Data-Aided Noise Reduction for Long-Range Fading Prediction in Adaptive Modulation Systems[1]

[*]Tao Jia, [*]Alexandra Duel-Hallen and [+]Hans Hallen
Department of [*]Electrical and Computer Engineering, [+]Physics
North Carolina State University
Raleigh, NC 27695, U. S. A.
Email: {tjia, sasha, hans_hallen}@ncsu.edu

Abstract

Long-range fading prediction (LRP) enables adaptive transmission methods for rapidly varying mobile radio channels. A data-aided noise reduction (DANR) method is proposed to enhance the accuracy of fading prediction and to improve the spectral efficiency of adaptive modulation systems enabled by the LRP. The DANR includes an adaptive pilot transmission mechanism, robust noise reduction (NR), and decision-directed channel estimation. Due to improved prediction accuracy and low pilot rates, the DANR results in higher spectral efficiency than previously proposed NR techniques, which rely on oversampled pilots. Since adaptive trellis-coded modulation (ATCM) is more sensitive to prediction errors than uncoded adaptive modulation, NR is necessary to maintain the coding gain for realistic signal-to-noise ratios (SNR) and prediction ranges. It is demonstrated that ATCM enabled by DANR-aided LRP provides significant coding gain for shorter prediction ranges often employed in practice. For realistic SNR, adaptive modulation aided by LRP has better performance for our realistic physical model than for the Jakes model.

[1] This research was supported by NSF grant CCR-0312294 and ARO grant W911NF-05-1-0311

I. INTRODUCTION

Adaptive modulation has been investigated extensively due to its effectiveness and simplicity [1, 2]. Moreover, the spectral efficiency of adaptive modulation schemes decreases as the accuracy of the channel state information (CSI) degrades at the transmitter [2, 3,8]. While the adaptive trellis-coded modulation (ATCM) achieves significant coding gain over the uncoded adaptive modulation for reliable CSI [4], the coding gain diminishes when the CSI accuracy degrades, and uncoded adaptive modulation outperforms ATCM when CSI is unreliable [3, 21].

To enable adaptive modulation, the CSI needs to be predicted to compensate for the feedback and data processing delay and system constraints. In this paper, we focus on the autoregressive model-based linear prediction, referred to as the long-range fading prediction (LRP) method, which was shown to achieve superior performance for realistic channel models and measured channels [5, 6]. Since performance of fading prediction degrades severely at low and medium signal-to-noise ratio (SNR), highly oversampled pilot method was proposed and utilized in adaptive modulation systems in [7, 23]. Utilization of highly oversampled pilots in noise reduction (NR) prior to prediction was proposed in [7]. These noise reduction techniques, referred to as the *high-rate pilot* methods, improve the prediction accuracy for practical SNR over the low-rate *raw pilot* prediction algorithm, which relies only on the pilot symbols necessary for noiseless fading coefficient prediction [7]. Fading prediction aided by the high-rate pilot method was utilized in adaptive modulation systems in [2, 3]. These methods were shown to improve on the *raw pilot* predictors that do not employ NR. However, the resources consumed by transmitting pilots were not considered in these investigations. In this paper, we demonstrate that the high-rate pilot method consumes excessive system bandwidth, so the overall spectral efficiency does not increase significantly despite better prediction accuracy. The rate and power of raw pilots are optimized to maximize the spectral efficiency under given power constraint in [8], the pilots still consume a large portion of the available bandwidth in that method.

We propose to exploit the data symbols in noise reduction to achieve high accuracy of prediction without significantly increasing the pilot rate and power, thereby improving the spectral efficiency [9]. While data-aided methods have been widely used for channel equalization, estimation, and tracking [10, 28]. In [5, 11], data-aided noise reduction (DANR) was employed in LRP to enable adaptive channel inversion. However, adaptive modulation presents unique challenges for decision directed noise reduction to aid the LRP. First, the adaptive modulator usually stops sending data symbols during deep fading to save energy. The decision directed algorithm is prone to divergence due to the interruption of data transmission. Second, error propagation is likely to happen since the LRP result is used to determine the data transmission.

In this paper, a new DANR method is proposed for the LRP-aided adaptive modulation systems. This method comprises an adaptive pilot transmission mechanism that compensates for outages, robust noise reduction and decision-directed channel estimation. The prediction accuracy and spectral efficiency of the proposed method are compared with those of the high-rate pilot and raw pilot methods, and the advantages of the DANR approach are demonstrated for practical SNR and prediction ranges. We employ both the standard Jakes model and our physical fading channel model in the simulations, and determine the set of SNR values where realistic distribution of reflectors aids prediction performance.

II. SYSTEM MODEL

We consider a single-carrier flat fading channel and assume the symbol rate f_s. The equivalent lowpass received lth sample is

$$y(l)=x(l)h(l)+w(l), \tag{1}$$

where $x(l)$, $h(l)$, and $w(l)$ are the transmitted symbol (data or pilots), the complex fading channel coefficient with $E[|h(l)|^2]=1$, and complex white Gaussian noise with variance N_0, respectively. We employ the Jakes model, with the autocorrelation function of the channel coefficient approximated by that of the Rayleigh fading, $E[h(l)h^*(l-\Delta l)]=J_0(2\pi f_{dm}\Delta l/f_s)$ [12], where $J_0(*)$ is

the Bessel function of the first kind, and f_{dm} is the maximum Doppler frequency. The transmitted symbols are either pilots or data symbols. Without loss of generality, we assume that $x(l)=\sqrt{E_p}$ when pilots are transmitted (*??*).

