Autoregressive Time-Varying Flat-Fading Channels: Model Order and Information Rate Bounds

Parastoo Sadeghi
The Australian National University
Canberra 0200 ACT Australia
Email: parastoo.sadeghi@anu.edu.au

Predrag Rapajic
University of Greenwich at Medway
KENT ME4 4TB United Kingdom
Email: p.rapajic@greenwhich.ac.uk

Rodney Kennedy and Thushara Abhayapala
The Australian National University
Canberra 0200 ACT Australia
Email: firstname.lastname@anu.edu.au

***Abstract*— In this paper, we study the effect of channel memory order on the information rate bounds in time-varying flat-fading (FF) channels. We model time variations of the FF channel with autoregressive (AR) processes with varying degrees of model order. We observe that in high SNR conditions (SNR $\gtrsim$ 20 dB), the information rate penalty of not knowing the AR channel is a non-increasing function of the AR model order. This is expected, since the AR channel predictability cannot decrease with increasing its order. However, in low SNR conditions, the information rate penalty in low-order AR channels can be lower than those in high-order AR channels. Likewise, the intuitive and universal monotonic increase of the information rate bounds with the AR model order is only observed in almost noiseless conditions. In the low SNR regime, however, the achievable information rate bounds in low-order AR channels can be higher than those in high-order AR channels.**

I. Introduction

A. Motivation and Background

The information capacity of time-varying fading channels has been recently investigated in the literature due to its practical significance in mobile communication systems [1]–[3]. The main goal of this paper is to study the effect of channel memory order on the achievable information rates in time-varying fading channels. The time-varying channel is assumed to be a random process, whose realization is unknown at the transmitter and at the receiver sides. The memory order of a random process is the number of previous realizations of the process that statistically determine its current realization. A well-known model to represent a time-varying channel process is the autoregressive (AR) model [2], [4], [5]. In an AR channel model of order P, hereafter denoted by AR(P), the current channel realization is determined from P previous channel realizations and a sample from a white random process with variance σ_P^2.

A smaller variance σ_P^2 in an AR(P) process signifies a higher predictability of the current realization of the process from its P previous realizations. From the theory of AR models, it is known that σ_P^2 is a non-increasing function of the model order P [6, p. 600]. That is, for a given correlation function for the process, an AR($P+1$) process is at least as predictable as an AR(P) process. In the theoretical limit, an AR(∞) process with a bandlimited spectrum is *deterministic*, in the sense that its current realization can be predicted, with no error, given its infinite past [6, p. 600], [7].

In this paper, we aim to investigate whether the non-decreasing predictability of AR time-varying channels with the AR model order has a non-decreasing effect on the achievable information rates through the channel. From an engineering point of view, *estimation* and *tracking* of a theoretically more predictable channel could potentially be more accurate, and hence, higher information rates should be achievable at the same noisy channel observation conditions.

In an earlier work by the authors [8], it was shown that decreasing the Markov channel memory order does not necessarily decrease the capacity of the finite-state Markov channel (FSMC) [9], [10]. The capacity comparison of FSMC models with different orders showed that when the FSMC states are deeply hidden in noise, a high-order FSMC usually has a lower capacity than its low-order FSMC counterpart. The intuitive capacity increase with increasing channel memory order only happens when the FSMC states are *highly observable* at the receiver. Whether a similar phenomenon happens in other time-varying channels deserves further analysis.

B. Approach and Contributions

Unlike [8], which considered finite-level partitioning of the flat-fading (FF) channel gain into FSMC models, we consider continuous-level FF channels. Therefore, the findings in this paper cannot be attributed to the artifacts of finite-level presentation of the FF channel gain in FSMC models or the choice of FF channel gain partitioning thresholds. With no channel state information (CSI) assumption, the channel capacity and the capacity-achieving input distribution for time-varying FF channels are essentially open problems [4], [11], [12]. Hence, we apply the bounds in [12], [13] to our problem to study the achievable information rate bounds in AR-FF channels.

