# Improved Long-Range Prediction with Data-Aided Noise Reduction for Adaptive Modulation Systems

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

***Abstract*—A novel data-aided noise reduction (DANR) method is proposed to enhance the accuracy of long-range prediction (LRP) for wireless fading channels, thereby improving the spectral efficiency (SE) of adaptive modulation (AM) system enabled by the LRP. This method includes an adaptive pilot transmission mechanism, robust noise reduction and decision-directed channel estimation. An improved practical AM scheme is used to test the proposed DANR method. Since this method maintains low pilot rates, it results in higher SE than previously proposed noise reduction (NR) techniques, which rely on oversampled pilots. These conclusions are confirmed for practical prediction ranges using the standard Jakes model and our realistic physical model.**

## I. Introduction

Adaptive modulation (AM) has been investigated extensively in wireless communication literature due to its effectiveness and simplicity [1]. To ensure high spectral efficiency (SE) for AM systems, reliable channel state information (CSI) is required. To compensate for the feedback delay, data processing, and system constraints, this CSI needs to be predicted ahead for rapidly time variant fading channels. Several fading prediction methods have been proposed to enable mobile radio AM system [2], [3].

In this paper, we focus on the autoregressive (AR) model-based linear prediction (LP), referred to as the long-range prediction (LRP) method, which was shown to achieve superior performance for realistic channel models and measured channels [2]–[4]. The LRP method achieves very low mean-square error (MSE) at prediction ranges over one carrier wavelength $\lambda$ when the signal to noise ratio (SNR) of the observations is sufficiently high [2]. However, at low SNR, this method suffers from significant performance degradation. To improve the prediction accuracy at low SNR, utilization of highly oversampled pilots to perform noise reduction (NR) prior to LP was proposed in [5]. We refer to this NR technique as the *high-rate pilot* method. This method improves the prediction accuracy over the low-rate (on the order of maximum Doppler shift) *raw pilot* method that relies only on the pilot symbols necessary for fading coefficient prediction. The NR method in [5] requires perfect knowledge of channel statistics. A practical high-rate pilot NR method was developed and shown to provide reliable prediction for SNR$\geq$20dB in [6]. However, these methods consume excessive system bandwidth due to transmitting oversampled pilots, so the overall SE does not increase significantly despite better prediction accuracy. The optimal rate and power of raw pilots that maximize the SE of an AM system under given power constraint was investigated in [7]. However, the pilots still consume a large portion of the available bandwidth and power in that method [7].

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

In this paper, we propose to exploit the data symbols in NR to achieve high accuracy of prediction without significantly increasing the pilot rate and power, thereby improving the SE. The data-aided methods have been widely used for channel equalization [8], estimation, and tracking [9]. In [10], decision-directed short-range prediction of fading channels was investigated, and in [2], [11], data-aided NR was employed in LRP to enable adaptive channel inversion. However, these decision-directed methods are not applicable to the AM system directly.

We propose a novel data-aided noise reduction (DANR) method for the LRP-aided AM system. This method comprises an adaptive pilot transmission mechanism that compensates for outages, robust noise reduction and decision-directed channel estimation. The prediction accuracy and SE 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 prediction ranges. The NR techniques are tested using the standard Jakes model and our realistic physical model.

The remainder of this paper is organized as following. The AM system enabled by LRP is discussed in section II. The proposed DANR method is described in section III. Finally, numerical results and conclusions are presented in sections IV and V, respectively.

## II. AM AIDED BY PREDICTED CSI

The proposed system model is illustrated in Fig. 1, and frequently used notation is summarized in Table I. We consider a single-carrier flat fading channel and assume the symbol rate $f_s$. The equivalent lowpass received $l$th sample is

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

where $x(l)$, $h(l)$, and $w(l)$ are the transmitted symbol, the complex fading channel coefficient, and complex white Gaussian noise with variance $N_0$, respectively. The channel coefficient is modeled as complex zero-mean Gaussian process

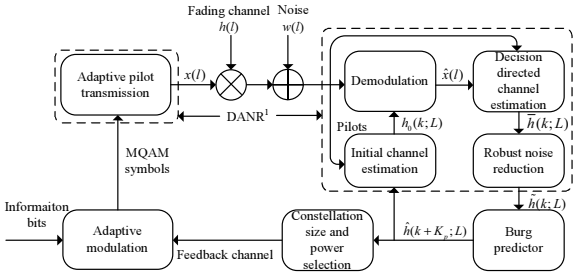


Fig. 1: System Diagram

with unit variance. We employ the widely used 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 can be either pilots or data symbols. Without loss of generality, we assume that $x(l) = E_p$ for pilots, where $E_p$ is the energy of the pilot symbol.

