STOCHASTIC LEARNING ALGORITHMS FOR ADAPTIVE MODULATION

Anup Misra, Vikram Krishnamurthy, and Robert Schober

Department of Electrical and Computer Engineering
University of British Columbia
Vancouver, Canada

ABSTRACT

In this paper we present re-enforcement learning algorithms for adaptive modulation in flat fading channels for reconfigurable, agile wireless communications devices. We derive the dynamical stochastic control model, convexity properties of the stated optimization problem, learning based feedback control optimization and numerical simulations of the designed system. We show how this technique can be applied independently of channel model, error correction coding, and modulation constellation options. In addition, we demonstrate the algorithm's learning and tracking capabilities.

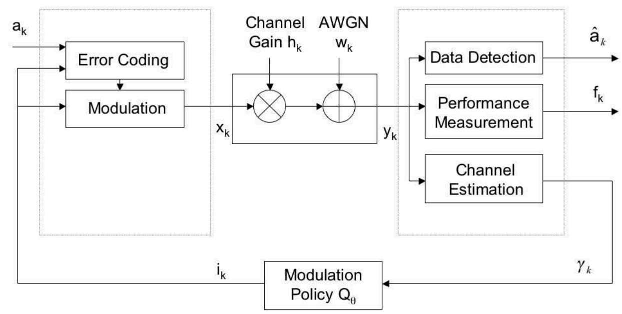

Figure 1: System block diagram

1. INTRODUCTION

In wireless communications systems employing adaptive modulation, the transmitter is able to vary its transmit power, constellation size, data rate, error coding rate and scheme in response to a flat fading channel. Adaptive modulation has been shown to greatly increase the spectral efficiency of wireless communications systems compared to fixed schemes [1] and has been actively researched as a key component in many proposed wireless standards [2].

Research into the design of adaptive modulation systems has relied primarily on analytical approaches [3] [4] where a model of the input/output characteristics of the communication system is derived as a function of the modulation policy used. In some cases functional approximations are used where the exact relationship is not convenient for analysis. It is then possible to find an optimal set of parameters that define the modulation policy to maximize the spectral efficiency or overall throughput. Techniques such as Lagrange optimization have been used to solve for the modulation policy [3]. This is a powerful approach that finds optimal solutions to a large class of adaptive modulation systems.

We propose re-enforcement learning based adaptive modulation techniques. Learning based solutions are ideal for highly adaptable devices that seek to best use the wireless channel given limited a priori knowledge of performance models. The learning based approach is motivated by the fact that, in certain cases, in may be difficult to define the input/output characteristics of the system given a broad set of communications options. Rather than using explicit models, we model the performance of the system as the output of a 'grey box'. In real-time, machine learning algorithms are used to assess the performance of the current system and iteratively tune communications parameters in order to optimize system performance.

2. STOCHASTIC DYNAMICAL MODEL FOR ADAPTIVE MODULATION

2.1. System Model for Adaptive Modulation with Packetized Transmission

Consider the single-transmit single-receive antenna communication system with adaptive modulation shown in Figure 1.

At the transmitter, we consider packetized transmission where the user data vector a_k is encoded and modulated using the modulation and coding scheme i_k to produce waveform x_k at time k. We assume that a_k is independent and identically distributed Bernoulli random vector taking values in $\mathcal{A} = \{0,1\}^L$ where L is the number of information bits per packet and p represents the probability of a individual component taking value 1. We also define $i_k \in \mathcal{I}$, where $\mathcal{I}$ is a discrete set of transmission modes. The resulting signal, $x_k \in \mathcal{X}$, is transmitted over the channel where $\mathcal{X}$ is the set of all possible modulated waveforms.

We consider a transmission scheme with variable rates and fixed power, where forward error correction (FEC), Grey mapping and M-ary Quadrature Amplitude Modulation (M-

QAM) is employed. Each packet has a fixed number of symbols and the modulation scheme used is fixed within a packet. We assume that each packet contains cyclic redundancy check (CRC) information that will allow the receiver to ascertain the success of transmission.

