Statistical Decision Making in Adaptive Modulation and Coding for 3G Wireless Systems

James Yang, Amir K. Khandani, *Member, IEEE*, and Noel Tin, *Member, IEEE*

***Abstract*—In this paper, we address the application of adaptive modulation and coding (AMC) for 3rd-generation (3G) wireless systems. We propose a new method for selecting the appropriate modulation and coding schemes (MCS) according to the estimated channel condition. In this method, we take a statistical decision making approach to maximize the average throughput while maintaining an acceptable frame error rate (FER). We use a first order finite state Markov model to approximate the time variations of the average channel signal to noise ratio (SNR) in subsequent frames. The MCS is selected in each state of this Markov model (among the choices proposed in the 3G standards proposals) to maximize the statistical average of the throughput in that state. Using this decision making approach, we also propose a simplified Markov model with fewer parameters, which is suitable in systems where changes in the fading characteristics need to be accounted for in an adaptive fashion. Numerical results are presented showing that both of our models substantially outperform the conventional techniques that use a memoryless threshold based decision making.**

***Index Terms*—Adaptive modulation and coding (AMC), first-order finite-state Markov model, 3rd-generation (3G) code division multiple access (CDMA), lognormal shadowing, spectral efficiency, turbo coding.**

I. Introduction

THE USE of adaptive modulation and coding (AMC) is one of the key enabling techniques in the standards for 3rd-generation (3G) wireless systems that have been developed to achieve high spectral efficiency on fading channels [1]–[5]. The core idea of AMC is to dynamically change the modulation and coding schemes (MCS) in subsequent frames with the objective of adapting the overall spectral efficiency to the channel condition. The decision about selecting the appropriate MCS is performed at the receiver side according to the observed channel condition, with the information fed back to the transmitter in each frame. Adaptation schemes may aim to adapt to the variations in the channel quality due to multipath fading (fast fading), or variations in the average signal to noise ratio (slow fading). An excellent review of various adaptation methods used in practice is given in [6]. In the following, we provide a brief description of some of the AMC techniques reported in the literature that are more relevant to the current article. Readers are referred to [6] and [7] for a more detailed list of references on this topic.

In [8] and [9], various rate and power adaptation schemes are investigated. The power adaptation policy found is essentially a water filling formula in time. In [9], a variable power variable rate modulation scheme using m-ary quadrature amplitude modulation (MQAM) is proposed. The presented results show that the proposed technique provides a 5–10 dB gain over variable rate fixed power modulation using channel inversion and truncated channel inversion techniques (where the received power is maintained constant), and up to 20 dB gain over the nonadaptive modulation.

In [10], the channel capacity of various adaptive transmission techniques is examined. The performance of these techniques employed with space diversity is also investigated. It is shown that the spectral efficiency for a fading channel can be improved by adaptive transmission techniques in conjunction with space diversity. It is also found that when the transmission rate is varied continuously according to the channel condition, varying the transmit power at the same time has minimal impact.

In [11], the adaptation technique from [8] and [9] is modified to take into account the effect of constrained peak power. Simulation results show that with a reasonable peak power constraint, there is a small loss in spectral efficiency as compared to the unconstrained case.

In [12], an AMC scheme is proposed based on the variable power variable rate technique from [8] and [9]. This technique superimposes a trellis code on top of the uncoded modulation. Simulation results show that with a simple four state trellis code, an effective coding gain of 3 dB can be realized.

In [13], a variable rate adaptive trellis coded QAM is discussed, offering lower average bit error rate (BER) as compared to fixed rate schemes.

In [14], another AMC technique is proposed using m-ary phase shift keying (MPSK) modulation, which offers 3–20 dB gain in BER performance.

In [15], an AMC scheme which utilizes a set of trellis codes originally designed for additive white gaussian noise (AWGN) channels is proposed. This scheme is applied to a model of fully loaded microcellular network for spectral efficiency comparisons against nonadaptive coded modulation. The results obtained show that the AMC schemes provide significant advantages over a traditional nonadaptive coded modulation scheme in terms of the average spectral efficiency and decoding delay.

Accurate prediction of channel coefficients is essential to realize the gain promised by adaptive schemes. Due to this reason,

Manuscript received September 3, 2002; revised December 30, 2004. This work was supported in part by the Bell Canada University Laboratory Program and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). This paper was presented in part at the IEEE Vehicular Technology Conference, Fall 2002. The review of this paper was coordinated by Prof. J. Shea.

J. Yang and A. Khandani are with the Coding & Signal Transmission Laboratory, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, N2L 3G1 (e-mail: james@cst.uwaterloo.ca; khandani@cst.uwaterloo.ca).

N. Tin is with Bell Mobility, 5099 Creekbank Road, 5N, Mississauga, ON, Canada, L4W 5N2 (e-mail: noel.tin@bell.ca).

Digital Object Identifier 10.1109/TVT.2005.853445

channel prediction in the context of adaptive transmission has been the subject of some excellent research work (see [16] and its references). Although channel coefficients can be predicted with higher accuracy using these methods, there will still be some prediction error present and, consequently, one would benefit by including the possibility of such prediction errors in the corresponding decision in an adaptive scheme. This is the main motivation behind the current article.

Turbo codes [17], which can achieve near-capacity performance on AWGN channels, have also been proposed in adaptive transmission systems to further improve performance [18]. Results show that a gain of about 3 dB can be obtained over an AMC scheme using trellis coded modulation.