To enable AM, a LP generates predicted channel coefficient $\hat{h}(l)$. The random variables $\hat{h}(l)$ and $h(l)$ are jointly gaussian distributed with the cross-correlation [13]

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

where $\sigma_{\hat{h}}^2 = E\left[|\hat{h}(l)|^2\right]$. Denote $\gamma(l) = |h(l)|^2$ and $\hat{\gamma}(l) = |\hat{h}(l)|^2$, where, without loss of generality, we omit the index $l$. The PDF of $\hat{\gamma}$ is $p(\hat{\gamma}) = \exp(-\hat{\gamma}/\sigma_{\hat{h}}^2)/\sigma_{\hat{h}}^2$.

Consider an AM system that employs M-quadrature amplitude modulation (MQAM) constellation sizes $\{M_1, M_2, \ldots, M_N\}$, $0 < M_{i_1} < M_{i_2}$ for $i_1 < i_2$, and switching thresholds $\{\gamma_0, \gamma_1, \ldots, \gamma_{N+1}\}$ used to select the constellation size used for transmission, where $0 \le \gamma_i \le \gamma_{i+1}$, $\gamma_0 = 0$, 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. When $0 \le \hat{\gamma} < \gamma_1$, an outage occurs. We employ AM with constant power allocation, i. e., the same average MQAM symbol energy $E_d$ is utilized for all $M_i > 0$, due to its excellent performance and low required feedback rate [7], [14]. Given fixed average SNR $\bar{E}_d/N_0$, where $\bar{E}_d$ is the average data symbol energy, $E_d$ is determined by $\bar{E}_d = E_d \int_{\gamma_1}^{\infty} p(\hat{\gamma})d\hat{\gamma} = E_d \exp\left[-\gamma_1/\sigma_{\hat{h}}^2\right]$ [7]. For the constellation size $M_i$, we obtain the instant BER $PE_i(E_d, \hat{\gamma})$ given $E_d$ and $\hat{\gamma}$ as in [7], [14]. When the SE loss caused by transmitting pilots is not taken into account, the SE of data symbols is evaluated as $S_d(\gamma_1, \ldots, \gamma_N) = \sum_{i=1}^{N}\left[log_2 M_i \int_{\gamma_i}^{\gamma_{i+1}} p(\hat{\gamma})d\hat{\gamma}\right]$. Using the average BER criterion, the thresholds $\{\gamma_1, \ldots, \gamma_N\}$ are selected so that $S_d(\gamma_1, \ldots, \gamma_N)$ is maximized under the constraint that the average BER is lower than the specified target BER $P_t$ [14].

Note that $\hat{\gamma}$ is a biased prediction, while unbiased power prediction $\hat{\gamma} + E(\gamma - \hat{\gamma})$ was employed in [5], [14]. However, the performance of the AM system is determined by the statistical characterization of the predicted channel coefficient $\hat{h}(l)$, not the predicted power [15]. Therefore, biased and unbiased power predictors result in the same SE.

In our simulations, the cross-correlation $\rho$ in (2) is measured from the observed data set. For the fixed SNR and target BER, the adjustment of thresholds in the AM system is determined by $\rho$. For the linear minimum MSE (LMMSE) predictor, the *normalized MSE* is NMSE $\triangleq E\left[|h(l) - \hat{h}(l)|^2\right]/E[|h(l)|^2]$ [13], and for general LP, NMSE $\ge 1 - \rho^2$ [13], [15]. While the NMSE is often used to measure prediction performance, it is not appropriate as an indicator of SE in AM aided by fading prediction. Since this SE is determined by $\rho$ , we define the *normalized prediction MSE* as NPMSE $\triangleq 1 - \rho^2$. The NPMSE is used as a performance measure of predictors in all our plots.

In this paper, we employ the set of constellation sizes $\{2, 4, 16, 64\}$. Fig. 2 shows the SE of this AM system for $P_t = 10^{-3}$ and several average SNR levels. For a NPMSE on the order of $10^{-2}$ or smaller, the SE is not sensitive to the variation of the NPMSE. On the other hand, the SE decreases rapidly when the NPMSE approaches $10^{-1}$. For prediction ranges over $0.1$ employed in adaptive transmission systems [3], NPMSE usually falls into the latter region for practical SNR levels (less than 30dB) [2], [5]. Thus, NR that can significantly lower the NPMSE is necessary to enable AM in practical systems.