To enable adaptive modulation, the fading coefficient is predicted using a linear predictor (LP) [5]. The actual fading coefficient $h(l)$ and its prediction $\hat{h}(l)$ are jointly Gaussian distributed with the cross-correlation [13]

$$\rho \triangleq \mathrm{E}[h(l)\hat{h}^*(l)]/\sqrt{\mathrm{E}[|h(l)|^2]\sigma_{\hat{h}}^2} \tag{2}$$

where $\sigma_{\hat{h}}^2=\mathrm{E}[|\hat{h}(l)|^2]$. Without loss of generality, we omit the time index l, and denote $\gamma=|h|^2$ and $\hat{\gamma}=|\hat{h}|^2$ in this section.

In practice, the cross-correlation ρ in (2) is measured from the observed data set. For the fixed SNR and target BER, the adjustment of thresholds in the adaptive modulation system is determined by ρ. For the linear minimum MSE (MMSE) predictor, the *normalized MSE* $\mathrm{NMSE}\triangleq\mathrm{E}[|h(l)-\hat{h}(l)|^2]/\mathrm{E}[|h(l)|^2]$ is given by $1-\rho^2$ [13], and for general linear predictor, $\mathrm{NMSE}\geq 1-\rho^2$. Since ρ determines the spectral efficiency in adaptive modulation aided by fading prediction, we employ the *normalized prediction MSE* as $\mathrm{NPMSE}\triangleq 1-\rho^2$ to measure performance of predictors in all our plots.

Consider a coded or uncoded adaptive modulation system that employs MQAM constellation sizes $\{M_1, M_2, \ldots, M_N\}$, and switching thresholds $\{\gamma_0, \gamma_1, \gamma_2, \ldots, \gamma_{N+1}\}$ with $\gamma_0=0$, and $\gamma_{N+1}=\infty$. When predicted CSI $\hat{\gamma}$ falls in the interval $[\gamma_i, \gamma_{i+1})$, the constellation size M_i is used in transmission. When $0\leq\hat{\gamma}<\gamma_1$, an outage occurs. We employ constant power allocation due to its excellent performance and low required feedback rate [2, 8]. Given fixed average $\mathrm{SNR}=E_s/N_0$, where E_s is the average overall symbol energy of data and pilot symbols. In all LRP methods investigated in this paper, the pilot energy $E_p=E_s$. Thus, $E_s=E_d\int_{\gamma_1}^{\infty} p(\hat{\gamma})d\hat{\gamma}=E_d\exp\left[-\gamma_1/\sigma_{\hat{h}}^2\right]$[2, 9],

where E_d is the average MQAM symbol energy, and $p()$ is the probability density function (PDF). When the spectral efficiency loss caused by transmitting pilots is not taken into account, the spectral efficiency is evaluated as

$$S_d = \sum_{i=1}^{N} \left[q_i \int_{\gamma_i}^{\gamma_{i+1}} p(\hat{\gamma}) d\hat{\gamma} \right], \tag{3}$$

where q_i is the number of information bits associated with the modulation level M_i.

In the presence of imperfect CSI, we employ the uncoded adaptive modulation method described in [2, 7], and the ATCM design in [3] to maintain the target BER. Given the reliability of prediction ρ (2), these methods select thresholds to maximize the spectral efficiency (3) while maintaining the average BER below the target BER P_t [2, 3]. In this paper, the set of constellation sizes {4, 16, 64, 256} and $P_t=10^{-4}$ are used for simulations unless specified otherwise. For ATCM, we adopt the scheme in [18, Fig. 2] that employs a simple rate 1/2 convolutional encoder with four states [19]. Fig. 2 shows the spectral efficiency of these systems S_d for several average SNR levels when the spectral efficiency degradation due to pilots is not taken into account. For the NPMSE on the order of 10^{-2} or smaller, prediction is sufficiently reliable to maintain near-optimal bit rate, and the spectral efficiencies are not sensitive to the variation of the NPMSE. On the other hand, these SEs decrease rapidly as the NPMSE approaches 10^{-1}. We refer to this phenomenon as "crossing the NPMSE threshold". Moreover, the rate of decrease is greater for ATCM than for uncoded adaptive modulation. This observation confirms the conclusion in [3,13] that ATCM is more sensitive to prediction errors than UAM. For NPMSE$\approx 10^{-2}$, ATCM achieves about 0.5 bps/Hz coding gain, while uncoded adaptive modulation outperforms ATCM when the NPMSE approaches 10^{-1}, which is often the case for realistic SNR values in practical mobile radio scenario [5, 13]. Therefore, to achieve positive coding gain in practical systems, the NPMSE should be decreased by employing NR.

III. LONG RANGE PREDICTION WITH DATA-AIDED NOISE REDUCTION

A. Frame Structure and Adaptive Pilot Transmission

The transmitted symbols are grouped into frames of L symbols, and the constellation size is fixed in each frame. Thus, we assume that the fading coefficient is approximately constant during each frame. To satisfy this assumption, the frame size L is chosen as [20]

$$L = \frac{f_s}{100 f_{dm}} \tag{4}$$

We use vectors $\boldsymbol{x}(k)=[x(kL), x(kL-1),\ldots, x(kL-L+1)]^{\mathrm{T}}$ and $\mathbf{y}(k)=[y(kL), y(kL-1), \ldots,y(kL-L+1)]^{\mathrm{T}}$ to represent the transmitted and received symbols of the kth frame, respectively (see (1)). In addition, we denote

$$h(k;L)\approx h(kL-l),\ l\in[0,\ldots,L-1] \tag{5}$$

as the channel coefficient associated with the kth frame.

In the proposed approach, which combines pilots and decisions to predict the fading coefficient, we group K consecutive frames into one superframe. The observed signals of the mth superframe are represented by a vector of length KL, $\mathbf{y}_s(m)=[\mathbf{y}^T(mK)\ \ \mathbf{y}^T(mK-1)\ \ \ldots\ \ \mathbf{y}^T(mK-K+1)]^{\mathrm{T}}$. In each superframe, one pilot is transmitted at the last symbol position. It is denoted $x(mKL)$, and referred to as *superframe pilot*. These pilots are used to obtain the initial channel estimates for the demodulation of data symbols. Moreover, if a frame does not contain the superframe pilot and experiences an outage, we also adaptively transmit one pilot symbol at the end of that frame. These additional pilot symbols are necessary to maintain the reliability of prediction during an outage, when decision directed estimation is not possible. This method is referred to as *adaptive pilot transmission*.