For packet data in the 3G standards, turbo codes are specified as the channel coding technique; throughout this paper, we follow the guidelines provided in one of the 3G standards proposals [1]–[4][1]. The MCSs considered include 16QAM with turbo code rate $R_c = 1/2$, 8PSK with $R_c = 1/2$, and BPSK with $R_c = 1/3$, where all of these MCSs have equal average symbol energy, E_s. Data is transmitted in successive frames. Each frame of bits has a constant duration of 5 ms, and consists of 384 coded symbols. This provides a constant data rate of 76.8 ksymbols/s regardless of the choice of MCS [1]–[4]. A common constituent code is used for the turbo code of different rates. The transfer function for the constituent code is:

$$G(D) = \left[1 \frac{n_0(D)}{d(D)} \frac{n_1(D)}{d(D)}\right]$$

where $d(D) = 1 + D^2 + D^3, n_0(D) = 1 + D + D^3$, and $n_1(D) = 1 + D + D^2 + D^3$.

A key factor determining the performance of an AMC scheme is the method used at the receiver to estimate the channel condition [in order to decide for the appropriate] MCS to be used in the next frame. The performance of turbo code in AMC systems depends heavily on the accurate prediction of the channel condition, which is usually a difficult task given the time-varying nature of the mobile environment. This is due to the fact that turbo codes operate close to the channel capacity and thus have steep performance curves. The sensitivity of turbo codes to prediction errors may cause the system to produce much less favorable results than expected. With much of the industry interest in 3G development, it is essential to overcome this shortcoming and find methods for using turbo codes in AMC systems under a more realistic environment, where prediction errors can often occur.

In the literature, many articles present AMC schemes without considering the effect of prediction errors in decision making; such is the case in [14]. In other research, authors have studied the effect of such errors on the resulting performance. This is the case with [9], where the effect of channel estimation errors is addressed for the first time. In some other articles, authors have employed more sophisticated predictors to improve the prediction accuracy. An example is [13], where the proposed scheme uses pilot symbols to estimate channel state at the receiver, and utilizes both an interpolation filter and a linear prediction filter to interpolate and predict channel conditions, respectively.

In other studies, authors have included the effect of prediction errors in the decision making. For example, in [19], the effect of fading channel variations is formally addressed, and the definition of strongly robust signaling is introduced. This is based on the idea of designing an adaptive signaling scheme that meets the BER requirements for a set of fading autocorrelation functions. This idea is applied to both uncoded modulation as well as trellis coded modulation. Results show that the proposed schemes provide a significant improvement in performance over the scheme that assumes a static channel. Reference [19] assumes that the random process of fading values follow a probability model for which the joint PDF of subsequent samples is captured in correlation values; specifically, a Raleigh or a Rician model with known Rician factor. This assumption allows the authors to capture the statistical dependency between subsequent samples in a single parameter, namely the corresponding correlation coefficient. This means the method proposed in [19] estimates a single parameter only if the underlying probability model is Raleigh or a Rician where the Rician factor is known; otherwise, it would first need to estimate the underlying probability model. Note that although we have used a correlated Gaussian probability model to generate the test data, the proposed method is not limited to this model and can be applied to any other fading process observed in practice. It should also be added that even if one follows the approach of [19] in modeling the statistical dependencies in time, one would still need to partition the space of decision variables into disjoint subsets corresponding to various modulation schemes as our proposed method does (without making any assumption on the underlying conditional probability models).

In most works [8]–[15], [18], the decision of which MCS to use for the next frame is based on the basic idea of partitioning the estimated channel signal-to-noise ratio (SNR) into regions using a set of threshold values. Each such region is associated with a particular MCS while the threshold values are optimized to maximize the overall throughput. This is what we refer to as the *memoryless threshold* method. In this paper, we take a different approach for selecting MCS with the objective of maximizing the statistical average of the channel throughput when there may exist an error in predicting the channel SNR. A simplified model with fewer parameters is also proposed, which can be used to account for the changes in the fading characteristics by updating the model parameters in an adaptive manner. Numerical results show that our method substantially outperforms the conventional memoryless threshold method. Reference [19] is the only other paper that includes the effect of channel variations in the decision making. However, since its decision making approach considers meeting the BER requirements for a set of fading autocorrelations, it only provides performance comparisons against the non adaptive scheme that assumes a static channel (it makes no comparisons with the memoryless threshold method). Therefore, unfortunately, we are not able to provide

[1]Note that this work started before the last release of the 3G standards and is based on the specifications outlined in the following proposals [1]–[4]. In the last release of the 3G standards [5], there are some slight changes with respect to the configurations considered in the current article. These changes do not affect the results reported in this article.

an appropriate comparison between our proposed method and the scheme in [19] using the results presented in [19].

The remainder of this paper is organized as follows. In Section II, we describe our system setup and channel model. In Section III, we discuss the conventional memoryless threshold method and its shortcomings. Our proposed method is presented in Section IV. Numerical results are presented in Section V, including throughput comparisons between the memoryless threshold method and our proposed method, as well as results obtained from some studies on the robustness of our proposed model. Finally, we conclude in Section VI.

II. System Setup and Channel Model

For our channel model, we consider a fading channel with time varying lognormal distributed complex gain, λ_κ where κ is the time index, and additive white Gaussian noise. This is similar to the model used in several other related papers including [9] and [12]. The lognormal complex gain represents the lognormal shadowing effect in the channel and is implemented by the following autoregressive model [20]

$$R(\tau) = e^{-v|\tau|/d} \tag{1}$$

where v is the speed of the vehicle, τ is the sampling period, and d is the effective decorrelation distance. This distance is in the order of 10–100 m as reported in [22].