TABLE I: Notation

| | |
|---|---|
| $f_{dm}$ | Maximum Doppler frequency |
| $f_s$ | Symbol rate |
| $h(l)$ | Channel coefficient of the $l$th symbol |
| $h(k;L)$ | Channel coefficient of the $k$th frame |
| $h_0(k;L)$ | Initial channel estimate for the $k$th frame |
| $\bar{h}(k;L)$ | Decision directed channel estimate for the $k$th frame |
| $\tilde{h}(k;L)$ | The noise-reduced channel estimate for the $k$th frame |
| $\hat{h}(k;L)$ | The predicted channel coefficient for the $k$th frame |
| $\tilde{h}(n)$, $\hat{h}(n)$ | Down-sampled $\tilde{h}(k;L)$ and $\hat{h}(k;L)$ at rate $f_{\text{pred}}=\frac{f_s}{LK}$ |
| $L$ | Frame size |
| $K$ | Superframe size |
| $E_p$ | Pilot symbol energy |
| $E_d$ | Average energy of MQAM symbols |
| $\bar{E}_d$ | Average data (outage and MQAM) symbol energy |
| $\bar{E}_s$ | Average energy of all symbols |

## III. LONG RANGE PREDICTION WITH DATA-AIDED NOISE REDUCTION

### A. Frame Structure and Adaptive Pilot Transmission

The frame structure employed in this paper is illustrated in Fig. 3. At the transmitter of the proposed AM system, the transmitted symbols are grouped into frames with $L$ symbols in each frame. To reduce the feedback rate and the complexity of NR, the constellation size is fixed in each frame. Thus, we assume that the fading coefficient is constant during each frame and choose $f_s/L = 100 f_{dm}$ to justify this assumption [16]. We use vectors $\boldsymbol{x}(k) = [x(kL), x(kL-1), \ldots, x(kL-L+1)]^T$ and $\boldsymbol{y}(k) = [y(kL), y(kL-1), \ldots, y(kL-L+1)]^T$ to represent the transmitted and received symbols of the $k$th frame, respectively (see (1)). In addition, we denote $h(k;L) \approx$

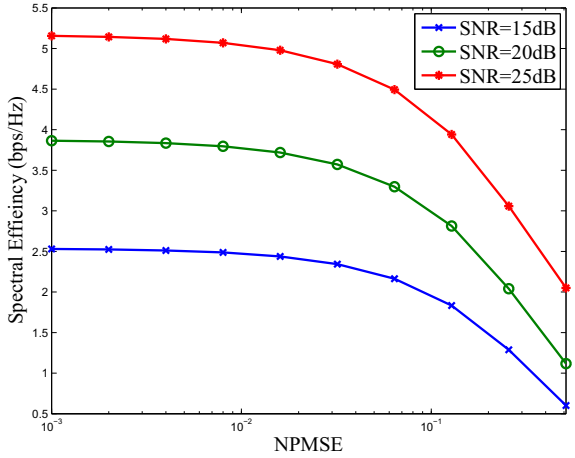


Fig. 2: Spectral Efficiency of AM vs. NPMSE for target BER $P_t = 10^{-3}$.

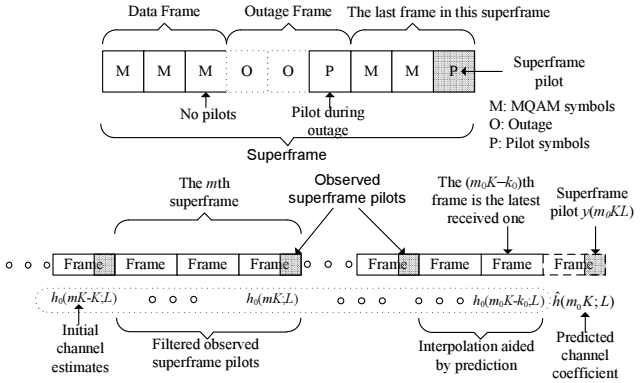


Fig. 3: Frame structure and initial channel estimates for demodulation (see equation (3)). L=3 symbols/frame, K=3 frames/superframe.

$h(kL-l)$, $l \in [0, L-1]$ as the channel coefficient associated with the $k$th frame.

If one pilot is transmitted in each frame to facilitate NR in LP, the pilot rate is $f_p = f_s/L = 100 f_{dm}$, which is the pilot rate in [5]. However, as stated earlier, this approach wastes a large portion of the available bandwidth due to transmitting oversampled pilots.