The contributions of this paper are summarized as follows.

1) In low SNR conditions, the information rate penalty due to not knowing a high-order AR channel is higher than the penalty due to not knowing its low-order AR channel counterpart (Fig. 1, SNR $\lesssim$ 7 dB). In fact, higher predictability of higher-order AR channels only has a non-increasing effect on the information rate penalty in very high SNR conditions.
2) The information rate bounds in AR channels also exhibit a non-monotonic behavior with the channel memory

order in the low SNR regime. Our analysis shows that in the low SNR regime, the achievable information rate bounds in low-order AR channels are often higher than those in high-order AR channels (Fig. 2, SNR $\lesssim$ 7 dB and Fig. 4, SNR $\lesssim$ 10 dB). The universal monotonic increase of the information rate bounds with AR model order and its predictability only happens in almost noiseless conditions.

II. System Model

With the FF channel assumption, the low-pass received signal at the discrete time index k is

$$y_k = h_k x_k + n_k, \tag{1}$$

where x_k is the transmitted signal, h_k is the complex-valued and Gaussian-distributed FF channel gain, and n_k is a sample of complex-valued, additive white Gaussian noise (AWGN). Furthermore, we assume that h_k has a zero mean and a normalized variance of 0.5 per dimension and n_k has a zero mean and variance of $N_0/2$ per dimension. The average transmitted power and SNR per symbol are denoted as $\mathcal{E}_s = E\{|x_k|^2\}$ and $\gamma_s = \mathcal{E}_s/N_0$, respectively.

We further assume that the time-varying FF channel gain h_k is related to its past realizations through the AR(P) model

$$h_k = -\sum_{p=1}^{P} a_p h_{k-p} + w_k, \tag{2}$$

where w_k is a complex-valued, white Gaussian noise with variance σ_P^2 and the vector $\mathbf{a} = \left[a_1\,, \cdots, a_P\right]^T$ contains the AR filter coefficients and is derived from the Yule-Walker equation [14, pp. 55-57]:

$$\mathbf{Ra} = -\mathbf{v}, \tag{3}$$

where $\mathbf{R}$ is the $P \times P$ Toeplitz covariance matrix of the process, whose element at row i and column j only depends on $|i-j|$ and is denoted as $R(|i-j|)$. The vector $\mathbf{v}$ is given as

$$\mathbf{v} = \left[R(1)\,, R(2)\,, \cdots, R(P)\right]^T. \tag{4}$$

If we assume that the FF channel follows the Clarke's model [15, Ch. 14], then $R(|i-j|)$ is given as

$$R_{\text{Clarke}}(|i-j|) = J_0(2\pi f_D |i-j|), \tag{5}$$

where J_0 is the zero-order Bessel function of the first kind and f_D is the Doppler frequency shift that is normalized by the transmitted symbol period. Alternatively, for the ease of spectral analysis in Section IV, we may assume that $R(|i-j|)$ is given as

$$R_{\text{U}}(|i-j|) = \text{sinc}(2\pi f_D |i-j|), \tag{6}$$

which is, in fact, equivalent to 'the Clarke's model' for 3D isotropic scattering. In an AR(P), the complete set of covariance elements $R_h(n)$ is given as [16]

$$R_h(n) = \begin{cases} R(n) & |n| \leqslant P \\ -\sum_{p=1}^{P} a_p R_h(n-p) & |n| > P \end{cases}, \tag{7}$$

where $R(n)$ for $|n| \leqslant P$ was defined in (5) or (6). The power spectral density (PSD) of AR(P) is a rational function

$$H_P(\omega) = \frac{\sigma_P^2}{\left|1 + \sum_{p=1}^{P} a_p e^{-jp\omega}\right|^2}. \tag{8}$$