Let η denote the fraction of bandwidth dedicated to pilots. The actual spectral efficiency of the system is $\overline{S}=(1-\eta)S_d$ bps/Hz, where S_d is the spectral efficiency when pilots are not taken into account (3). Thus, in DANR system, $\eta=1/KL$. In high-rate pilot systems, the pilot rate is on the

order of $100f_{dm}$, i.e., the frame rate in DANR or $\eta=1/K$ [21]. Thus, DANR reduces η by a factor of K. Note that pilots adaptively transmitted during outage frames in DANR method do not incur spectral efficiency penalty since data symbols are not transmitted in those frames.

In DANR system, we select $E_s=E_p$ as in a pilot symbol assisted modulation (PSAM) system [22]. Although optimally allocating power between data symbols and pilots would achieve higher spectral efficiency as in [8], the spectral efficiency loss of our selection relative to the optimal allocation is small due to the following reasons. First, the pilot symbols account for only a small percentage of the available bandwidth. Reducing the power allocated to pilot symbols would not significantly improve the energy of the data symbols. Second, when decision-directed data symbols are used to improve prediction accuracy, increasing the pilot power would not significantly improve the prediction accuracy provided the pilot SNR is maintained high enough to demodulate the received data symbols reliably.

The power constraint for our DANR method is given by

$$E_s=\frac{KL-1}{KL}E_d\int_{\gamma_1}^{\infty}p(\hat{\gamma})d\hat{\gamma}+\eta E_p+\frac{E_p}{KL}(K-1)\int_{0}^{\gamma_1}p(\hat{\gamma})d\hat{\gamma} \tag{6}$$

Where the first, the second, and the last term on the right correspond to the energy spent on M_i-QAM symbols ($M_i>0$), the energy spent on pilots transmitted at superframe rate, and the energy spent on pilots transmitted during outage, respectively. Using (6), we can calculate the optimal value of thresholds for adaptive modulation methods.

B. Initial Channel Estimation for Demodulation

To demodulate received data symbols, we obtain initial channel estimates as illustrated in Fig. 1. Without loss of generality, assume that the latest received frame has index (m_0K-k_0), $k_0\in[0, K-1]$. This frame belongs to the m_0th superframe. The initial estimate for frame $(mK-k)\le(m_0K-k_0)$ in superframe $m\le m_0$ is computed as

$$\text{(*any change?*) } h_0(mK-k;L)=\begin{cases}\dfrac{k}{K}h_0(m_0K-K;L)+\dfrac{K-k}{K}\hat{h}(m_0K;L), & \text{if } m=m_0 \text{ and } k_0\neq 0\\ \displaystyle\sum_{p=0}^{P_{est}-1}\omega_{est}^{*}(k,p)y(mKL-pKL), & \text{otherwise}\end{cases} \tag{7}$$

In the first case in (7), the superframe pilot $y(m_0KL)$ associated with superframe m_0 has not been received. Therefore, the initial channel estimates for frames in this superframe are obtained via linear interpolation between $h_0(m_0K-K;L)$ (the initial channel estimate for the last frame in the previous superframe) and $\hat{h}(m_0K;L)$ (the prediction of the channel coefficient for the last frame in the m_0th superframe), where the prediction method is discussed in section III-E.

In the second case in (7), the mth superframe pilot has been observed (see Fig. 1). Thus, either $m<m_0$ or $k_0=0$. In the latter case, the latest observed frame includes the superframe pilot. The estimation is made by filtering superframe pilot observations $\mathbf{y}_p=[y(mKL)\ y(mKL-KL)\ \ldots\ y(mKL-P_{est}KL+KL)]^T$, where P_{est} is the order of the estimation filter. Since channel statistics are unknown and time-variant in practice, we employ a robust filter that assumes a flat Doppler spectrum with support $[-f_{dm}, f_{dm}]$ [23]. The $P_{est}\times 1$ vector of filter coefficients is constructed as

$$[\boldsymbol{\omega}_{est}(k,0),\ \ldots,\boldsymbol{\omega}_{est}(k,P_{est}-1)]^T=(\mathbf{R}+N_0\mathbf{I})^{-1}\mathbf{r}(k), \tag{8}$$

where the $P_{est}\times P_{est}$ autocorrelation matrix $\mathbf{R}$ has elements $\mathbf{R}_{i_1,i_2}=\text{sinc}[(i_1-i_2)\hat{f}_{dm}KL/f_s]$, the $P_{est}\times 1$ cross-correlation vector $\mathbf{r}(k)$ has elements $[\mathbf{r}(k)]_i=\text{sinc}[(iK-k)\hat{f}_{dm}L/f_s]$, $\mathbf{I}$ is the $P_{est}\times P_{est}$ identity matrix, and $\hat{f}_{dm}$ is the estimated maximum Doppler frequency obtained by the FFT operation in [23] aided by decision-directed channel estimates [33]. Finally, $h_0(k;L)$ is used to obtain $\hat{x}(l)$, the decisions of data symbols $x(l)$ for kth frame. In the ATCM system, we select the survival path with the smallest metric, and use the information bits corresponding to this path as the decisions [24].