Using (1), the lognormal values can be generated by low-pass filtering of a discrete white Gaussian random process. With this model, we have [20]

$$\lambda_{\kappa+1} = \xi\lambda_\kappa + (1-\xi)\theta_\kappa \tag{2}$$

where λ_κ is the mean fading level (in dB) that is experienced at time κ, ξ is a parameter that controls the correlation of the lognormal shadowing, and θ_κ is a zero mean Gaussian random variable, which is independent of λ_κ. Note that although the above probability model is used to generate the test data, the proposed method is not limited to this model and can be applied to any other fading process observed in practice.

The variance of $\theta_\kappa, \sigma_\theta^2$, is related to the variance of the lognormal shadowing, σ_λ^2, and the parameter, ξ, through [20]

$$\sigma_\lambda^2 = \frac{1-\xi}{1+\xi}\sigma_\theta^2. \tag{3}$$

By selecting appropriate values for σ_λ^2 and ξ, lognormal shadowing with any desired standard deviation and correlation can be generated. In our simulations, we have chosen values for these parameters such that the correlation between subsequent fading values follow the results reported in [22] for reasonable values of vehicle speed. Note that a different fading value is generated for each symbol of duration 13 μs (a frame of 384 symbols corresponds to 5 ms resulting in a symbol duration of approximately 13 μs).

In this paper, we follow the guidelines provided in one of the 3G standards proposals[1]–[4] in terms of modulation schemes, code rates, and frame structure as outlined in the last section. Each coded symbol in a frame has a different lognormal gain, λ_κ, generated by (2), and the channel SNR of a coded symbol is defined, in decibel scale, as

$$\gamma = \lambda_\kappa + 10\log\left(\frac{E_s}{N_o}\right) \tag{4}$$

where N_o is the one sided noise spectral density, E_s is the average symbol energy, and λ_κ is as defined in (3). The per-frame average channel SNR, which is the basis of the MCS selection criterion for the subsequent frames, is the average of the channel SNR of all the coded symbols in the frame.

It is assumed that the average channel SNR is accurately estimated at the receiver and that no delay or transmission errors can occur in the feedback channel, so any discrepancy between the predicted and the actual SNR of the next frame can only result from channel SNR prediction errors caused by the time varying nature of the channel. The effects of transmission delay and transmission errors are beyond the scope of this paper.

The performance criterion used for evaluation of the memoryless threshold method and our proposed method is the statistical average of throughput per transmitted frame. This is determined by the corresponding probability of frame error rate (FER) and the spectral efficiency of the MCS selected in the frame. The use of FER for determining throughput instead of BER is due the fact that if errors are detected in a frame after decoding, the entire frame is retransmitted and thus any correctly decoded bits in that frame should not be included in the average throughput calculation.

III. Threshold Method

Conventionally, in what we call the memoryless threshold method, the AMC system has a set $\{M_0, \ldots, M_{n-1}\}$ of n MCSs. This MCS set has a corresponding throughput versus average channel SNR, denoted by $\{T_i(\gamma), i = 0, \ldots, n-1\}$, where γ is the per-frame average channel SNR as defined earlier. These throughput values can be graphically represented, where the curves intersect with each other. The average channel SNR values corresponding to the intersection points are chosen as the threshold values, denoted by $\{\gamma_0 = -\infty, \gamma_1, \ldots, \gamma_{n-1}, \gamma_n = \infty\}$. These threshold points partition the range of SNR into n regions, denoted by $[\gamma_i, \gamma_{i+1}]$ for $i = 0, \ldots, n-1$. The kth MCS, namely M_k, is assigned to the region $[\gamma_i, \gamma_{i+1}]$ if the following condition is satisfied

$$T_k(\gamma) \geq T_j(\gamma), \quad \forall j \neq k, \quad \forall\gamma \subset [\gamma_i, \gamma_{i+1}]. \tag{5}$$

This can be interpreted as quantizing the received SNR where a specific MCS is assigned to each quantization partition to maximize the expected throughout conditioned on falling in that partition.

With this correspondence between the MCSs and the channel SNR, M_k is selected for the next frame if the average channel SNR in the current frame lies in the region $[\gamma_i, \gamma_{i+1}]$.

Since it is assumed in the memoryless threshold method that the fading is slow enough such that the average channel SNR remains in the same region from the current frame to the next, the estimated channel SNR of the current frame is simply taken as the predicted channel SNR for the next frame. This simplifying assumption, however, is often not true in a mobile environment.

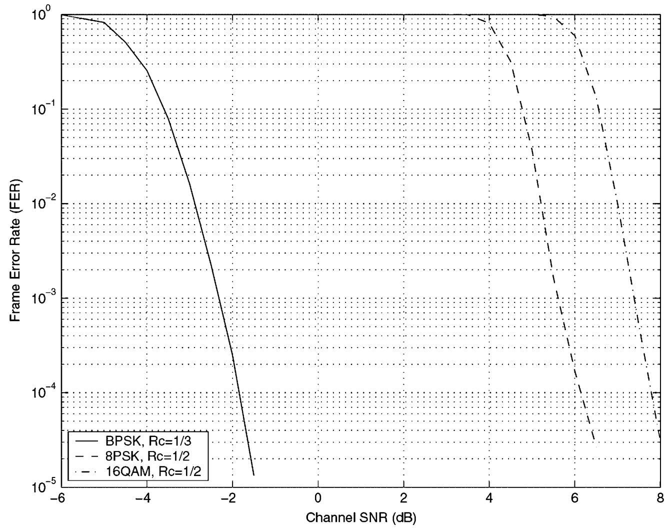

Fig. 1. FER versus SNR for turbo coded modulation schemes.