For the special case of AR(∞), we obtain the well-known bowl-shaped PSD for the Clarke's model, which is given as

$$H_{\text{Clarke}}(\omega) = \begin{cases} \frac{1}{\pi f_D \sqrt{1-(\omega/2\pi f_D)^2}} & |\omega| < 2\pi f_D \\ 0 & \text{otherwise} \end{cases}, \tag{9}$$

and a uniform PSD for R_{U} in (6) as

$$H_{\text{U}}(\omega) = \begin{cases} \frac{1}{2 f_D} & |\omega| < 2\pi f_D \\ 0 & \text{otherwise} \end{cases}. \tag{10}$$

The variance σ_P^2 of the white noise w_k in (2) is the minimum mean square error (MMSE) of the one-step predictor of order P for the fading process h_k [6, pp. 598-601] and is given as

$$\sigma_P^2 = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln H_P(\omega) d\omega\right), \tag{11}$$

where $\ln$ stands for natural logarithm. It is known that the MMSE is a non-increasing function of the prediction order P [6, p. 600]. That is,

$$\sigma_{P+1}^2 \leqslant \sigma_P^2, \quad \forall P. \tag{12}$$

In other words, the AR(P) channel is at most as equally *predictable* as the AR($P+1$) channel. From (11), we observe that if the asymptotic PSD of the fading process is zero over some frequency interval, the asymptotic MMSE prediction error is zero. In other words, a bandlimited fading process (as in (9) or (10)) is, in theory, *deterministic* in the sense that its future realization can be predicted from its infinite past with no error [7].

Now, reconsider the unknown FF channel with observation equation (1), where the FF channel gain varies with time according to the AR(P) model in (2). From the predictability discussion following (11)-(12), one might be inclined to conclude that the achievable information rates in an AR(P) channel are generally lower than those in an AR($P+1$) channel. In the limit, it might be even conjectured that an AR(∞) model with the highest predictability has the highest achievable information rate among its lower-order and less predictable AR counterparts. However, the predictability analysis in (11)-(12) lies on the important assumption that the channel's infinite past is already observed without error and is available. This is equivalent to assuming a noiseless channel observation equation in (1) and a known transmitted signal ($N_0 = 0$ and known x_k). The analysis in Section IV reveals, however, that in low to medium SNR conditions, when the channel realizations are not directly observable, the effect of increasing channel memory order on the information rate bounds is not monotonic with the AR model order.

III. Information Rate Bounds

The information capacity and the capacity-achieving input distribution of temporally correlated, time-varying FF channels with no CSI at the transmitter and at the receiver are generally open problems [3], [4], [11], [12]. Therefore, we limit ourselves to analyze and compare the information rate upper and lower bounds in correlated FF channels with varying degrees of AR model order. We start by briefly reviewing the information rate bounds in [12]. The advantage of these bounds is that they have closed-form and single-letter expressions and provide valuable insight into our problem. Later in Section IV, we will verify our observations by using tighter bounds in [13], which can only be numerically computed. All mutual information rates are computed in nats per channel use.

Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{h}$ be $N \times 1$ sequences of transmitted signal, received signal, and channel realization, respectively. Using the chain rule on the mutual information $I(\mathbf{y};\mathbf{x},\mathbf{h})$, we can write the mutual information between $\mathbf{y}$ and $\mathbf{x}$ as

$$I(\mathbf{y};\mathbf{x}) = I(\mathbf{y};\mathbf{x},\mathbf{h}) - I(\mathbf{y};\mathbf{h}|\mathbf{x}) \tag{13}$$

$$= I(\mathbf{y};\mathbf{h}) + I(\mathbf{y};\mathbf{x}|\mathbf{h}) - I(\mathbf{y};\mathbf{h}|\mathbf{x}), \tag{14}$$