C. Decision-Directed Channel Estimation

Next, the decisions $\hat{x}(l)$ and robust noise reduction technique are employed in the estimation of channel coefficient $h(k;L)$ as in [33]. To derive the estimation filter, we assume perfect decisions

$\hat{x}(l)$. This assumption is realistic since error propagation is negligible for $P_t \leq 10^{-2}$ employed in practical adaptive modulation systems [27, 28]. When kth frame is not an outage frame, the unbiased least square (LS) estimation of $h(k;L)$ is employed [33]. The variance of this decision-directed estimate is N_0 / LE_d [33] , thus reducing the noise power by approximate factor of L relative to the high-pilot rate method. To reduce noise further, we also employ a robust NR filter that removes the noise power falling outside of the frequency band of the fading signal [6].

Note that the initial channel estimate $h_0(k;L)$ of each frame except the last frame in a superframe is computed by interpolation first, then using filtering (see (7)) to improve accuracy. The associated data symbols are also re-demodulated to reduce decision errors, and channel estimates are updated using these recomputed initial estimates.

The channel estimates are sampled at the frame rate. They are decimated by a factor of K. The one-step prediction of channel coefficient is obtained as [5],

$$\hat{h}(n+1) = \sum_{p=0}^{P_{\text{Pred}}-1} \omega_{\text{pred}}^{*}(p)\tilde{h}(n-p), \tag{9}$$

where P_{pred} is the order, and $\omega_{\text{pred}}(p)$, $p \in [0, P_{\text{pred}}-1]$, are the coefficients of the predictor. The memory span of the predictor is defined as the time interval spanned by the samples used to perform one prediction, i. e., $\tilde{h}(n-p)$, $p \in [0, P_{\text{pred}}-1]$. For this predictor, the memory span is $P_{pred}KL/f_s$ seconds. Note that the inputs are sampled at the superframe rate in (9), while its outputs are computed at the frame rate. This approach improves prediction accuracy relative to prediction followed by interpolation [5, 33]. As in [23], we employ the Burg method to estimate $\omega_{\text{pred}}(p)$ using an observation window that contains B past noise-reduced samples at the rate f_{pred}, where usually $B >> P_{\text{pred}}$. The normalized spatial prediction range in (9) is $f_{dm}KL/f_s$, usually expressed in multiples of carrier wavelength λ [6]. Finally, when the desired prediction interval is longer than K frames, we iterate (9) by using previously predicted samples instead of $\tilde{h}(n)$ [5].

As discussed in section II, the thresholds of adaptive modulation scheme depend on ρ, or the NPMSE of the predictor.

IV. NUMERICAL RESULTS

The standard Jakes model with nine oscillators (34 scatterers) [12] and our realistic physical model [5, 6, 29] are employed in simulations. The geometry of physical model data set used for testing LRP enabled by DANR is shown in [9, Fig. 4]. It represents a typical flat fading urban mobile radio channel. For both models, the rms delay spread σ_d=1 us and f_{dm}=100 Hz. In all simulations, the symbol rate f_s=100 Ksps, the frame size L=10, and K=10 frames per superframe are selected. The sampling frequency of the predictor is f_p=1 KHz=10f_{dm} and the order P_{pred}=20. The observation window of Burg predictor contains 200 and 100 samples for the Jakes and physical models, respectively. The low-rate raw pilot [5] and high-rate pilot methods [7] have pilot rates 10f_{dm} and 100f_{dm}, respectively. Both methods employ Burg predictor. In the high rate pilot method, four filters with smoothing lags {0, 2, 5, 10} [7] and order 20 perform NR. These NR filters are constructed from estimated f_{dm} as proposed in [23]. For fair comparison, the spatial memory span is given by 2λ for all prediction methods.

To illustrate the potential of our DANR method, the performance of several techniques that employ ideal assumptions is illustrated for the Jakes model. The first method employs perfect CSI at the transmitter of the adaptive modulation system. This method provides an upper bound to the achievable spectral efficiency of all other techniques. Second, we show the performance of the outdated CSI scheme, where one noiseless channel coefficient delayed by the normalized prediction range is employed to enable the adaptive modulation [31].

Fig. 2 and 3 illustrate the NPMSE and spectral efficiency of several prediction methods for the Jakes model when the SNR is 20dB. Due to accurate decisions of recently received data symbols, the DANR-aided prediction for the uncoded adaptive modulation has slightly better MMSE than for the adaptive trellis coded modulation. At short to medium prediction ranges (under 0.3λ), the DANR for both uncoded adaptive modulation and ATCM exhibits excellent

prediction accuracy and outperforms the pilot-aided methods. As discussed in [6, 7, 33], the benefit of NR diminishes as the prediction range increases. These figures also show performance for the outdated CSI method.

The performance of the following "MMSE bound" predictor is also shown in the plots for the Jakes model. Suppose channel statistics are known, and consider a linear MMSE predictor that employs pilots transmitted at the data rate. Let the memory span of this predictor be the same as for DANR-aided LRP. Thus, the corresponding filter order of this MMSE predictor is $P_{pred}KL$. Suppose this predictor enables an adaptive modulation scheme with parameters described in section II and bandwidth loss $\eta=1/KL$ as for the DANR method. The resulting MMSE and spectral efficiency of this MMSE predictor serve as lower and upper bounds on the achievable NPMSE and spectral efficiency of our practical technique, respectively. Note that a MMSE predictor that experiences outages at the same rate as in the LRP-aided adaptive modulation system (i.e. utilizes adaptive pilot transmission where one pilot symbol is transmitted in each outage frame) has similar performance to our ideal MMSE predictor [25,33]. Thus, the impact of outages on the performance of DANR is small. This result is due to the utilization of adaptive pilot transmission that significantly improves performance of the DANR method [33].