In such a case, an error in the estimation of average channel SNR can cause inappropriate selection of MCS, resulting in a degradation in FER performance.

As mentioned, turbo codes are specified as the channel coding technique for packet data in the 3G standards. One of the main characteristics of turbo codes is that they operate close to the channel capacity and the corresponding FER vs. SNR curves have a steep slope (Fig. 1). This means that even a small prediction error in the average channel SNR can result in a large degradation in FER. Therefore, it is essential to take into account the possible prediction errors when designing an AMC system where turbo codes are employed.

IV. Markov Model

Markov modeling has been successfully applied in some earlier works to capture time variations in wireless channels (see [21] as an example). In this work, noting that the exponential autocorrelation function in (1) decays very fast for practical values of its parameters motivates us to consider a first order finite state Markov model to approximate the time variations in the average channel SNR. The states in this model represent the average channel SNR of a frame (in dB) uniformly quantized (in dB scale) with a given step size Δ, and they form a set $\{S_0, \ldots, S_{m-1}\}$ of m states.

As in the memoryless threshold method, assume that there are n MCSs. We denote N_i as the number of information bits in a frame of 384 coded symbols that uses the ith MCS, namely M_i. Table I shows the values of N_i for the three MCSs used in this paper as specified in the 3G standards proposals. The turbo code interleaver acts on N_i information bits as given in Table I and the resulting coded bits are mapped via a separate interleaver to the constellation points. We also define F_{ij} as the FER of M_i in state j, and T_{ij} as the expected throughput of M_i in state j.

In the following, we propose a method for selecting the appropriate MCS based on the states of a first order Markov model, and evaluate its expected throughput. The basic strategy is to assign an MCS to each state such that the expected throughput is maximized in that state.

TABLE I
Values of N_i for the Three MCSs Used in the 3G Standards

Modulation Scheme, $M_i, i=0,1,2$	Turbo Code Rate, R_c	Number of Information Bits, N_i
16QAM	1/2	768
8PSK	1/2	576
BPSK	1/3	128

A. Full-Scale Model

We simulate a channel with lognormal shadowing according to (1)–(3) where the average SNR corresponding to each frame is uniformly quantized with a given step size Δ. We have selected appropriate values for ξ and σ_λ^2 in (1)–(3) such that the correlation between subsequent fading values follow the results reported in [22] for reasonable values of vehicle speed, e.g., between 60 km/hr and 100 km/hr (a different fading value is generated for each symbol of duration 13 μs). An appropriate offset is added to the fading values so that they result in an acceptable FER performance.

The calculation of the expected throughput for each MCS in each state of the Markov model requires the knowledge of the corresponding transitional probabilities between the Markov states. For a given number of states, m, and a given Δ, the transitional probabilities can be obtained by simulating the transmissions of a large number of frames of bits. These transitional probabilities form a set $\{P_{\alpha\beta}, 0 < \alpha, \beta < m-1\}$, where $P_{\alpha\beta}$ is the transitional probability from state α to state β.

The stationary probabilities of the states, denoted by $\{\Pi_\beta, 0\beta < m-1\}$, can be computed using the following well known system of equations [23],

$$\Pi_\beta = \sum_{\alpha=0}^{m-1} \Pi_\alpha P_{\alpha\beta}, \quad 0 \le \beta \le m-1, \quad \sum_{\beta=0}^{m-1} \Pi_\beta = 1. \tag{6}$$

The expected throughput of M_i in state j, namely T_{ij}, is therefore

$$T_{ij} = \sum_{k=0}^{m-1} N_i P_{jk}(1 - F_{ik}) \tag{7}$$

where N_i is the number of information bits in a frame of 384 coded symbols using the ith MCS, P_{jk} is the transitional probability from state j to state k, and $(1 - F_{ik})$ is the probability of correct transmission if the ith MCS is selected when the Markov chain is in state k.

For each state, we assign the MCS that has the highest expected throughput in that state according to (7) and select this MCS for the next frame if the estimated channel SNR falls in this state. In other words, M_i is assigned to S_j, if

$$T_{ij} \ge T_{kj}, \quad \forall k \ne i. \tag{8}$$

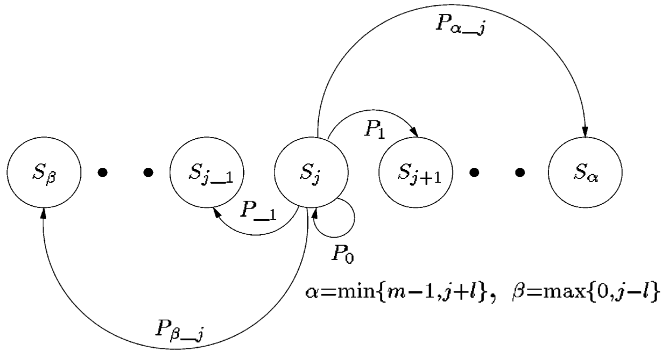

Fig. 2. Simplified Markov model.

We denote the expected throughput in state j as $\bar{T}_j$. The expected throughput averaged over all states is computed using

$$\sum_{j=0}^{m-1} \Pi_j \bar{T}_j. \tag{9}$$

B. Simplified Model

A drawback of the full scale model is that it involves many parameters (m^2 transitional probabilities) and, consequently, it is difficult to train the model on the fly to adapt to the changes in the fading characteristics (for example, caused by the variations in vehicle speed). To accommodate such an adaptation, we need a simplified Markov model with fewer parameters, which allows us to dynamically recalculate the transitional probabilities over a window of past symbols of a reasonable size.