where $I(\mathbf{y};\mathbf{x}|\mathbf{h})$ is the mutual information with perfect CSI and $I(\mathbf{y};\mathbf{h}|\mathbf{x}) - I(\mathbf{y};\mathbf{h})$ is the penalty in mutual information when the CSI is not available. Since $I(\mathbf{y};\mathbf{h}) \geqslant 0$, it is concluded that $I(\mathbf{y};\mathbf{h}|\mathbf{x})$ is an upper bound on the mutual information penalty due to unknown CSI. An upper bound on $I(\mathbf{y};\mathbf{h}|\mathbf{x})$ can be derived as (see [12] for more details)

$$I(\mathbf{y};\mathbf{h}|\mathbf{x}) \leqslant \sum_{n=1}^{N} \ln\Big(1 + \frac{\mathcal{E}_s}{N_0}\lambda_n\Big), \tag{15}$$

where $\{\lambda_n\}_{n=1}^{N}$ are the eigenvalues of the $N \times N$ Toeplitz covariance matrix of channel vector $\mathbf{h}$ with elements defined in (7). The asymptotic information rate per symbol for $N \to \infty$ is then upper bounded as

$$\Delta_P \triangleq \mathcal{I}_P(y;h|x) = \lim_{N\to\infty} \frac{1}{N} I(\mathbf{y};\mathbf{h}|\mathbf{x}) \tag{16}$$

$$\leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\Big(1 + \frac{\mathcal{E}_s}{N_0}H_P(\omega)\Big)d\omega. \tag{17}$$

The equality in (17) is achieved for constant power signaling such as M-ary PSK schemes [15, Ch. 4-5]. Therefore, it is possible to upper bound $I(\mathbf{y};\mathbf{x})$ for constant power signaling by using (13) and upper bounding $I(\mathbf{y};\mathbf{h},\mathbf{x}) = H(\mathbf{y}) - H(\mathbf{y}|\mathbf{h},\mathbf{x}) = H(\mathbf{y}) - H(\mathbf{n})$ by assuming a Gaussian distribution for $\mathbf{y}$. The following asymptotic information rate upper bound is derived for constant power signaling

$$\mathcal{I}_P^C(y;x) \leqslant \ln\Big(1 + \frac{\mathcal{E}_s}{N_0}\Big) - \Delta_P \tag{18}$$

$$= \ln\Big(1 + \frac{\mathcal{E}_s}{N_0}\Big) - \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\Big(1 + \frac{\mathcal{E}_s}{N_0}H_P(\omega)\Big)d\omega.$$

Using the fact that $I(\mathbf{y};\mathbf{h})$ is non-negative in (14), we obtain the following lower bound on $I(\mathbf{y};\mathbf{x})$

$$I(\mathbf{y};\mathbf{x}) \geqslant I(\mathbf{y};\mathbf{x}|\mathbf{h}) - I(\mathbf{y};\mathbf{h}|\mathbf{x}). \tag{19}$$

In order to evaluate the asymptotic information rate per symbol, we use (16)-(17) and the fact that the mutual information rate with CSI is independent of time index. This yields

$$\mathcal{I}(y;x|h) = H(y|h) - H(y|x,h) = H(y|h) - H(n) \tag{20}$$

$$= H(y|h) - \ln(\pi e N_0),$$

and therefore,

$$\mathcal{I}(y;x) \geqslant \mathcal{I}(y;x|h) - \Delta_P \tag{21}$$

$$\geqslant \mathcal{I}(y;x|h) - \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\Big(1 + \frac{\mathcal{E}_s}{N_0}H_P(\omega)\Big)d\omega.$$

For M-ary PSK signaling, the CSI entropy rate $H(y|h)$ only depends on the channel gain amplitude $r = |h|$ and is given as

$$H(y|h) = \int_0^{\infty}\int_{-\infty}^{\infty} f(y|r)f(r)\ln f(y|r)dydr, \tag{22}$$

where $f(y|r)$ is the conditional pdf of y given r and r is Rayleigh-distributed. For equiprobable, constant power, M-ary PSK signaling $f(y|r)$ is given by

$$f(y|r) = \frac{1}{M}\sum_{m=0}^{M-1} f(y|x_m, r), \tag{23}$$

where $f(y|x_m, r)$ is the pdf of AWGN in (1) with mean $rx_m = r\sqrt{\mathcal{E}_s}e^{j2\pi m/M}$ and variance $N_0/2$ per dimension.