The channel is modeled as a slowly varying, flat Rayleigh fading channel with channel gain $h_k \in \mathcal{H}$ and autocorrelation $R_h[\tau] = J_0(2\pi f_d T_s \tau)$ where; $f_d T_s$ is the normalized Doppler spread; J_0 is the zeroth order Bessel function of the first kind; and τ is the time offset. The received signal is impaired by zero-mean additive white Gaussian noise (AWGN) w_k with variance σ_w^2. We approximate the fading coefficient by a first order Markov chain such that $h_k = \alpha h_{k-1} + v_k$ where v_k is zero-mean Gaussian noise with variance σ_v^2 independent of $h_{k-1}$and $|\alpha| < 1$ [5]. It should be noted that the fading realizations were simulated using the modified Jakes method of [6]. We also consider a block fading model where the fading remains constant over the packet interval.

The output of the channel in the discrete time baseband formulation is given by $y_k = h_k x_k + w_k$ with $y_k \in \mathcal{Y}$.

The instantaneous signal to noise ratio (SNR), γ_k, is

$$\gamma_k = \frac{\bar{P}|h_k|^2}{\sigma_w^2}. \tag{1}$$

where $\bar{P}$ is the average transmit power.

The receiver outputs the detected data, $\hat{a}_k$, and calculates the performance function. The performance function is an empirically measured quantity that reflects the success of transmission.

To enable adaptive transmission the receiver transmits channel state information (CSI) to the transmitter in order to select a new modulation and coding scheme (MCS) via a feedback path. We assume error free channel estimation and an error free, instantaneous feedback path.

The MCS used to encode x_k, is selected by the modulation policy Q_θ, which maps the measured channel quality at the receiver to a new modulation and coding scheme, $i_k \in \mathcal{I}$, at the transmitter. Each i_k is associated with a particular rate r_k in bits per symbol.

γ is partitioned into R rate regions, each region is assigned a rate in increasing order, $r_0 < r_1 < \ldots < r_{R-1}$. The rate region thresholds are parameterized by $\theta \in \Theta$ where Θ is the set $\{\mathbb{R}_+^{R-1}$ such that $\theta_1 < \theta_2 < \ldots < \theta_{R-1}\}$.

As the system operates, the transmitter is fedback empirical packet error rate information which is used to update Q_θ and drive the system to maximum performance.

2.2. Markovian Dynamical Formulation of Adaptive Modulation Problem

With the system model in place we turn our attention to the statistical dynamics of the time evolution of the system. The overall system dynamics are given by

$$y_k = h_k x_k + w_k; \quad x_k = a_k \odot Q_\theta\Big[\gamma_{k-1}\Big] \tag{2}$$

In (2), '$\odot$' denotes the error coding and modulation of the user data. The stochastic process $\{Z_k\}$ defined by

$$z_k = \begin{pmatrix} y_k \\ h_k \\ x_k \\ a_k \end{pmatrix} \tag{3}$$

on state space $\mathcal{Z} = \{\mathcal{Y} \times \mathcal{H} \times \mathcal{X} \times \mathcal{A}\}$ is Markovian with transition kernel

$$\begin{aligned} p_\theta(z_k|z_{k-1}) = \frac{p^L}{2\pi\sigma_w\sigma_v} \exp\left(-\frac{1}{2\sigma_w^2}(y_k - h_k x_k)^2\right) \\ \cdot \exp\left(-\frac{1}{2\sigma_v^2}(h_k - \alpha h_{k-1})^2\right) \\ \cdot\, \delta\left(x_k - a_k \odot \mathbf{Q}_\theta\left[\frac{\bar{P}|h_{k-1}|^2}{\sigma_w^2}\right]\right) \end{aligned} \tag{4}$$

In addition, the Markov process $\{Z_k\}$ is positively Harris recurrent and is geometrically ergodic with invariant distribution π_θ [7]. For further discussion and proof see [8]

In order to conduct statistical inference and stochastic optimization of the average behavior of the system we require the Markov Chain, Z_k, to be geometrically ergodic and possesses a unique invariant distribution π_θ.