As in the full scale model, the set $\{S_0, \ldots, S_{m-1}\}$ represents the m states in the simplified model with a step size Δ between neighboring states. We assume that the connectivities between states are as shown in Fig. 2, where

1) The maximum number of transitions from a given state is determined by $r = 2l + 1$, where $r \leq m$. The choice of the value of r provides a tradeoff between achievable throughput and the complexity of the model. An appropriate value of r is given in our numerical results.
2) $\alpha = \min\{m-1, j+l\}$ reflecting the fact that states above S_{m-1} do not exist; transitional probabilities to those states are added to the transition probability to S_{m-1}.
3) $\beta = \max\{0, j-l\}$ reflecting the fact that states below S_0 do not exist; transitional probabilities to those states are added to the transition probability to S_0.
4) The transitional probabilities are averaged over all states, and consequently are independent of the state index.

The transitional probabilities in the simplified model are specified by $\mathbf{P} = \{P_a, -l \leq a \leq l\}$. Note that the transitional probabilties exist in pairs, and this allows us to set P_a equal to P_{-a}, where $0 \leq a \leq l$ (each state is symmetrically connected to $2l+1$ states), further reducing the number of parameters in the model. We have observed (through numerical simulations) that in the full scale model these probabilities are almost equal for the majority of the states. This has motivated us to further reduce the complexit of the Markov model by seetting these probabilities equal to each other. This approximation is justified as here we are concerned with maximizing throughput. Note that although there will be cases (occuring with a small probability) in which the above approximation behaves poorly; however, as such cases occur with a small probability, the final impact on the average throughput is negligible. This observation is supported through simulation results.

The calculation of expected throughput in each state follows a relation similar to (7), slightly modified according to the structure of the simplified model shown in Fig. 2.

Since the average and approximated probabilities are now used instead of the true probabilities, this model is expected to yield smaller throughput than the full scale model. However, as will be seen in our numerical results, a very good performance can be achieved (at an appropriate step size) while substantially reducing the number of parameters in the model.

As there are fewer parameters in this simplified model (maximum of m transitional probabilities in $\mathbf{P}$), it requires a window of past symbols of a much smaller size for on-the-fly adaptation to the changes in the fading characteristics.

We call this model the *simplified* model with parameters $\{\mathbf{P}, r, \Delta\}$.

C. Robustness of Simplified Model

Setting P_a to equal to P_{-a} and/or using a smaller window size of past frames results in approximation errors in the computation of the transitional probability set, $\mathbf{P}$. Since the MCS is selected from a finite set of candidates, a small error in the transitional probability values does not necessarily result in selecting a suboptimal MCS. The robustness of the simplfied model is, therefore, determined by how much error in the transitional probability values can be tolerated before a suboptimal MCS is selected.

Assuming that P_a is set to equal to P_{-a}, then the elements of $\mathbf{P}$ is the set $\{P_a, a = 0, \ldots, l\}$. Suppose that the corresponding approximation errors are denoted by $\{\epsilon_a, a = 0, \ldots, l\}$. Then, the approximated $\mathbf{P}$ is $\{P_a + \epsilon_a, a = 0, \ldots, l\}$. Note that since $\{P_a + \epsilon_a, a = 0, \ldots, l\}$ is a probability set, the following holds

1) $\sum_a \epsilon_a = 0$,
2) $0 \leq P_a + \epsilon_a \leq 1$.

Suppose that for the jth state, S_j, the two MCSs that offer the highest expected throughput based on $\mathbf{P}$ are M_i and M_k, where $T_{ij} > T_{kj}$, meaning M_i is selected for S_j. The difference between these two expected throughputs is equal to $T_{ij} - T_{kj}$. Using $\{P_a + \epsilon_a, a = 0, \ldots, l\}$ in a relation similar to (7) (modified according to the structure of the simplified model), we can easily find the change in $T_{ij} - T_{kj}$ due to an error $\{\epsilon_a, a = 0, \ldots, l\}$ as follows

$$\delta(T_{ij} - T_{kj}) = \sum_a \epsilon_a K_a \tag{10}$$

where $K_a = N_i(1 - F_{ij+a}) - N_k(1 - F_{kj+a})$, denotes the difference between the expected throughput of M_i and M_k in state $j + a$. Therefore, we can readily identify all the $\{\epsilon_a, a = 0, \ldots, l\}$ that decreases $T_{ij} - T_{kj}$ such that it results in $T_{ij} - T_{kj} = 0$, beyond which point $T_{kj} > T_{ij}$, meaning a suboptimal MCS, M_k, will be selected instead of M_i.

D. Algorithm for Implementation of Simplified Model

The goals of using the simplified model is to take into account the changes in fading characteristics of the mobile channel in an adaptive fashion. Such a model with parameters $\{\mathbf{P}, r, \Delta\}$ can be implemented using the following algorithm:

1) FER versus SNR curves are obtained for each MCS (offline). An example of such curves is shown in Fig. 1.
2) A sufficient number of frames are passed through the channel with the average SNR recorded for each frame. As will be seen in our numerical results, using 500 frames results in a very good performance.
3) The average SNR values are uniformly quantized based on a given step size, Δ, to set up a first order finite state Markov model of m states.
4) The transitional probability set $\mathbf{P}$ of the Markov model is computed based on a given r.
   a) Set $P_a = P_{-a}$ (optional).
   b) The transitions which are not allowed are deleted, and the corresponding P_a s are modified as explained earlier.
5) The expected throughput in each state of the Markov model for each MCS is calculated using a relation similar to (7), modified according to the structure of the simplified model.
6) MCSs are assigned to each of the states in the Markov model according to (8).
7) Steps 2) through 6) are repeated for every 500 frames (corresponding to 2.5 s) for the adaptive case.