In Fig. 3, the DANR method has higher spectral efficiency than pilot-aided methods for prediction ranges below 0.4λ. Its gain over the raw pilot method is as high as 1.2 bps/Hz when UAM is used. Even though the ACM is more sensitive to prediction errors than the UAM, the ACM aided by DANR still achieves excellent performance for prediction ranges under 0.2λ, where its spectral efficiency loss relative to the perfect CSI case is at most 0.43 bps/Hz. This prediction range is sufficient for most adaptive modulation applications [6, 8]. For prediction ranges above 0.2λ, the NPMSE of all practical prediction methods simulated in this paper approaches 10^{-1} (see Fig. 2). Thus, as illustrated in , they cross the NPMSE threshold, and the spectral efficiency drops rapidly with the increasing NPMSE at these larger prediction ranges in Fig. 3. We also observe that although the high-rate pilot scheme has better prediction accuracy

than the raw pilot scheme, they have the same spectral efficiency at short prediction ranges (0.1λ) since the former is penalized by high η values, i.e., oversampled pilots. Finally, we note that excellent performance achieved by the MMSE bound method illustrates the potential of the data-aided prediction approach.

Fig. 4 and 5 illustrate the dependency of NPMSE and spectral efficiency, respectively, on the system SNR for the prediction range of 0.2λ for the Jakes and physical models when the UAM is employed. For both models, the proposed DANR method outperforms the pilot-aided methods. The spectral efficiency gain of DANR over the raw pilot method is as large as 0.75 bps/Hz for the Jakes model. We observe that the physical model data set is easier to predict than the Jakes model for most SNR values in Fig. 4, while the opposite conclusion was reached in [5, 6, 30] for high pilot SNR. This can be explained as follows: for lower SNR, noise dominates prediction performance, and the prediction is aided by lower number and non-uniform distribution of reflectors in the realistic physical model data set [33]. On the other hand, for pilot SNR≥30dB, non-stationarity of fading coefficient dominates prediction errors and limits the accuracy of prediction for the physical model relative to the Jakes model [6, 30].

In Fig. 4, the NPMSE degradation of the DANR-aided LRP relative to the MMSE bound predictor is due to several practical limitations, including decision errors, AR coefficient mismatch and robust NR filtering. Moreover, in the DANR method, decomposition of the prediction into three stages (decision-directed NR, robust filtering, and Burg predictor) results in simple but suboptimal solution relative to a single MMSE predictor with very large filter order. Note that in this case the SEs of the DANR and the ideal MMSE bound method are very similar (not shown) since the NPMSE is close to or below the NPMSE threshold for both methods (see Fig. 4).

For the Jakes model, the spectral efficiency gain of trellis-code is illustrated in Fig. 6 for different SNR and prediction ranges. The maximum gain, 0.5 bps/Hz, is achieved at SNR=24dB and the shortest prediction range (0.02λ). As the additive noise level and/or prediction range

increase, the spectral efficiency gap reduces due to less accurate prediction for the adaptive trellis-coded modulation system. To achieve a positive gain, the prediction range has to be less than 0.35λ.

While previous simulations utilized a simple four-state encoder, a more complex Ungerboeck encoder [15] with 64 states is considered in Fig. 7. For short prediction range at 0.1 λ and target BER $P_t=10^{-5}$, the 64-state encoder achieves an additional coding gain of 0.2 bps/Hz over the 4-state encoder. However, as the prediction range grows, this coding gain diminishes, and 64-state encoder is outperformed by the 4-state encoder for prediction range larger than 0.2 λ. Similar result was also observed in [13]. It can be explained intuitively as follows. Increasing the number of states improves the coding gain for medium to high SNR, but degrades the BER at low SNR for the AWGN channel [4]. Therefore, the sensitivity of ACM to prediction errors increases as the number of states grows. This conclusion indicates that the selection of encoder for ACM depends on the prediction range of the application. When the CSI can be reliably predicted, selection of large constraint length improves the spectral efficiency. On the other hand, when the prediction is unreliable, it is unnecessary to use complex encoders/decoders as they will only reduce the spectral efficiency.

V. CONCLUSIONS

A novel DANR method for fading channel prediction is proposed for UAM and ACM systems. It is demonstrated that, compared with previously investigated pilot-aided NR techniques, DANR achieves higher throughput by improving the prediction accuracy while maintaining relatively low pilot rate. The set of SNR values and prediction ranges where positive coding gain is achieved by ACM enabled by DANR-aided LRP is determined, and the effect of the prediction range on the choice of the encoder is analyzed. Finally, it is shown that adaptive modulation aided by LRP has better performance with the realistic physical model than for the Jakes model in the practical SNR range.

REFERENCES

[1] A. J. Goldsmith, and S.-G. Chua, “Variable-rate variable-power MQAM for fading channels,” IEE Trans. Commun., vol. 45, no. 10, pp. 1218-1230, Oct. 1997.

[2] S. Falahati, A. Svensson, T. Ekman and M. Sternad, "Adaptive modulation systems for predicted wireless channels," IEEE Trans. Commun., vol. 52, no. 2, pp. 307-316, Feb. 2004.

[3] S. Falahati, M. Hong, A. Svensson, and M. Sternad, “Adaptive trellis-coded modulation over predicted flat fading channels,” Proc. IEEE VTC, 2003-fall, vol. 3, pp. 1532-1536.

[4] A. J. Goldsmith, and S.-G. Chua, “Adaptive coded modulation for fading channels,” IEEE Trans. Commun., vol. 46, no. 5, pp. 595-602, May 1998.

[5] A. Duel-Hallen, Shengquan Hu and H. Hallen, “Long-range prediction of fading signals,” IEEE Signal Process. Mag., vol. 17, no. 3, pp. 62-75, May. 2000.

[6] A. Duel-Hallen, “Fading channel prediction for mobile radio adaptive transmission systems,” Proc. IEEE, vol. 95, no. 12, pp. 2299-2313, Dec. 2007.

[7] T. Ekman, “Prediction of mobile radio channels: modeling and design,” Ph. D. Dissertation, Uppsala Univ., Sweden, 2002.

[8] X. Cai and G. B. Giannakis, “Adaptive PSAM accounting for channel estimation and prediction errors,” IEEE Trans. Wireless Commun., vol. 4, no. 1, pp. 246-256, Jan. 2005.