IV. Analysis of Information Rate Bounds

We first examine the information rate upper and lower bounds in (18) and (21) in more detail. We observe that the first terms in (18) and (21) are independent of the dynamics of time-varying AR channel and only depend on SNR. On the other hand, the second term Δ_P, given in (17) depends on both the AR(P) spectrum $H_P(\omega)$ and SNR $\mathcal{E}_s/N_0$.

The information rate penalty Δ_P for the special case of AR(1) has a closed form. To see this, we use (3) to rewrite the AR spectrum in (8) as

$$H_1(\omega) = \frac{\sigma_1^2}{\left|1 + a_1 e^{-j\omega}\right|^2} \tag{24}$$

$$= \frac{1 - R^2(1)}{1 + R^2(1) - 2R(1)\cos\omega}. \tag{25}$$

Upon using (25) in (17) and the integral identity [17, p. 526]

$$\int_{-\pi}^{\pi} \ln(1 + a\cos x)dx = 2\pi \ln\frac{1 + \sqrt{1 - a^2}}{2}, \tag{26}$$

we obtain

$$\Delta_1 = \ln\frac{A + \sqrt{A^2 - B^2}}{C + \sqrt{C^2 - B^2}}, \tag{27}$$

where $A \triangleq 1 + R^2(1) + \mathcal{E}_s/N_0(1 - R^2(1))$, $B \triangleq 2R(1)$, and $C \triangleq 1 + R^2(1)$.

The information rate penalty Δ_∞ for AR(∞) with the uniform spectrum (10) is simply derived to be

$$\Delta_\infty = 2f_D \ln\Big(1 + \frac{\mathcal{E}_s}{N_0}\frac{1}{2f_D}\Big). \tag{28}$$

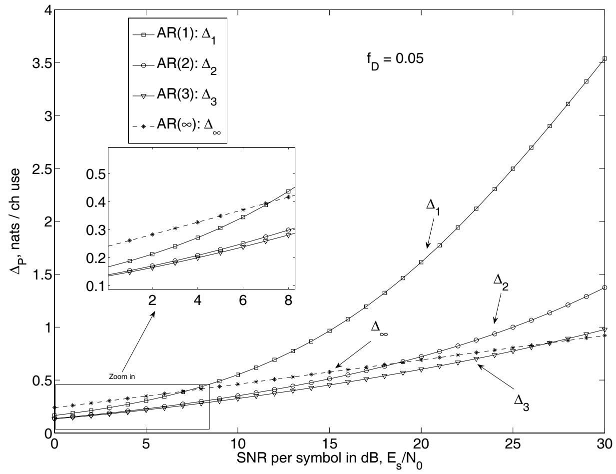

Fig. 1. The information rate penalty due to unknown CSI Δ_P, given in (17), as a function of SNR and the AR channel model order P. As shown in the enlarged graphs, for SNR $\lesssim 7$ dB, the penalty in information rate due to unknown CSI is highest in the most "ideally predictable" channel model AR(∞). In fact, higher channel predictability has the expected monotonically decreasing effect on the information rate penalty only in very high SNR (ideally noiseless) conditions. This is observed for SNR $\gtrsim 27$ dB.