We have not addressed the issue of the training of the Markov chain in this article. Instead, we have selected the number of frames used for such training large enough to guarantee a very good estimate of the underlying transitional probabilities. Specifically, our simulations indicate that one could obtain a good approximation of the underlying transitional probabilities using far fewer than 500 frames (as used in the reported numerical results). A practical method to implement such a training algorithm would be based on using a sliding window to update the probabilities. Note that similar training problems exist in any alternative method one chooses to use. The main contributions of our work are: 1) to use a model with a small number of parameters (a model which is easier to compute as compared to other alternative techniques involving a larger number of parameters), and 2) demonstrating through computer simulations that such a simple model performs well in practice and is able to track the channel variations with required precision. Also, note that unlike [19], our proposed approach does not make any assumption on the underlying conditional probability models.

V. Results and Discussions

A. Performance of Full Scale and Simplified Models

The expected throughput per frame computed using (7) and (9) for both the memoryless threshold method and our proposed method based on the full scale model and the simplified model are shown in Fig. 3 for typical values of Δ, r, ξ (corresponding

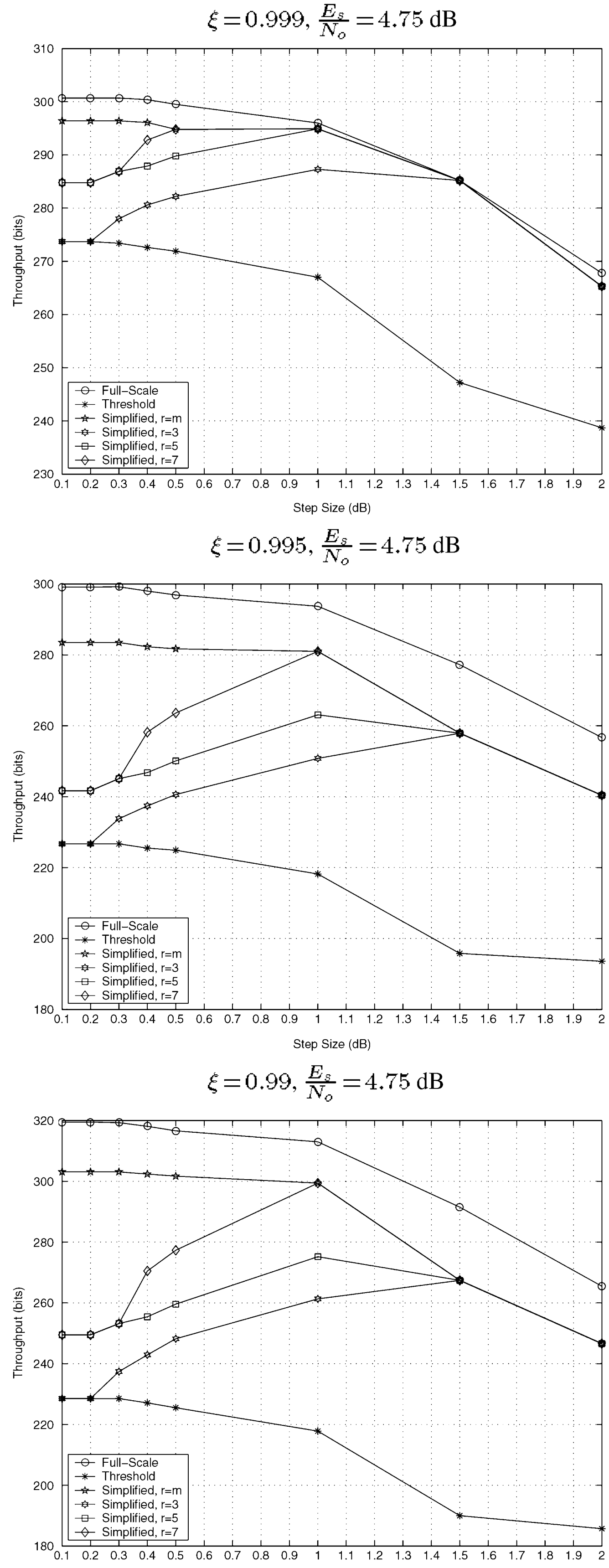

Fig. 3. Throughput versus step size for $\xi = 0.999, \xi = 0.995, \xi = 0.99$, and $(E_s/N_o) = 4.75$ dB (note that for the simplified model $P_a \neq P_{-a}$).

to different fading characteristics[2]). The value of Δ determines the number of states in the model. A typical value for the number of states is on the order of 10–20.

From Fig. 3, it can be seen that both the full scale model and the simplified model outperform the memoryless threshold method. These results, therefore, prove that our proposed method accomplishes the goal of capturing the transitional behavior of the average channel SNR that is lacking in the memoryless threshold method, and in doing so it increases the average throughput.

Referring to Fig. 3 for both the full scale model and the memoryless threshold method, the expected throughput reaches a saturation point at approximately $\Delta = 0.5$ dB, below which it stays relatively constant. It is observed that the simplified model also reaches this saturation point (corresponding to $\Delta = 0.5$ dB) when $r = m$. When $r \ll m$, the maximum expected throughput occurs at step size of 1 dB, below which the expected throughput decreases as Δ decreases due to the fact that when $r \ll m$, using a smaller Δ means a bigger portion of state transitions is ignored. In particular, when $r = 3$ and $\Delta < 0.2$ dB, the simplified model yields the same throughput as the memoryless threshold method.