[9] T. Jia, A. Duel-Hallen, and H. Hallen, “Improved long-range prediction with data-aided noise reduction for adaptive modulation systems,” Proc. 43rd Annual Conference on Information Sciences and Systems, Princeton, NJ, 2008.

[10] D. Schafhuber and G. Matz, “MMSE and adaptive prediction of time-varying channels for OFDM systems,” IEEE Trans. Wireless Commun., vol. 4, no. 2, pp. 593-602, Mar. 2005.

[11] T. Eyceoz, S. Hu, A. Duel-Hallen, H. Hallen, “Performance analysis of long range prediction for fast fading channels,” Proc. 33rd Annual Conference on Information Sciences and Systems (CISS), 1999, vol. 2, pp. 656-661.

[12] W. C. Jakes, *Microwave Mobile Communications*, New York, NY : Wiley, 1974.

[13] S. Zhou, G. B. Giannakis, “How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?” IEEE Trans. Wireless Commun., vol. 3, no. 4, pp. 1285-1294, July 2004

[18] K. J. Hole, H. Holm, and G. E. Øien, "Adaptive multidimensional coded modulation over flat fading channels," IEEE J. Selected Areas Commun., vol. 18, no. 7, pp. 1153-1158, July 2000.

[19] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets part II: state of the art," IEEE Commun. Mag., vol. 25, no. 2, pp. 12-21, Feb. 1987

[20] C. H. Aldana and J. Cioff, "Channel tracking for multiple input, single output systems using EM algorithm," in Proc. IEEE Int. Conf. Commun., 2001, vol. 2, pp. 586-590.

[21] S. Zhou, and G. B. Giannakis, "Adaptive modulation for multiantenna transmissions with channel mean feedback," IEEE Trans. Wireless Commun., vol. 3, no. 5, pp. 1626-1636, Sept. 2004.

[22]J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," IEEE Trans. Vech., vol. 40, no. 4, pp. 686-693, Nov. 1991.

[23] K.E. Baddour, N. C. Beaulieu, "Improved pilot-assisted prediction of unknown time-selective Rayleigh channels," Proc. IEEE Int. Conf. Commun., June 2006, vo. 11, pp. 5192-5199

[24] R. Raheli, A. Polydoros and C. K. Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environments," IEEE Trans. Commun., vol. 43, pp.354-364, Feb.-Apr. 1995.

[27] J. Proakis, *Digital Communications*, 4th ed. New York, NY: McGraw-Hill, 2001.

[28] L. Lindbom, A. Ahlen, M. Sternad, and M. Falkenstrom, "Tracking of time-varying mobile radio channels - part II: a case study," IEEE Trans. Commun., vol. 50, no. 1, pp. 156-167, Jan. 2002.

[29] H. Hallen, A. Duel-hallen, S. Hu, T.-S. Yang, and M. Lei, "A physical model for wireless channels to provide insights for long range prediction," Proc. IEEE MILCOM, Oct. 2002, vol. 2, pp. 627-631.

[31] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," IEEE Trans. Commun., vol. 47, no. 6, pp. 844-855, June 1999.

[33] T. Jia, "Single and multicarrier adaptive transmission systems with long-range prediction aided by noise reduction," Ph. D. Thesis, North Carolina State University, in preparation.

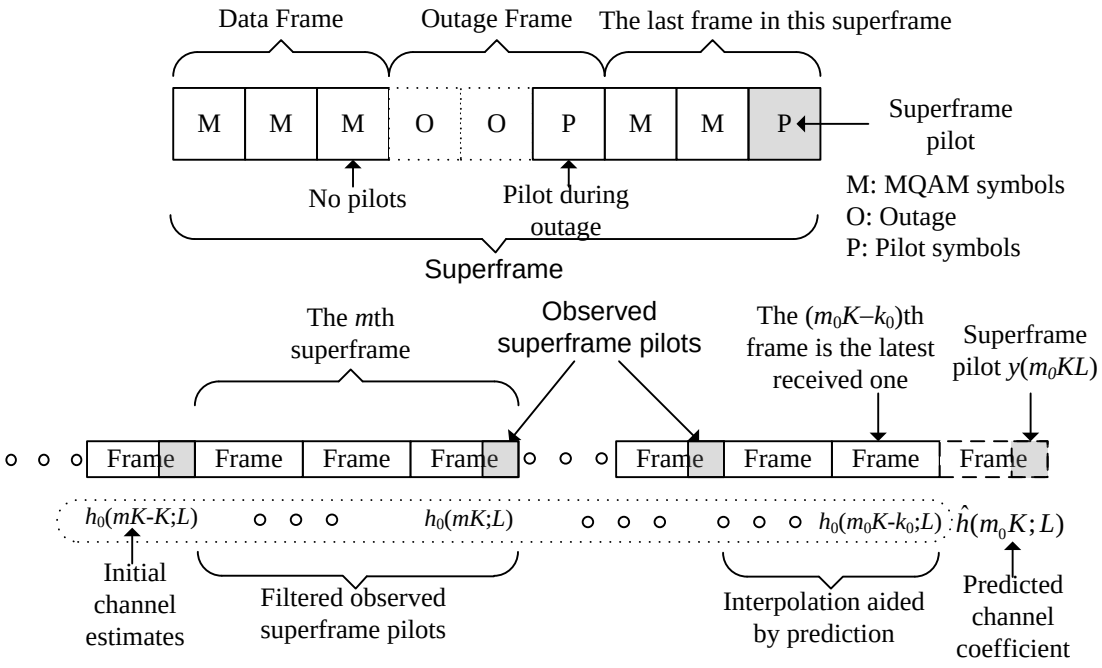

Fig. 1. Frame structure and initial channel estimates for demodulation (see equation (7)). L=3 symbols/frame, K=3 frames/superframe. (Keep this figure. Delete Fig. 1 and 2?)