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 $m$th superframe are represented by a vector of length $KL$, $\boldsymbol{y}_s(m) = \left[\boldsymbol{y}^T(mK) \quad \ldots \quad \boldsymbol{y}^T(mK-K+1)\right]^T$. 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 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. Denote the percentage of bandwidth occupied by pilots as $\eta = f_p/f_s$. When pilot symbols are taken into account, the average SE of the system is $\bar{S} = (1-\eta) S_d(\gamma_1, \ldots, \gamma_N)$ bps/Hz. For our method, $\eta = \left[1 + (L-1)\int_0^{\gamma_1} p(\hat{\gamma}) d\hat{\gamma}\right]/KL$, which satisfies $1/KL \le \eta \le 1/L$, while $\eta = 0.1$ for the high-rate pilot method [5], and $\eta$ varies from 0.005 to 0.01 for the raw pilot method [2].

While the loss in power and bandwidth caused by transmitting pilots is not considered in [5], [14], it needs to be taken into account in the system performance optimization. Given the average symbol energy $\bar{E}_s$, the pilot energy and the average energy of data symbols have to satisfy the constraint $\bar{E}_s = \eta E_p + (1-\eta)\bar{E}_d$. In this paper, we select $\bar{E}_s = E_p = \bar{E}_d$ as in a pilot symbol assisted modulation (PSAM) system [17]. Although higher SE is achieved by optimally allocating power to data symbols and pilots as in [7], the SE loss of our method relative to the optimal allocation is small [15]. Our selection greatly simplifies power allocation since $E_p$ and $E_s$ are constant for given $\bar{E}_d$. In addition, the selected pilot energy does not depend on the accuracy of prediction, while the pilot energy is a function of NPMSE in the variable pilot energy method [7]. Thus, our methods reduces the required feedback relative to [7] since the pilot energy is known at the transmitter.

### B. Initial Channel Estimation for Demodulation

To demodulate received data symbols, we obtain initial channel estimates as illustrated in Fig. 3. 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 $m_0$th superframe. The initial estimate for frame $(mK-k) \le (m_0K-k_0)$ in superframe $m \le m_0$ is computed as

$$h_0(mK-k;L) = \quad (3)$$
$$\begin{cases} \frac{k}{K} h_0(m_0K-K;L) + \frac{K-k}{K}\hat{h}(m_0K;L), & \text{if } m=m_0 \text{ and } k_0 \neq 0 \\ \sum_{p=0}^{P_{\text{est}}-1} \omega^*_{est}(k,p)\, y(mKL - pKL), & \text{otherwise} \end{cases}$$

In the first case in (3), 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 simple 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 $m_0$th superframe), where the prediction method is discussed in section III-E.

In the second case in (3), the $m$th superframe pilot has been observed (see Fig. 3). 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 $\boldsymbol{y}_p = [y(mKL) \quad y(mKL-KL) \quad \ldots \quad y(mKL-P_{\text{est}}KL+KL)]^T$, where $P_{\text{est}}$ 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 $[-f_{dm}, f_{dm}]$ [6]. The $P_{\text{est}} \times 1$ vector of filter

coefficients is constructed as

$$[\omega_{est}(k,0) \quad \dots \quad \omega_{est}(k,P_{\text{est}}-1)]^T = (\boldsymbol{R}+N_0\boldsymbol{I})^{-1}\boldsymbol{r}(k) \quad (4)$$

where the $P_{\text{est}}\times P_{\text{est}}$ autocorrelation matrix $\boldsymbol{R}$ has elements $\boldsymbol{R}_{i_1,i_2}=\text{sinc}[(i_1-i_2)\hat{f}_{dm}KL/f_s]$, the $P_{\text{est}}\times 1$ cross-correlation vector $\boldsymbol{r}(k)$ has elements $[\boldsymbol{r}(k)]_i=\text{sinc}[(iK-k)\hat{f}_{dm}L/f_s]$, $\boldsymbol{I}$ is the $P_{\text{est}}\times P_{\text{est}}$ identity matrix, and $\hat{f}_{dm}$ is the estimated maximum Doppler frequency obtained by the FFT operation in [6] aided by decision-directed channel estimates [15].

Finally, $h_0(k;L)$ is used to obtain $\hat{x}(l)$, the decisions of data symbols $x(l)$ for $k$th frame. By definition, $\hat{x}(l)=\sqrt{E_p}$ for pilot symbols, and $\hat{x}(l)=0$ during outage.