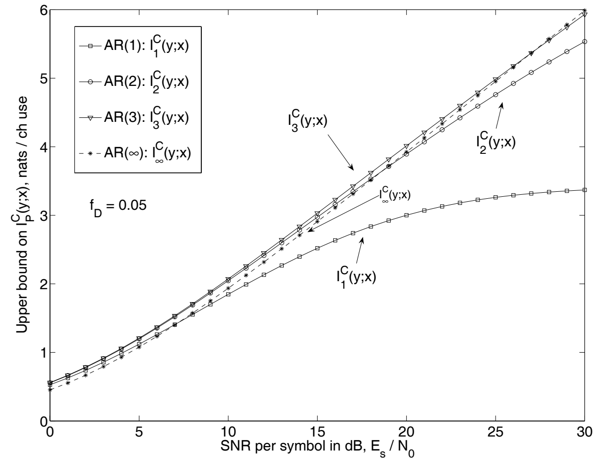

Fig. 2. The non-monotonic behavior of the information rate penalty with the AR model order in Fig. 1 is translated into the non-monotonic behavior of information rate upper bounds for constant power signaling. For example, the information rate upper bound in the AR(∞) channel, $\mathcal{I}_\infty^C(y;x)$ is lower than $\mathcal{I}_3^C(y;x)$ in the AR(3) channel for SNR conditions $\mathcal{E}_s/N_0 \lesssim 27$ dB.

The information rate penalty Δ_P for $P > 1$ is numerically evaluated by directly using (8) in (17).

Fig. 1 shows the information rate penalty Δ_P at the normalized fading rate of $f_D = 0.05$. The elements of $\mathbf{R}$ in (3) are chosen from (6) (uniform fading spectrum). Four different AR models with $P = 1$ to $P = 3$ and $P = \infty$ are shown. From this figure, it is clearly observed that the information rate penalty Δ_P due to unknown CSI does not necessarily decrease with increasing the channel memory order (increasing channel predictability). For example, the penalty in the achievable information rate in an AR(∞) channel, Δ_∞, is larger than the penalty in the achievable information rate in an AR(1) channel, Δ_1, for $\mathcal{E}_s/N_0 \lesssim 7$ dB. This is despite the fact that AR(1) is the least predictable channel according to (12). Similarly, $\Delta_\infty > \Delta_3$ for $\mathcal{E}_s/N_0 \lesssim 27$ dB. Only in very high SNR conditions, do we observe a monotonically decreasing information rate penalty with increasing P. In fact, this phenomenon can be qualitatively explained by referring to the integrand in (17). In very high SNR conditions, the constant term 1 in the logarithm is negligible, compared to the second term and may be ignored. In this case, the behavior of Δ_P with memory order would be very much like the behavior of σ_P^2 in (11), which is non-increasing as P grows large. However, in low to medium SNR conditions, the constant term in the integrand in (17) may not be ignored and the real behavior of Δ_P depends on SNR and AR(P) spectrum. The corresponding information rate upper bounds for constant power signaling, $\mathcal{I}_P^C$ in (18), are shown in Fig. 2. It is verified that the non-monotonic behavior of Δ_P translates into non-monotonic information rate upper bounds. For example, the information rate upper bound in the AR(3) channel is higher than that in the AR(∞) channel for $\mathcal{E}_s/N_0 \lesssim 27$ dB.

In Fig. 3, we have shown the information rate upper bounds for constant power signaling, $\mathcal{I}_P^C$ in (18), as a function of the normalized Doppler frequency shift f_D for two low SNR conditions of $\mathcal{E}_s/N_0 = 0$ and $\mathcal{E}_s/N_0 = 3$ dB and two AR models AR(1) and AR(∞). The elements of $\mathbf{R}$ in (3) are chosen from (6) (uniform fading spectrum). $\mathcal{I}_1^C$ and $\mathcal{I}_\infty^C$ both have closed-form expressions according to (18) and (27)-(28). From this figure it is observed that $\mathcal{I}_1^C$ in AR(1) is higher than $\mathcal{I}_\infty^C$ in AR(∞) for slow to medium-speed fading rates. However, for very fast fading rates $\mathcal{I}_\infty^C$ is higher. It is also verified from this figure that $\mathcal{I}_\infty^C$ is higher than $\mathcal{I}_1^C$ for a wider range of fading rates at higher SNR conditions. As expected, when the SNR is sufficiently high ($\mathcal{E}_s/N_0 \gtrsim 10$ dB), $\mathcal{I}_\infty^C$ is universally higher than $\mathcal{I}_1^C$ for almost all fading rates. This is not shown in Fig. 3 for more clarity of the figure.