2.3. Constrained and Unconstrained Performance Metrics for Adaptive Modulation

The throughput f is defined as

$$f(\theta) = \sum_{i=0}^{R-1} r_i \int_{\theta_i}^{\theta_{i+1}} (1 - P_{e,r_i}(\gamma)) p(\gamma) d\gamma \tag{5}$$

where, γ is SNR with distribution $p(\gamma)$. $\theta_0 = 0$, $\theta_R = \infty$ are fixed with $\theta_i < \theta_{i+1}$ and $r_i < r_{i+1}$ for $i = 1, 2 \ldots R-1$. $P_{e,r_i}(\gamma)$ represents the packet error rate under transmission rate r_i as a function of received SNR γ.

We also consider maximization of throughput, f, subject to average packet error rate constraint h given by

$$\text{s.t. } h(\theta) = \sum_{i=0}^{R-1} \int_{\theta_i}^{\theta_{i+1}} P_{e,r_i}(\gamma) p(\gamma) d\gamma < \psi \tag{6}$$

where ψ is the target packet error rate. In the constrainted case for R possible rates we require $\theta \in \mathbb{R}^R$.

Our aim is to compute the optimal threshold vector θ^* as a solution to the stochastic optimization problem

$$\begin{array}{cc} \max_\theta & f(\theta) \\ \text{s.t.} & h(\theta) - \psi \leq 0 \end{array} \tag{7}$$

2.4. Convexity of Throughput and Error Rate Constraint

We begin by making certain assumptions on the structure of the packet error rate functions, $P_{e,r}(\gamma)$.

A1) P_e is an exponentially decreasing function of received SNR γ

A2) For a given SNR $P_{e,r_i}(\gamma) < P_{e,r_{i+1}}(\gamma)$

A3) $\frac{P'_{e,r_{i+1}}(\gamma)}{P'_{e,r_i}(\gamma)} < \frac{r_i}{r_{i+1}}$ where $P'_{e,r_i}(\gamma)$ denotes $\frac{\partial P_{e,r_i}(\gamma)}{\partial \gamma}$.

Assumption A1 has been studied by [2] as an effective means for approximating packet error rates as a function of channel SNR. In addition [3] has similar studies for bit error rate. Assumption A3 follows from assumptions A1 and A2 and is empirically valid. For example, the approximations to packet error rate curves used in HIPERLAN/2 system design exhibit this property [2]. For our purposes the specific numerical approximations to the packet error rate curves are not required, only that the assumptions outlined hold.

The functions $f(\theta)$ and $h(\theta)$ defined by (5) and (6) respectively, are convex with respect to θ. For proof refer to [8].

Let the set $\mathcal{K}$, where

$$\mathcal{K} = \{\theta^* \in \Theta, \exists \lambda^* \leq 0, \nabla_\theta f + \lambda^* \nabla_\theta h = 0, \lambda^* \nabla_\theta h = 0\} \quad (8)$$

define the Karush-Kuhn-Tucker points. Assuming f and h are twice differentiable, the point $(\theta^*, \lambda^*) \in \mathcal{K}$ is unique. Moreover, θ^* meets the second order sufficient condition $\nabla^2_\theta f(\theta^*) + \lambda^* \nabla^2_\theta h(\theta^*) = 0$ [9].

3. PRIMAL DUAL STOCHASTIC GRADIENT APPROXIMATION FOR ADAPTIVE MODULATION

As the system operates the performance is empirically measured and fed back to the transmitter. Using this information, stochastic estimates of the gradient of the cost function may be computed based on sample path information. In this section we frame the problem as an iterative stochastic optimization and present stochastic approximation techniques for sample path gradient estimation.