### C. Decision-Directed Channel Estimation

Next, the decisions $\hat{x}(l)$ are employed in the estimation of channel coefficient $h(k;L)$ at the frame rate. To derive the estimation filter, we assume perfect decisions $\hat{x}(l)$. This assumption is realistic since error propagation is negligible for $P_t \le 10^{-2}$ employed in practical AM systems [8], [9]. When $k$th frame is not an outage frame, the unbiased least square (LS) estimation of $h(k;L)$ is obtained as [15].

$$\bar{h}(k;L)=\frac{\sum_{l=kL-L+1}^{kL} y(l)\hat{x}^*(l)}{\sum_{l=kL-L+1}^{kL}|\hat{x}(l)|^2}=h(k;L)+\frac{\sum_{l=kL-L+1}^{kL} w(l)\hat{x}^*(l)}{\sum_{l=kL-L+1}^{kL}|\hat{x}(l)|^2}. \quad (5)$$

The variance of this decision-directed estimate is [15]

$$v(k;L)=E\left[|\bar{h}(k;L)-h(k;L)|^2\right]=\frac{N_0}{\sum_{l=kL-L+1}^{kL}|\hat{x}(l)|^2}. \quad (6)$$

Since $\sum_{kL-L+1}^{kL}|\hat{x}(l)|^2 \approx LE_d \gg E_p$ for large frame size $L$, the noise power is greatly reduced at the output of (5). For outage frames, the noisy pilot observation is employed for channel estimation, i. e., $\bar{h}(k;L)=y(kL)/\sqrt{E_p}=h(k;L)+w(kL)/\sqrt{Ep}$ and $v(k;L)=N_0/E_p$.

### D. Robust Noise Reduction

Since the sampling rate in (5) is given by the frame rate, the channel coefficients are highly correlated, while the noise samples are independent at the output $\bar{h}(k;L)$ of the decision-directed estimator. Thus, the power spectrum density of the noise samples is much wider than that of the fading signal. To reduce noise further, we pass $\bar{h}(k;L)$ through a robust NR filter that removes the noise power falling outside of the frequency band of the fading signal. In vector form, the input to this NR filter is $\bar{\boldsymbol{h}}(k;L)=[h(k;L) \quad \dots \quad h(k-P_{NR}+1;L)]^T$, where $P_{NR}$ is the filer order. Similarly to (4), the coefficients of this filter are expressed as a $P_{NR}\times 1$ vector

$$\boldsymbol{\omega}_{NR}(k)=\left[\boldsymbol{R}_{NR}+\boldsymbol{V}(k)\right]^{-1}\boldsymbol{r}_{NR} \quad (7)$$

where the $P_{NR}\times P_{NR}$ matrix $\boldsymbol{R}_{NR}$ has elements $[\boldsymbol{R}_{NR}]_{i_1,i_2}=\text{sinc}[(i_1-i_2)\hat{f}_{dm}L/f_s]$, the $P_{NR}\times 1$ vector $\boldsymbol{r}_{NR}$ has elements $[\boldsymbol{r}_{NR}]_i=\text{sinc}[i\hat{f}_{dm}L/f_s]$, and $\boldsymbol{V}(k)=\text{diag}[v(k;L),\dots,v(k-P_{NR}+1;L)]$. Note that $v(k;L)$ is time-variant and depends on whether given frame experiences an outage. The outputs of the NR filter are

$$\tilde{h}(k;L)=\boldsymbol{\omega}_{NR}^H(k)\bar{\boldsymbol{h}}(k;L) \quad (8)$$

Our filter is similar to the NR filter used in [6], but we employ decision-directed channel estimates instead of the over-sampled pilot observations.

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 (3)) to improve accuracy. The associated data symbols are also re-demodulated to reduce decision errors, and (5) and (8) are updated using these recomputed initial estimates.