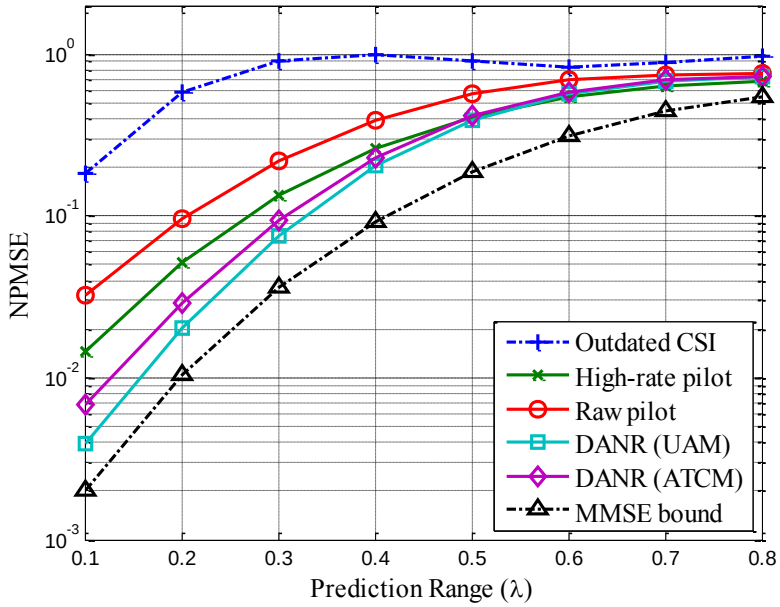

Fig. 2. NPMSE vs. prediction range for SNR=20dB, Jakes model, and P_t=10^{-4}. (UAM: uncoded adaptive modulation, ATCM: adaptive trellis coded modulation)

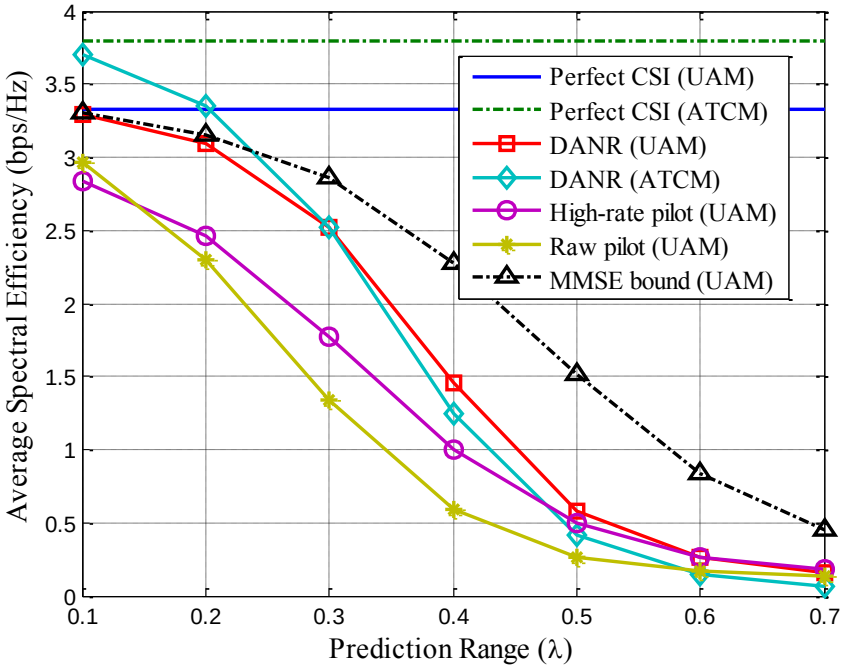

Fig. 3. Spectral efficiency vs. prediction range for SNR=20dB, Jakes model, and P_t=10^{-4}

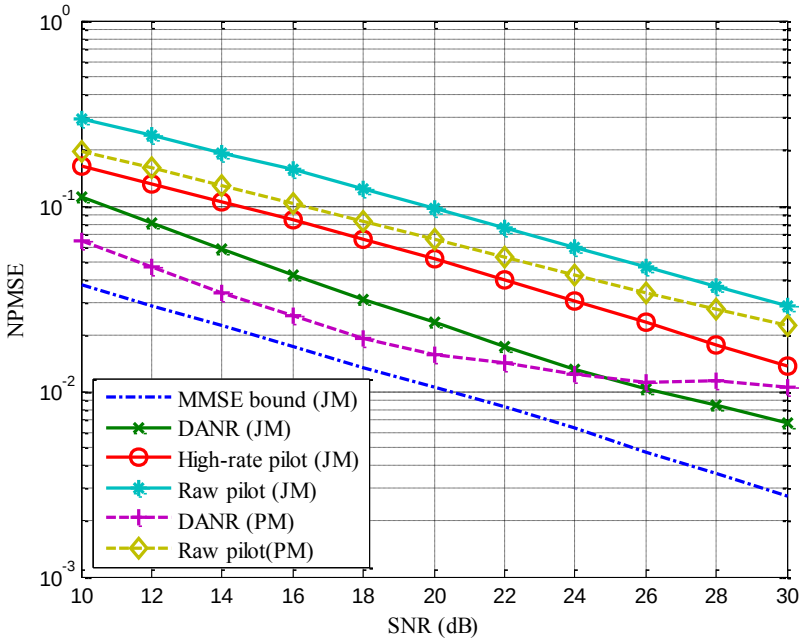

Fig. 4. NPMSE vs. SNR for prediction range at 0.2λ, UAM, P_t =10^{-4}, Jakes model (JM) and physical model (PM) (New Fig. 4)