Let $\hat{\theta}_n$, $\hat{\lambda}_n$, $\hat{\nabla} f_n$ and $\hat{\nabla} h_n$ represent estimates of θ, λ, ∇f and ∇h available at iteration n. The primal dual stochastic gradient algorithm is

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \epsilon_n [\hat{\nabla}_\theta f_n(\hat{\theta}_n) + \hat{\lambda}_n \hat{\nabla}_\theta h_n(\hat{\theta}_n)] \quad (9)$$

$$\hat{\lambda}_{n+1} = \min\{\hat{\lambda}_n - \epsilon_n (h_n(\hat{\theta}_n) - \psi), 0\} \quad (10)$$

where ϵ_n is a decreasing step size sequence of the form ϵ/n, where ϵ is a constant greater than zero. For tracking purposes we may also select a constant step size.

3.1. Gradient Estimation and SPSA Algorithm

Simultaneous Perturbation Stochastic Approximation [10] (SPSA) is a technique for gradient approximation. In SPSA an approximation of $\nabla_\theta f(\theta)$ can be obtained using only two measurements of the cost function. The SPSA approach has all components of the vector θ randomly perturbed simultaneously with a random vector Δ. If $\hat{\nabla}_\theta f_n$ and $\hat{\nabla}_\theta h_n$ are given by SPSA estimates, then

$$(\hat{\theta}_n, \hat{\lambda}_n) \rightarrow (\theta^*, \lambda^*) \text{ with probability 1} \quad (11)$$

The constants used in the application of the SPSA algorithm have been selected in accordance with the conditions of [11]. We assume that $f(\theta)$ is thrice differentiable for all values of θ. Further details about sufficient conditions and proof of convergence can be found in [10].

Apart from finite difference methods there are 3 other methods for gradient estimation; pathwise (IPA), score function and weak derivatives. For a Markovian process, score function methods have large variance and are not suitable. The weak derivative approach is a topic of future research.

4. SIMULATION RESULTS

In this section, we present Monte Carlo simulations that were conducted to test the effectiveness of the stochastic learning algorithms in the presence of convolutional error correction coding. In addition, we investigate the tracking capabilities of the stochastic optimization. Further analysis of uncoded transmission performance and turbo coded systems is available in [8].

4.1. Convolutional Coding

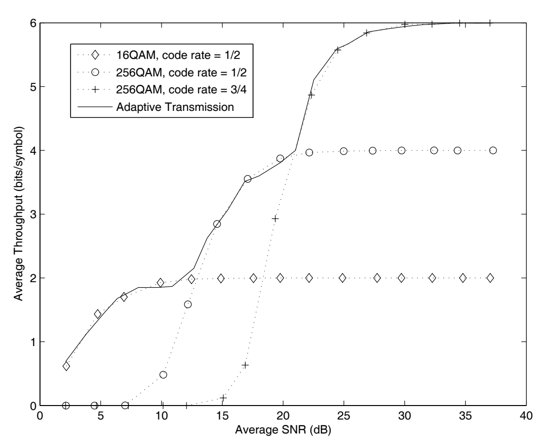

Figure 2: Average throughput versus SNR for fixed and adaptive modulation policies with convolutional coding in Rayleigh fading

We begin by considering MQAM transmission with convolutional error correction coding of constraint length 7, generator polynomial $[133, 171]_8$. Figure 2 shows the average throughput performance versus received SNR for unconstrained

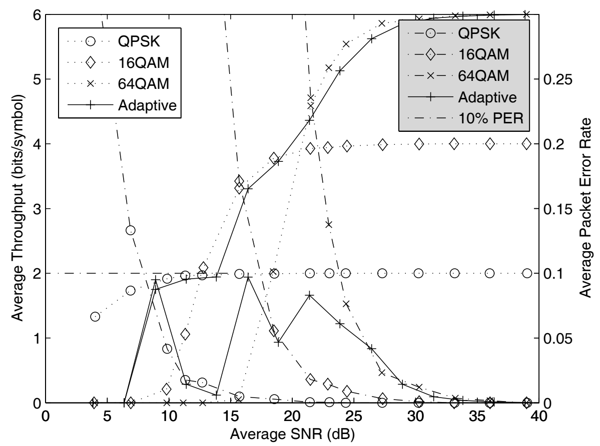

Figure 3: Average throughput versus iteration number of adaptive modulation learning process under convolutional coding with 10dB average received SNR