### E. Long-Range Prediction

Since the outputs of (8) are generated at the frame rate, they are decimated by factor $J$ as each new sample $\tilde{h}(k;L)$ is computed, so the effective sampling rate used for prediction $f_{\text{pred}}=f_s/(JL)$ is on the order of $f_{dm}$ [2], [3], [5]. In our simulation, $J=K$, the number of frames per superframe. We denote the sequences $\tilde{h}(k;L)$ and $\hat{h}(k;L)$, decimated at rate $f_{\text{pred}}$, as $\tilde{h}(n)$ and $\hat{h}(n)$, respectively. The one-step prediction of channel coefficient is obtained as [2],

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

where $P_{\text{pred}}$ is the order of predictor, and $\omega_{\text{pred}}(p)$, $p \in [0,P_{\text{pred}}-1]$, are the coefficients of the predictor. Note that a low sampling rate $f_{\text{pred}}$ is employed in (9), while its outputs are computed at the frame rate. This approach improves prediction accuracy relative to prediction followed by interpolation [2], [15]. As in [6], 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 $f_{\text{pred}}$, where usually $B \gg P_{\text{pred}}$. The normalized spatial prediction range in (9) is $f_{dm}JL/f_s$, usually expressed in multiples of carrier wavelength $\lambda$ [3]. Finally, when the desired prediction interval is longer than $J$ frames, we iterate (9) by using previously predicted samples instead of $\tilde{h}(n)$ [2].

The NPMSE and SE of this LRP form a feedback loop. Clearly, the system SE is a function of the NPMSE. Moreover, the accuracy of DANR-aided LRP increases with the effective data rate of the system. Since the theoretical solution for the equilibrium point of this feedback system is very complex, we resort to simulations. First, we employ a NPMSE guess to calculate the thresholds of AM. If the simulated NPMSE is higher/lower than the initial guess, the guessed NPMSE is increased/decreased by a small step $\Delta$. This process is iterated until the guess and the simulated NPMSE are sufficiently close. The final NPMSE and SE represent the equilibrium point of this adaptive system.

TABLE II: Parameters used in the simulation for Jakes model

| | |
|---|---|
| $f_{dm}$ | 100Hz |
| $f_s$ | 100Ksps |
| Frame rate $f_m$ | 10K frames/s |
| $L$ | 10 |
| $K$ | 10 |
| $f_{\text{pred}}$ | 1KHz |
| FFT Window size | 100 |
| Order of PSAM filter $P_{\text{est}}$ | 10 |
| Order of NR filter $P_{NR}$ | 20 |
| Order of predictor $P_{\text{pred}}$ | 20 |
| Observation window size $B$ | 200 |
| Target BER $P_t$ | $10^{-3}$ |

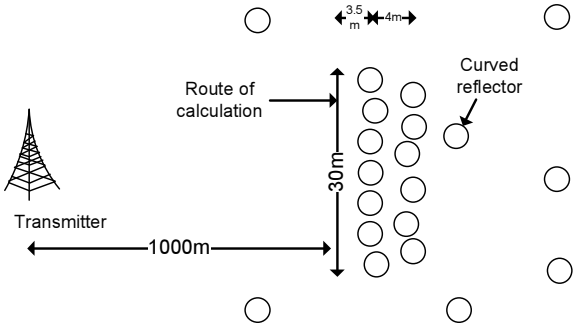


Fig. 4: Geometry of physical model

## IV. Numerical Results

The simulation parameters that represent a typical flat fading urban mobile radio channel are listed in Table II. The standard Jakes model with nine oscillators (34 scatterers) [12] is employed. In addition, we utilize our realistic physical model [18]. It was demonstrated in [2], [3], [18] that this model is more suitable than the Jakes model for validating the LRP. We employ the physical model geometry in Fig. 4. This data set represents a challenging realistic scenario [15]. The same parameters are utilized as for the Jakes model (Table II) with the exception of the observation window size $B$ of Burg predictor, which is reduced to 100 samples to accommodate the time-variant channel statistics.

In addition to the proposed DANR method, we also simulated two pilot-aided LRP techniques, the raw pilot and high-rate pilot methods. In the raw pilot method without NR, noisy superframe pilots at rate $10f_{dm}$ are employed in the Burg predictor. In the high-rate pilot method [5], one pilot is transmitted in each frame, i. e., the pilot rate is $f_p = 100f_{dm}$ and $\eta = f_p/f_s = 0.1$. The NR is performed using four filters with smoothing lags $\{0, 2, 5, 10\}$ [5] and filter order 50. The outputs of these NR filters are down-sampled at a rate of $10f_{dm}$ and then fed into the MMSE predictor. As in [5], perfect knowledge of channel statistics is assumed in the design of this method, and it is only tested for the Jakes model. For all prediction methods, we choose the order of the predictor $P_{\text{pred}} = 20$.

