of samples increases and the algorithm seeks the optimal energy normalized backlog. The improvement in performance is not monotone, since the channel is changing randomly.

Figure 3 shows that the constraints are closely met. The moving average power curve initially deviates in the wrong direction and then quickly moves to within a few percent of $\bar{S} = 1$ as the algorithm learns the channel. The deviation is a result of the random initial value chosen for the energy normalized backlog. The source and link rate curves also initially deviate and then converge. This deviation is also a result of the initial conditions chosen for the simulation. The area between the two curves is the delay or backlog of the link.

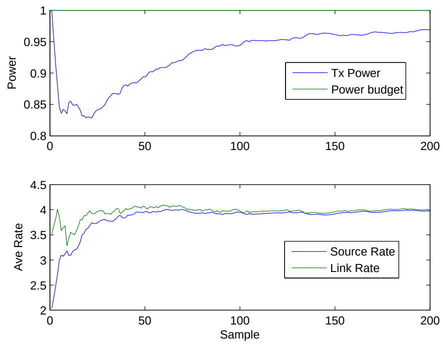

Fig. 3: Rate and Power

V. Multiple Links

In this section we consider multiple interfering links. Unfortunately (5) is not a convex problem for $L \geq 2$ and global policies may not exist. It can be made convex by assuming $SIR_l >> 1$ and transforming the variables $S_l = \exp(x_l)$, $G_{ij} = \exp(g_{ij})$, $N = \exp(n)$, where x_l g_{ij} and n are proportional to transmitter power, channel gain, and noise in dB. The link rate model can now be expressed as

$$R_i(G,S(G)) = -\ln(\mathrm{e}^{(-x_i - g_{ii})}(\textstyle\sum_{j \neq i} \mathrm{e}^{(x_j + g_{ij})} + \mathrm{e}^{n})). \quad (19)$$

The estimated Lagrange multipliers become $\lambda_q^k \in \mathbf{R}^L$ and $\lambda_s^k \in \mathbf{R}^L$ have similar interpretations to the single link case, with $[\lambda_q^k]_l$ proportional to the backlog of the l^{th} queue at the k^{th} channel sample and λ_q^k interpreted similarly. The ratios $\frac{[\lambda_q^k]_i}{[\lambda_s^j]_j}$ can be interpreted as the backlog of the i^{th} queue normalized by the j^{th} transmitter energy, we term this the cross-link energy normalized backlog.

The transmitter power optimization has necessary condition

$$\lambda_q^T \mathbf{D}[R(S(G),G)] = \lambda_s^T \quad (20)$$

where $\mathbf{D}$ is the Jacobian operator. Equation (20) relates the set of link delays to the set of transmitter powers through a non-linear operator. It is difficult to solve for $S(G)$ analytically, but it can be readily solved numerically. The FROEC approach yields a sequence of values that are the optimal policies at the the current channel state G.

The optimal source rate policy is identical to the single link case, since inter-link interference only effects the link capacities:

$$[\dot{U}_i]^{-1}(\lambda_q^k) = r_i^k(G^k, \lambda_q^k) = r_i^k(\lambda_q^k) \quad (21)$$

where $[\dot{U}]^{-1}$ is the inverse function of the derivative of the utility function.

A. Numerical Example

Figures 4 and 5 depict a numerical example with $L = N = 6$ links and sources and $\overline{S_l} = 1$. The simulation is over 200 time periods, and the channel state matrix is drawn iid Rayleigh, with diagonal elements scaled to yield an average of 20 dB over all links Figure 4 shows the performance of each link in bits/sec . As in the single link case the curves are running averages. As the network samples the channel the performance improves significantly for all links.