It is easy to observe from Fig. 3 that by setting $\Delta = 1$ dB and $r = 7$ in the simplified model, we can achieve a throughput that is very close to the maximum value while using far fewer parameters than needed in the full scale model. Note that the value of m does not affect the implementation complexity of the model, while the value of r determines the window size of past symbols for on-the-fly adaptation.

Numerical results show that for the case of $P_a = P_{-a}, \Delta = 1$ dB and $r = 7$, using a window of 500 past frames (corresponding to 2.5 s) to recompute the transitional probabilities results in only 0.5% loss in the expected throughput as compared to using 100000 past frames for this purpose. This shows that the simplified model can be easily adapted to the changes in the fading characteristics with a reasonable delay. Also, by only requiring 500 frames, the simplified model significantly reduces the memory requirement for buffering average channel SNR values of past frames.

As an example, Table II shows the expected throughput of the three MCSs in each state of the simplified model along with the resultant MCS assignments for each state when $\xi = 0.999, (E_s/N_o) = 6.75$ dB, $r = 7$ and $\Delta = 1$ dB. It also shows a comparison between the MCS assignments made by the simplified model and the memoryless threshold method. As shown, in some states, the MCS assignments made by these two methods are the same, while in other states they are different.

B. Effects of Approximation Errors on the Robustness of the Simplified Model

In the following, we use a simple example (based on some simplifying assumptions) as an indication that the proposed scheme has some degree of robustness against possible errors in the calculation of the transitional probabilities.

[2] $\xi = 0.999$ corresponds to a fading level experienced by a vehicle traveling at approximately 60 km/hr, while $\xi = 0.99$ corresponds to a fading level experienced by a vehicle traveling at approximately 100 km/hr.

TABLE II
MCS ASSIGNMENT FOR $\xi = 0.999, (E_s/N_o) = 6.75$ dB USING SIMPLIFIED MODEL AND THRESHOLD METHOD

States, S_j	Expected Throughput, T_{ij} (bits)			MCS Assignment	
	BPSK $R_c = 1/3$	8PSK $R_c = 1/2$	16QAM $R_c = 1/2$	Simplified Model	Threshold Method
0	128	0	0	BPSK	BPSK
1	128	0	0	BPSK	BPSK
2	128	4.1	0	BPSK	BPSK
3	128	33.3	0.02	BPSK	BPSK
4	128	107.8	11.8	BPSK	BPSK
5	128	236.4	63.7	8PSK	BPSK
6	128	383.2	182.0	8PSK	8PSK
7	128	496.6	363.6	8PSK	8PSK
8	128	555.7	551.6	8PSK	16QAM
9	128	575.2	685.1	16QAM	16QAM
10	128	576	749.6	16QAM	16QAM
11	128	576	767.7	16QAM	16QAM
12	128	576	768	16QAM	16QAM
13	128	576	768	16QAM	16QAM
14	128	576	768	16QAM	16QAM

Since it is suggested in the last section that the appropriate selections for Δ and r in the *simplified* model are $\Delta = 1$ dB and $r = 7$, in the following, we examine the effect of approximation errors (due to using a smaller window size of past frames for calculating $\mathbf{P}$ and/or setting $P_a = P_{-a}$) on the overall expected throughput for this particular case.

As can be seen in Table II, the difference between the two highest calculated expected throughput in state 8 is the smallest among all the states (8PSK with $R_c = 1/2$ is selected over 16QAM with $R_c = 1/2$ for a difference of 4.1 bits, i.e., $T_{ij} - T_{kj} = 4.1$ bits), and therefore this state has the lowest error tolerance level, and thus determines the robustness of the simplified model.

If $P_a = P_{-a}$, then the corresponding errors for $\mathbf{P}$ are $\{\epsilon_a, a = 0, 1, 2, 3\}$. To simplify calculation, we assume that the magnitude of ϵ_a is constant for $a = 0, 1, 2, 3$, say $|\epsilon_a| = \epsilon_c, a = 0, 1, 2, 3$. To find the maximum value for ϵ_a (denoted by $\epsilon_{\max}$) such that the optimal MCS is still selected, we solve the following equation for ϵ_a

$$(T_{ij} - T_{kj}) + \delta(T_{ij} - T_{kj}) = 0. \tag{11}$$

To find $\epsilon_{\max}$, we compute K_as and arrange them in decreasing order. Then, noting that (10) is a linear function of $\{\epsilon_a, a = 0, 1, 2, 3\}$, we simply associate $\epsilon_{\max}$ with the two smaller values, and $-\epsilon_{\max}$ with the two larger values. This selection results in the most negative value for $\delta(T_{ij} - T_{kj}) = \sum_a \epsilon_a K_a$ while maintaining the condition that $\sum_a \epsilon_a = 0$. Then, we substitute the results in (11) and solve for $\epsilon_{\max}$.

We find that there are six cases to consider. The results are tabulated in Table III for all six cases. Note that '-' denotes the case where the particular error type cannot result in $T_{ij} - T_{kj} = 0$, and consequently does not result in selection of a suboptimal

TABLE III
MAXIMUM ERRORS ALLOWED FOR ERROR TYPES IN STATE 8

Error Types $\{\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3\}$	ϵ_{max}
$\{-\epsilon_c, -\epsilon_c, \epsilon_c, \epsilon_c\}$	0.0072
$\{-\epsilon_c, \epsilon_c, \epsilon_c, -\epsilon_c\}$	-
$\{-\epsilon_c, \epsilon_c, -\epsilon_c, \epsilon_c\}$	-
$\{\epsilon_c, \epsilon_c, -\epsilon_c, -\epsilon_c\}$	-
$\{\epsilon_c, -\epsilon_c, \epsilon_c, -\epsilon_c\}$	-
$\{\epsilon_c, -\epsilon_c, -\epsilon_c, \epsilon_c\}$	0.0046

MCS; the last row in the table represents the worst case scenario in this state for the system as described earlier.