To illustrate the potential of our NR method, the performance of several techniques that employ ideal assumptions is illustrated for the Jakes model. First, we plot the SE when perfect CSI is available at the transmitter of the AM system, and $\eta = 0$. This method provides an upper bound to the achievable SE of all other techniques. We also simulated the performance of the high-SNR pilot method, i.e., the raw pilot method where the pilot SNR=60dB is much higher than for the data symbols without incurring power penalty. This method was employed in [2] and provides a lower bound on the achievable NPMSE for NR methods. Finally, we plot the performance of the outdated CSI scheme, where one noiseless channel coefficient delayed by the normalized prediction range is employed to enable the AM [19].

Fig. 5 and 6 illustrate the variation of NPMSE and SE, respectively, as the prediction range increases. At short to medium prediction ranges (under $0.6\lambda$), the DANR method exhibits excellent prediction accuracy, and its NPMSE is lower than those of the pilot-aided methods. At larger prediction ranges, the improvement in NPMSE provided by NR is less significant, and the errors caused by imperfect estimation of predictor coefficients in Burg method result in slight degradation of the raw pilot and the DANR methods relative to the high-rate pilot method, which has perfect knowledge of channel statistics. The DANR also has higher SE than pilot-aided methods for prediction ranges below 0.6 $\lambda$. Its gain over the raw pilot method is as high as 1.0 bps/Hz. Moreover, our DANR method has excellent performance for prediction ranges$\leq 0.2\lambda$, where its SE loss relative to the perfect CSI case is at most 0.22 bps/Hz. This prediction range is sufficient for most AM applications [3], [7]. For prediction ranges$> 0.2\lambda$, the NPMSE of all practical prediction methods simulated in this paper approaches $10^{-1}$ (see Fig. 5). Thus, as illustrated in Fig. 2, they cross the NPMSE threshold, and their SE drops rapidly with the increasing NPMSE at these larger prediction ranges in Fig. 6. Although the high-rate pilot scheme has better prediction accuracy than the raw pilot scheme, its SE is lower at short prediction ranges ($0.1\lambda$) since it is penalized by high $\eta$ values, i.e., oversampled pilots. Finally, much higher SE is maintained by the high-SNR pilot method relative to other prediction methods for medium to high prediction ranges. This result demonstrates that there is significant potential for performance improvement in the area of NR for fading prediction.

Fig. 7 and 8 illustrate the dependency of NPMSE and SE, respectively, on the system SNR for the prediction range of $0.2\lambda$ for the Jakes and physical models. For both models, the proposed DANR method outperforms the pilot-aided methods. The SE gain of DANR over the raw pilot method is as large as 0.6 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. 7, while the opposite conclusion was reached in [2], [3], [18] for high-SNR pilot method. 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 [15]. On the other hand, for pilot SNR$\geq$30dB, non-stationarity dominates prediction errors and limits the accuracy of prediction for the physical model relative to the Jakes model [3], [18].

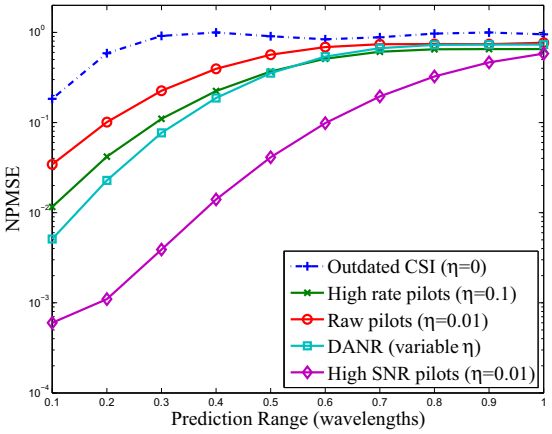


Fig. 5: NPMSE vs. prediction range for SNR=20dB, Jakes model

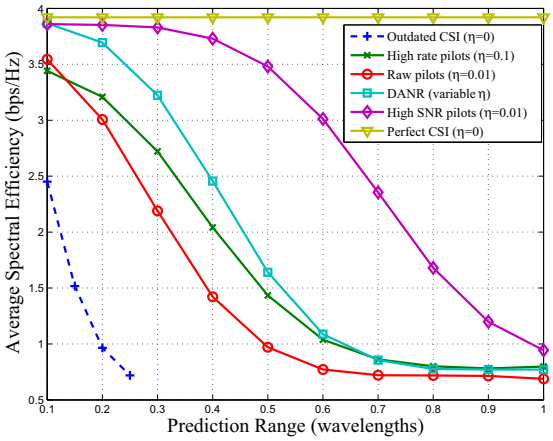


Fig. 6: Spectral efficiency vs. prediction range for SNR=20dB; Jakes model;

## V. Conclusions

A novel DANR method for fading channel prediction is proposed. This technique achieves higher throughput than the previously investigated oversampled pilot methods for the standard Jakes model and for our realistic physical model.