In Fig. 4, we have shown the information rate lower bounds in (21) by assuming QPSK constant power signaling. The normalized fading rate is $f_D = 0.01$. The bounds tend to become loose in high SNR conditions. The reason is that in (21), the first term is specifically calculated for QPSK and hence, $\mathcal{I}(y;x|h) \leqslant 2\ln 2 = 1.39$ nats/ch use, whereas the second term in (21) generally holds for all constant power signaling schemes and becomes large for high SNR. Nevertheless, the behavior of the lower bound at low SNR conditions is noteworthy. In fact, the information lower bounds in AR(1) to AR(3) models are all higher than that of AR(∞) for $\mathcal{E}_s/N_0 \lesssim 10$ dB. At higher SNR conditions, the information rate lower bound in AR(3) is still higher than that in AR(∞).

We have also computed the lower and upper bounds in [13] for two AR(1) and AR(∞) channels. The bounds are tighter than those in [12], but can only be numerically computed. The upper bound in [13] can be computed for constant power

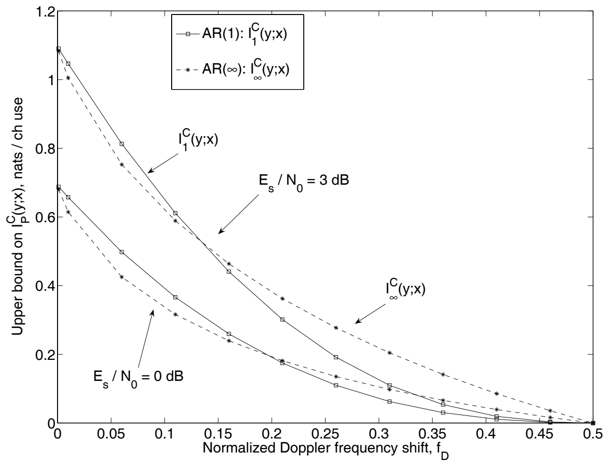

Fig. 3. The information rate upper bounds, $\mathcal{I}_P^C$ in (18), for constant power signaling as a function of the normalized Doppler frequency shift. In very fast fading rates, $\mathcal{I}_\infty^C$ is higher than $\mathcal{I}_1^C$. In relatively slow to medium-speed fading conditions, $\mathcal{I}_1^C$ is higher. The transition point depends on the SNR. In lower SNR conditions, $\mathcal{I}_1^C > \mathcal{I}_\infty^C$ for a wider range of fading rates.

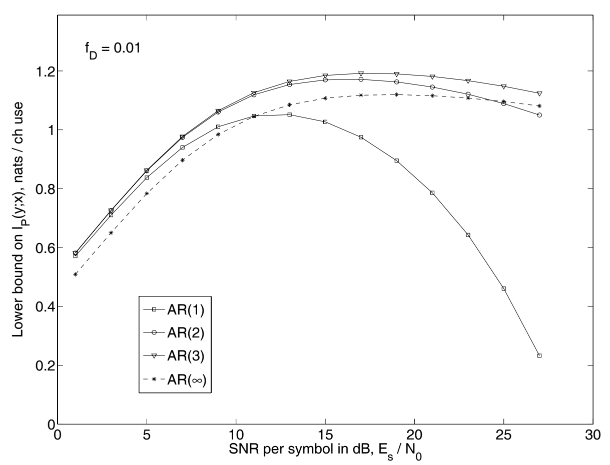

Fig. 4. The information rate lower bounds for QPSK signaling, given in (21), as a function of SNR. The information rate lower bounds in AR(1) to AR(3) channels are all higher than that in the AR(∞) channel for $\mathcal{E}_s/N_0 \lesssim 10$ dB. The information rate lower bound in the AR(3) channel is higher than that in the AR(∞) channel for all considered SNR values.