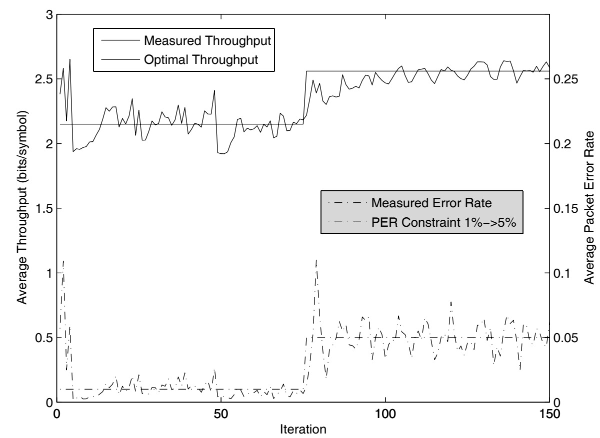

Figure 4: Tracking of adaptive modulation system with convolutional coding under 16dB average SNR where PER constraint is adjusted at iteration 75 form 1% to 5%

optimization. We see the performance of fixed rate transmission schemes (dashed lines) as well as the performance of the adaptive modulation system (solid line). There is strong adherence to the optimal transmission mode in each rate region. Figure 3 shows the average throughput performance where the PER is constrained to be 10 %. Figure 4 shows the tracking capabilities of the learning algorithm as the PER constraint is relaxed. Each iteration corresponds to transmission of 200 packets.

5. DISCUSSION AND CONCLUSION

In this paper we have designed and tested stochastic learning algorithms for the design of adaptive modulation systems. We have presented a detailed system architecture and dynamics where adaptive learning algorithms can be applied. We have found the application of stochastic learning techniques can be used to design flexible and agile adaptive modulation protocols. For future development more sophisticated gradient estimation techniques such as measure valued differentiation can be used to compute consistent estimates of the derivative.

6. REFERENCES

[1] A. J. Goldsmith and S. G. Chua, “Variable-rate variable-power MQAM for fading channels,” *IEEE Trans. Wireless Communications*, vol. 45, no. 10, Oct. 1997.

[2] Q. Liu, S. Zhou, and G. B. Giannakis, “Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links,” *IEEE Trans. Wireless Communications*, vol. 3, Sept. 2004.

[3] S. T. Chung and A. J. Goldsmith, “Degrees of freedom in adaptive modulation: A unified view,” *IEEE Trans. Comms.*, vol. 49, no. 9, Sept. 2001.

[4] A. J. Goldsmith and S. G. Chua, “Adaptive coded modulation for fading channels,” *IEEE Trans. Comms.*, vol. 46, no. 5, May 1998.

[5] K.E. Baddour and N.C. Beaulieu, “Autoregresive models for fading channel simulation,” *Proc. IEEE Globecom*, vol. 2, 2001.

[6] D. Young and N. C. Beaulieu, “The generation of correlated rayleigh random variates by inverse discrete fourier transform,” *IEEE Transactions on Communications*, vol. 48, July 2000.

[7] S.P. Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, 1993.

[8] A. Misra, V. Krishnamurthy, and R. Schober, “Stochastic optimization algorithms for adaptive modulation in software defined radio,” *IEEE Trans. Wireless Communications (In Preparation)*, 2006.

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

[10] J. C. Spall, *Introduction to Stochastic Search and Optimization: Estimation, Simulation and Control*, Wiley-Interscience, 2003.

[11] J.C. Spall, “Adaptive stochastic approximation by the simultaneous perturbation method,” *Automatic Control, IEEE Transactions on*, vol. 45, no. 10, Oct. 2000.