By assuming that all errors have the same magnitude, this approach finds the worst case scenario for the errors such that (10) is reduced to zero with the least value of magnitude of ϵ_a.

VI. Conclusion

In this paper, we evaluated the performance of turbo code based adaptive modulation and coding in 3G wireless systems. We proposed a new method for selecting the appropriate modulation and coding schemes according to the estimated channel condition where we use a first order finite state Markov model to represent the average channel SNR. We take a statistical decision making approach to address the potential problems caused by the sensitivity of turbo code to the errors in predicting the channel SNR. Numerical results are presented showing that our method substantially outperforms the conventional techniques that use a memoryless threshold-based decision making approach. We also propose a simplified model with fewer parameters which is suitable in systems where changes in the fading characteristics need to be accounted for in an adaptive manner. It is shown that the simplified model has some degree of robustness against possible errors in the calculation of the transitional probabilities.

Acknowledgment

The authors would like to thank the Associate Editor, J. Shea, and the anonymous reviewers for their technical input which substantially improved the quality of this article.

References

[1] "Proposal for standardization of very high rate mixed voice-data traffic capabilities, based on extending and enhancing 1X systems," Motorola and Nokia, 3GPP2, S00-200003210-020, Mar. 2000.

[2] 3GGP2, S00-20000321-0191XTREME.Motorola and Nokia, Mar., 2000.

[3] 3GPP2, C50-20010611-008aJoint 1XTREME Proposal for 1XEV-DV.Motorola, Nokia, Philips, TI and Altera, Jun. 2001.

[4] Motorola, Nokia, Texas Instruments, Altera, and Philips Semiconductors, 3GPP2, C50-200010611-013R11XTREME Physical Specification for Integrated Data and Voice Services in cdma2000 Spread Spectrum Systems., Jun., 2001.

[5] 3GPP2 standardC.S0001-C: Introduction to CDMA2000 . . ., C.S0002-C: Physical Layer Standard . . ., C.S0003-C: Medium Access Control (MAC) . . ., C.S0004-C: Signaling Link Access Control (LAC) . . ., C.S0005-C: Upper Layer (Layer 3) Signaling . . ., C.S0006-C: Analog Signaling Standard . . .., May, 2002.

[6] S. Nanda, K. Balachandr, and S. Kumar, "Adaptation techniques in wireless packet data services," *IEEE Commun. Mag.*, vol. 38, no. 1, pp. 54–64, Jan. 2000.

[7] L. Hanzo, C. H. Wong, and M. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA, and OFDM Systems*. New York: Wiley, 2002.

[8] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.

[9] A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.

[10] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 1165–1181, Jul. 1999.

[11] W. Choi, K. Cheong, and J. Cioffi, "Adaptive modulation with limited peak power for fading channels," *Proc. IEEE*, pp. 2568–2571, Mar. 2000.

[12] A. J. Goldsmith and S. G. Chua, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 595–602, May 1998.

[13] V. Lau and M. Macleod, "Variable rate adaptive trellis coded QAM for high bandwidth efficiency applications in Rayleigh fading channels," *Proc. IEEE*, pp. 348–352, Apr. 1998.

[14] S. M. Alamouti and S. Kallel, "Adaptive trellis-coded multiple-phased-shift keying for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 2305–2314, Jun. 1994.

[15] K. J. Hole and G. E. Oien, "Spectral efficiency of adaptive coded modulation in urban microcellular networks," *IEEE Trans. Veh. Technol.*, vol. 50, pp. 205–220, Jan. 2001.

[16] A. Duel-Hallen, H. Shengquan, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 62–75, May 2000.

[17] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: turbo-codes," in *Proc. ICC'93*, Geneva, May 1993, pp. 1064–1070.

[18] S. Vishwanath and A. J. Goldsmith, "Exploring adaptive turbo coded modulation for flat fading channels," in *Proc. VTC*, vol. 4, Boston, MA, Fall 2000, pp. 1778–1783.

[19] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. Commun.*, vol. 47, pp. 844–855, Jun. 1999.

[20] G. L. Stuber, *Principles of Mobile Communication*. Norwell, MA: Kluwer, 1996.

[21] H. Wang and N. Moayeri, "Finite-state markov channel-a useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, 1995.

[22] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, pp. 2145–2146, Nov. 1991.

[23] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed., New York: McGraw-Hill, 1991.

James Yang, photograph and biography not available at time of publication.

Amir K. Khandani (M'93) received the M.A.Sc. degree from the University of Tehran, Tehran, Iran, and the Ph.D. degree from McGill University, Montreal, QC, Canada, in 1985 and 1992, respectively.

Subsequently, he was a Research Associate at INRS Telecommunications, Quebec University, Montreal, for one year. In 1993, he joined the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, where he is currently a Professor.

Dr. Khandani is currently an NSERC Industrial Research Chair (funded by NSERC and Nortel Networks) on "Advanced Telecommunications Technologies" and a Canada Research Chair (Tier I) on "Wireless Systems."

Noel Tin (M'00) received the B.Sc., B.A., and M.Sc. degrees from Queen's University, London, ON, Canada, in 1998 and 2000, respectively.

He has been with Bell Mobility, Mississauga, ON, since 2000, working as network architect, senior architect, and, most recently, as technology advisor. His research interests include wireless access technology, smart antenna and signal processing, and wireless device technology and applications.