signaling, since $H_P(y|x) = \mathcal{I}_P(y;h|x) + H_P(y|x,h) = \Delta_P + H(n) = \Delta_P + \ln(\pi e N_0)$. For the fading rate $f_D = 0.05$, BPSK scheme, and SNR of 0 dB, the information rate lower bound in [13] for the AR(1) channel yields $\underline{I}_1 = 0.3062$ nats/ch use, whereas for the AR(∞) channel $\underline{I}_\infty = 0.2361$ nats/ch use. The upper bound for the AR(1) channel yields $\overline{I}_1 = 0.3830$ nats/ch use, whereas for the AR(∞) channel $\overline{I}_\infty = 0.3308$ nats/ch use. The upper bounds in Fig. 2 for this SNR were $I_1^C = 0.5275$ and $I_\infty^C = 0.4534$ nats/ch use, respectively. This confirms the non-increasing behavior of information rates with increasing the AR model order at low SNR values.

We also observed that similar conclusions as in Figs. 1-4 can be drawn by using the Clarke's spectrum in (9), instead of (10). Therefore, the numerical results are omitted here.

V. Conclusions

The theoretical higher channel predictability of high-order AR channel models is only well-defined in almost noiseless conditions (high SNR regime). In this case, one can expect that the achievable information rates in time-varying channels are non-decreasing functions of the channel memory order. On the other hand, in the low SNR regime, higher channel memory order and ideal channel predictability do not affect the achievable information rates as expected. Our analysis showed that in many low SNR conditions and for a wide range of channel fading rates, the information rate bounds in low-order AR channels are higher than those in high-order AR channels. Our analysis is in accordance with the findings in [8] on the non-monotonic effect of memory order on the FSMC capacity.

References

[1] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[2] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[3] Y. Liang and V. V. Veeravalli, "Capacity of noncoherent time-selective Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3095–3110, Dec. 2004.

[4] J. Zhang;, "On the bounds of the non-coherent capacity of Gauss-Markov fading channels," in *Proc. IEEE Int. Conf. on Acoustic, Speech, and Signal Processing (ICASSP)*, Mar. 2005, pp. 757–760.

[5] R. Chen, B. Hajek, R. Koetter, and U. Madhow, "On fixed input distributions for noncoherent communication over high-SNR Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3390–3396, Dec. 2004.

[6] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. New York: Mc Graw Hill, 2002.

[7] K. E. Baddour and N. C. Beaulieu, "Autoregressive models for fading channel simulation," in *IEEE Global commun. Conf. (GLOBECOM)*, San Antonio, TX, Nov. 2001, pp. 1187–1192.

[8] P. Sadeghi, P. B. Rapajic, and Z. B. Krusevac, "The effect of memory order on the capacity of finite-state markov and flat-fading channels," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, Adelaide, Australia, Sept. 2005, pp. 525–529.

[9] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[10] A. J. Goldsmith and P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 868–886, May 1996.

[11] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[12] X. Deng and A. M. Haimovich, "Information rates of time varying Rayleigh fading channels," in *IEEE Int. Conf. Commun.(ICC)*, Paris, France, June 2004, pp. 573–577.

[13] D. Arnold, H. A. Loeliger, and P. O. Vontobel, "Computation of information rates from finite-state source/channel models," in *Proc. 40th Allerton Conf. on Commun., Control, and Computing*, Monticello, Illinois, Oct. 2002.

[14] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River: Prentice Hall, 2002.

[15] J. G. Proakis, *Digital Communications*, 4th ed. New York: Mc Graw Hill, 2000.

[16] S. M. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[17] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed. San Diego: Academic Press, 2000.