## References


[1] 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.

[2] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Process. Mag.*, vol. 17, pp. 62–75, May 2000.

[3] A. Duel-Hallen, "Fading channel prediction for mobile radio adaptive transmission systems," *Proc. IEEE*, vol. 95, pp. 2299–2313, Dec. 2007.

[4] S. Semmelrodt and R. Kattenbach, "Investigation of different fading forecast schemes for flat fading radio channels," in *Proc. IEEE VTC*, 2003-Fall, pp. 149–153.

[5] T. Ekman, "Prediction of mobile radio channels: modeling and design," Ph.D. dissertation, Uppsala University, 2002.

[6] K. E. Baddour and N. C. Beaulieu, "Improved pilot-assisted prediction of unknown time-selective Rayleigh channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2006, pp. 5192–5199.

[7] X. Cai and G. B. Giannakis, "Adaptive PSAM accounting for channel estimation and prediction errors," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 246–256, Jan. 2005.

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

[9] L. Lindobm, A. Ahlen, M. Sternad, and M. Falkenstrom, "Tracking of time-varying mobile radio channels - part II: a case study," *IEEE Trans. Commun.*, vol. 50, pp. 156–167, Jan. 2002.


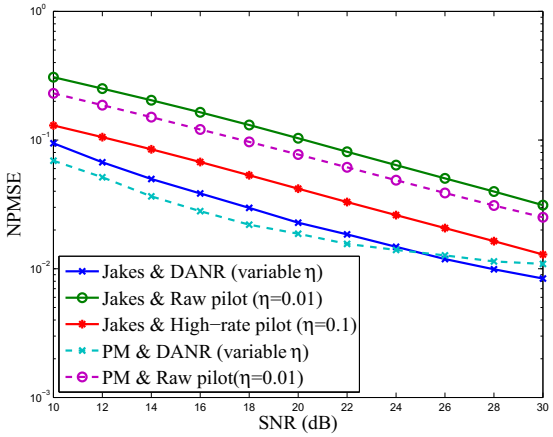


Fig. 7: NPMSE vs. SNR, prediction range of $0.2\lambda$; Jakes and physical model (PM)

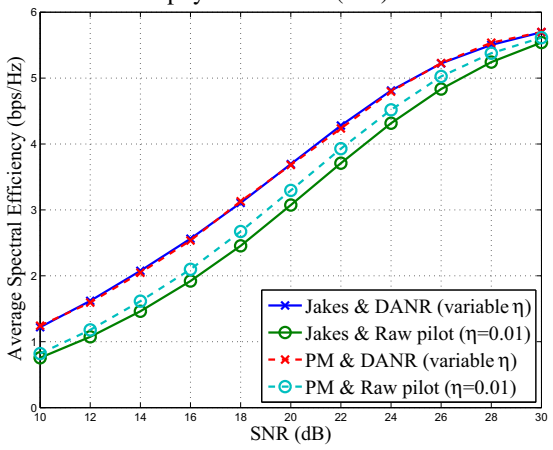


Fig. 8: Spectral efficiency vs. SNR, prediction range of $0.2\lambda$; Jakes and physical model (PM)


[10] D. Schafhuber and G. Matz, "MMSE and adaptive prediction of time-varying channels for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 593–602, Mar. 2005.

[11] T. Eyceoz, S. Hu, and A. Duel-Hallen, "Performance analysis of long range prediction for fast fading channels," in *Proc. 33rd Annual Conference on Information Sciences and Systems (CISS'99)*, 1999, pp. 656–661.

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

[13] S. Zhou and 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, pp. 1285–1294, May 2004.

[14] S. Falahati, A. Svensson, T. Ekman, and M. Sternad, "Adaptive modulation systems for predicted wireless channels," *IEEE Trans. Commun.*, vol. 52, pp. 307–316, Feb. 2004.

[15] T. Jia, "Single and multicarrier adaptive transmission systems with long-range prediction aided by noise reduction," Ph.D. dissertation, NC State University, Raleigh, NC, in preparation.

[16] C. H. Aldana and J. Cioffi, "Channel tracking for multiple input, single output systems using EM algorithm," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2001, pp. 586–590.

[17] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.

[18] 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," in *Proc. IEEE MILCOM*, Oct. 2002, pp. 627–631.

[19] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 844–854, July 1999.