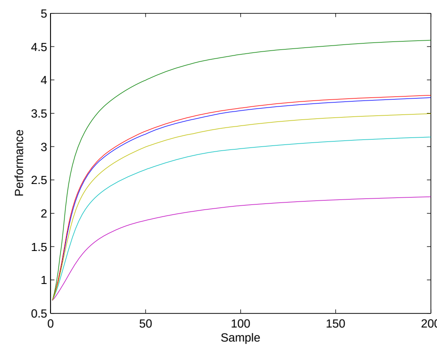

Fig. 4: Network Performance

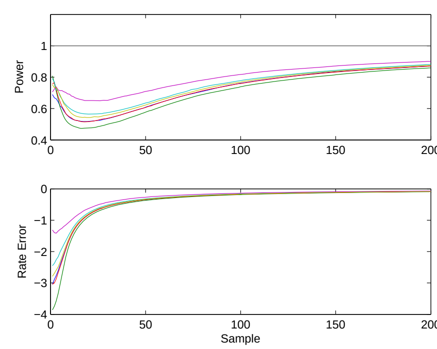

Fig. 5: Network Constraint Error

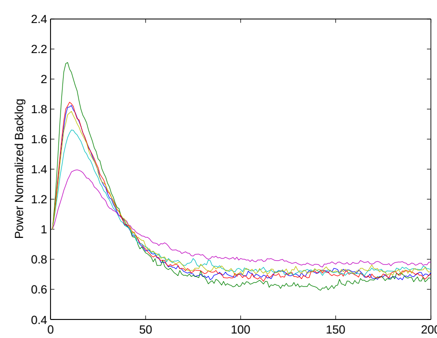

Fig. 6: Network Power Normalized Backlog

Figure 5 shows the transmitter power used by each link. It initially drops and then asymptotically converges to within a few percent of its target expected value. The initial drop is an artifice of the starting point or initial conditions of the network. The lower chart shows the difference between the source rates and link rates for each link, which also converge. Figure 6 shows the power normalized backlog for each link. The large initial peak in this curve is also caused by the initial conditions used.

Figure 6 shows the energy normalized backlog for each link. The large initial peak in this curve is caused by the initial conditions used. The curves settle in to a narrow region about their optimal values, but do not converge, reflecting the random variation of the channel.

VI. Conclusion

We have developed cross-layer adaptive transmission policies to optimize network performance based on tradeoffs between data rate, delay, and energy. The policies determine the optimal source rate, transmit power and transmission rate based on performance metrics associated with the application layer or network protocol. We also introduce the concept of energy-normalized backlog, which corresponds to the cost of transmitting packets in the next timeslot. For a single link, our optimal adaptive policies are expressed in closed form, while for multiple interfering links the policies are described using a numerical algorithm. Simulations illustrate an initial transient response, followed by convergence to a steady state region. These numerical results provide significant insight into transmission adaptation based on network performance, and also show how optimization relative to link layer metrics alone can lead to highly suboptimal policies. Future areas of research include extending our approach to correlated channels, investigating more general classes of policies using past channel information, and investigating adaptive policies based on link and network reliability.

References

[1] D. ONeill, A. J. Goldsmith, and S. Boyd, “Optimizing adaptive modulation in wireless networks via utility maximization,” *IEEE ICC 2008*, 2008.

[2] R. Hoefel and C. de Almeida, “Performance of ieee 802.11-based networks with link level adaptive techniques,” *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 2.

[3] K.-B. Song, A. S. T. C. Ekbal, and J. Cioffi, “Adaptive modulation and coding (amc) for bit-interleaved coded ofdm (bic-ofdm),” *IEEE Communications, 2004 IEEE International Conference on*, vol. 6, no. 11, 2004.

[4] X. Qiu and K. Chawla, “On the performance of adaptive modulation in cellular systems,” *IEEE Trans. Comm.*, pp. 884–895, June 1999.

[5] A. Goldsmith, “The capacity of downlink fading channels with variable rate and power,” *Vehicular Technology, IEEE Transactions on*, vol. 46, no. 3, 1997.

[6] S. Chung and A. J. Goldsmith, “Degrees of freedom in adaptive modulation: A unified approach,” *IEEE Trans. Commun.*, vol. 49, 2001.

[7] J.-W. Lee, M. Chiang, and A. R. Calderbank, “Price-based distributed algorithms for rate-reliability tradeoff in network utility maximization,” *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 5, 2006.

[8] A. R. C. M. Chiang, S. H. Low and J. C. Doyle, “’layering as optimization decomposition: A mathematical theory of network architectures’,” *Proceedings of the IEEE*, vol. 95, pp. 255–312, January 2007.

[9] J. C. Ramming, “Control-based mobile ad-hoc networking (cbmanet) program,” unpublished.

[10] R. Srikant, Ed., *The Mathematics of Internet Congestion Control*. Boston: Birkhauser, 2003.

[11] S. Low, L. Peterson, and L. Wang, “Understanding vegas: A duality model,” *Journal of ACM*, vol. 49(2), pp. 207–235, Mar. 2002.

[12] V. S. Borkar and S. P. Meyn, “The o.d.e. method for convergence of stochastic approximation and reinforcement learning,” *SIAM Journal of Control Optimization*, vol. 38, no. 2, 2000.