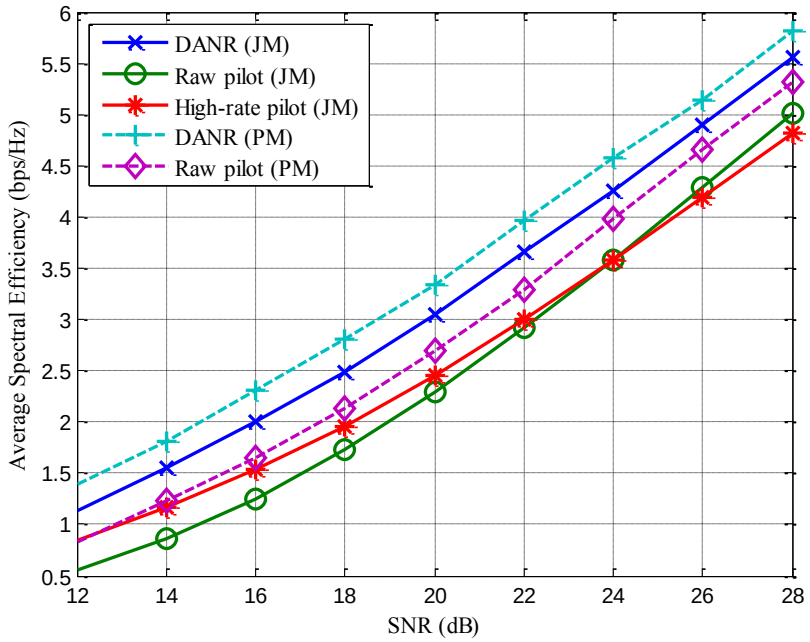

Fig. 5. Spectral efficiency vs. SNR for prediction range at 0.2λ, uncoded adaptive modulation, P_t =10^{-4}, Jakes model (JM) and physical model (PM) (*They are all about UAM*)

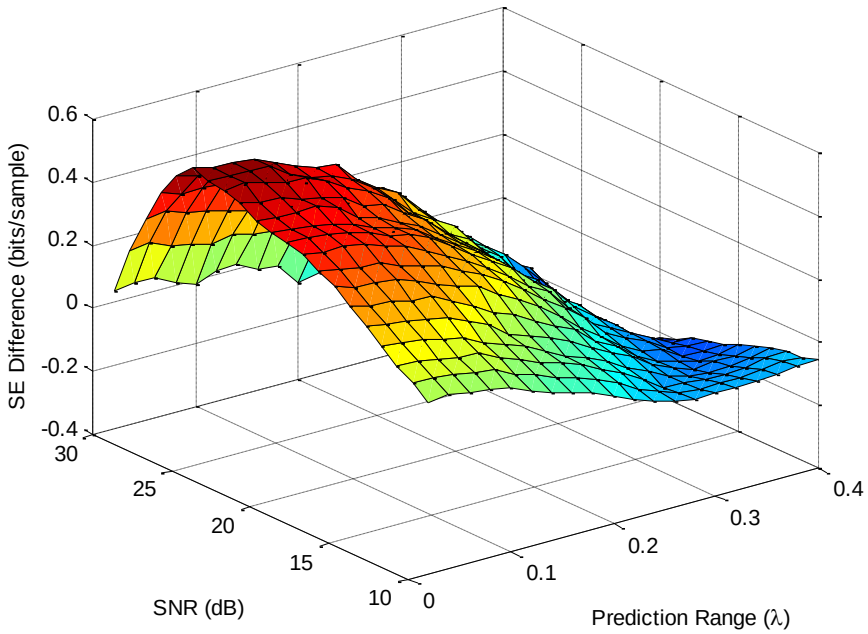

Fig. 6 (New Fig. 6) Difference in spectral efficiency between ATCM and uncoded adaptive modulation with LRP enabled by DANR for Jakes model and P_t =10^{-4}

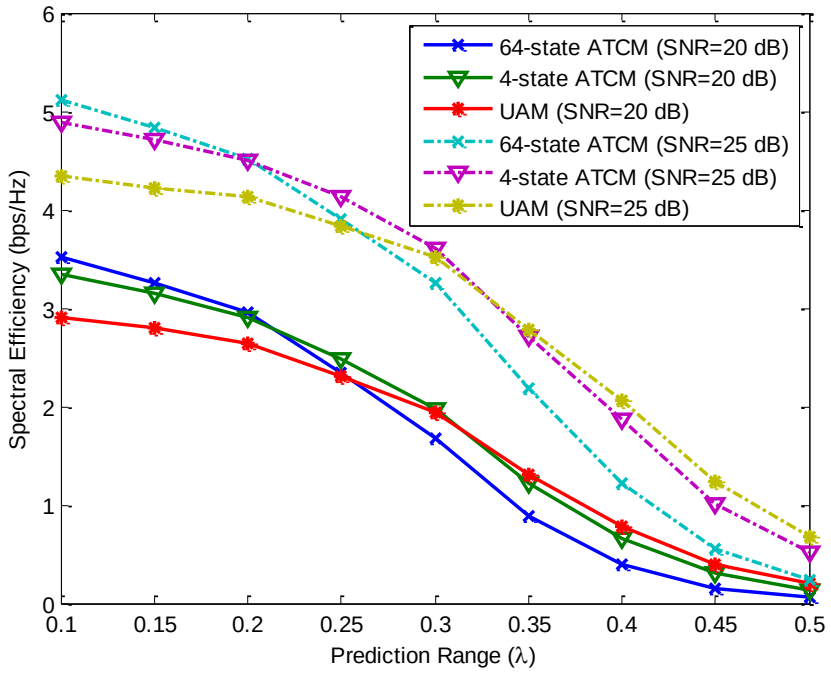

Fig. 7 Spectral efficiency vs. prediction range for Jakes model and P_t =10^{-5} (UAM: uncoded adaptive modulation, ATCM: adaptive trellis coded modulation)