Pilot Assisted Estimation of MIMO Fading Channel Response and Achievable Data Rates *

Dragan Samardzija
Wireless Research Laboratory,
Bell Labs, Lucent Technologies,
791 Holmdel-Keyport Road,
Holmdel, NJ 07733, USA
dragan@bell-labs.com

Narayan Mandayam
WINLAB, Rutgers University,
73 Brett Road,
Piscataway NJ 08854, USA
narayan@winlab.rutgers.edu

Abstract

We analyze the effects of pilot assisted channel estimation on achievable data rates (lower bound on information capacity) over a frequency flat time-varying channel. Under a block-fading channel model, the effects of the estimation error are evaluated in the case of the estimates being available at the receiver only (open loop), and in the case when the estimates are fed back to the transmitter allowing water pouring transmitter optimization (closed loop). Using a characterization of the effective noise due to estimation error, we analyze the achievable rates as a function of the power allocated to the pilot, the channel coherence time, the background noise level as well as the number of transmit and receive antennas. The analysis presented here can be used to optimally allocate pilot power for various system and channel operating conditions, and to also determine the effectiveness of closed loop feedback.

***Keywords*: Estimation, Ergodic Capacity, Water Pouring, MIMO.**

*This paper was presented, in part, at the DIMACS Workshop on Signal Processing for Wireless Transmission, Rutgers University, October 2002 . This work is supported in part by the NJ Commission on Science and Technology under the NJCWT program.

1 Introduction

Fading channels are an important element of any wireless propagation environment [1]. Different aspects of fading channels have been studied and publicized. It has been recognized that the inherent temporal and spatial variations of wireless channels impose stringent demand on design of a communication system to allow it approach the data rates that are achievable in, for example, wire-line systems. A number of different solutions exploit variations in wireless channels. For example, a transmitter optimization scheme (using power control), known as the water pouring algorithm, maximizes the capacity for the constrained average transmit power [2] (see also [3]). In addition to power control, recent applications of variable coding rate and modulation formats illustrate a wide range of resource allocation techniques used to exploit and combat effects of fading channels in multiuser wireless systems [4]. An extensive review of the information theoretical aspects of communications in fading channels is given in [5]. Furthermore, modulation and channel coding for fading channels is also being studied (see [5] and references therein).

Multiple-transmit multiple-receive antenna systems represent an implementation of the MIMO concept in wireless communications [6]. This particular multiple antenna architecture provides high capacity wireless communications in rich scattering environments. It has been shown that the theoretical capacity (approximately) increases linearly as the number of antennas is increased [6,7]. This and related results point to the importance of understanding all aspects of MIMO wireless systems. For example, the studies regarding propagation [8–10], detection [11, 12], space-time coding and implementation aspects [13–15] are well publicized.

Next generation wireless systems and standards are supposed to operate over wireless channels whose variations are faster and/or further pronounced. For example, using higher carrier frequencies (e.g., 5 GHz for 802.11a) results in smaller scale of spatial variations of the electro-magnetic field. Also, compared to SISO channels, MIMO channels have greater number of parameters that a receiver and/or transmitter has to operate with, consequently pronouncing the channel variations. In addition, there has been a perpetual need for supporting higher mobility within wireless networks. These are just a few motivations for studying the implications of channel variations on

achievable data rates in wireless systems.

In this paper we analyze how the estimation error of the channel response affects the performance of a MIMO wireless system. Considering the practical importance of single-input single-output (SISO) systems, we analyze them as a subset of MIMO systems. Considering terminology in literature (see [5] and references therein), the channel response estimate corresponds to *channel state information* (CSI). We assume a frequency-flat time-varying wireless channel with additive white Gaussian noise (AWGN). More precisely, a quasi-static block-fading channel model is used. Furthermore, the temporal variations of the channel are characterized by the correlation between successive channel blocks. The above system may also correspond to one subchannel (i.e., carrier) of an OFDM wireless system [16]. We consider two pilot (training) arrangement schemes in this paper. The first scheme uses a single pilot symbol per block with the different power than the data symbol power. The second scheme uses more than one pilot symbol per block, whose power is the same as the data symbol power. For the given pilot schemes, in both cases, *maximum-likelihood* (ML) estimation of the channel response is considered [17]. In the MIMO case, the orthogonality between the pilots assigned to different transmit antennas is assumed. The effects of the estimation error are evaluated in the case of the estimates being available at the receiver only, and in the case when the estimates are fed back to the transmitter allowing water pouring optimization. The presented analysis may be viewed as a study of mismatched receiver and transmitter algorithms in MIMO systems. The analysis connects results of information theory (see [18, 19] and references therein) with practical wireless communication systems (employing pilot assisted channel estimation) and generalizing it to MIMO systems. Previously published studies on MIMO channel estimation and its effects include [20] and [21]. An elaborate information-theoretical study analyzing different training schemes, and optimizing their parameters to maximize the open loop MIMO capacity lower bounds, is also presented in [22]. We will highlight the similarities and differences of the work presented here to that in [22] in the subsequent sections of this paper. We believe that the results presented here are directly applicable to current and next generation wireless systems [13–15, 23]. Furthermore, the results may be used as baseline benchmarks for

performance evaluation of more advanced estimation and transmitter optimization schemes, such as anticipated in future systems.

2 System Model

In the following we present a MIMO communication system that consists on M transmit and N receive antennas (denoted as a $M \times N$ system). At the receiver we assume sampling with the period $T_{smp} = 1/B$, where B is the signal bandwidth, thus preserving the sufficient statistics. The received signal is a spatial vector $\mathbf{y}$

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{n}(k), \qquad \mathbf{y}(k) \in \mathcal{C}^N, \mathbf{x}(k) \in \mathcal{C}^M, \mathbf{n}(k) \in \mathcal{C}^N, \mathbf{H}(k) \in \mathcal{C}^{N \times M} \tag{1}$$

where $\mathbf{x}(k) = [g_1(k) \cdots g_M(k)]^{\mathrm{T}}$ is the transmitted vector, $\mathbf{n}(k) = [n_1(k) \cdots n_N(k)]^{\mathrm{T}}$ is the AWGN vector with $(\mathrm{E}[\mathbf{n}(k)\mathbf{n}(k)^{\mathrm{H}}] = N_0\, \mathbf{I}_{N \times N})$, and $\mathbf{H}(k)$ is the MIMO channel response matrix, all corresponding to the time instance k. We assign index $m = 1, \cdots, M$ to denote the transmit antennas, and index $n = 1, \cdots, N$ to denote the receive antennas. Thus, $h_{nm}(k)$ is the n-th row and m-th column element of the matrix $\mathbf{H}(k)$. Note that it corresponds to a SISO channel response between the transmit antenna m and the receive antenna n. $g_m(k)$ is the transmitted signal from the m-th transmit antenna The n-th component of the received spatial vector $\mathbf{y}(k) = [y_1(k) \cdots y_N(k)]^{\mathrm{T}}$ (i.e., signal at the receive antenna n) is

$$y_n(k) = \sum_{m=1}^{M} h_{nm}(k) g_m(k) + n_n(k). \tag{2}$$

To perform estimation of the channel response $\mathbf{H}(k)$, the receiver uses a pilot (training) signal that is a part of the transmitted data. The pilot is sent periodically, every K sample periods. We consider the transmitted signal to be comprised of two parts: one is the data bearing signal and the other is the pilot signal. Within the pilot period consisting of K symbols, L symbols (i.e., signal dimensions) are allocated to the pilot, per transmit antenna. As a common practical solution (see [13–15, 24]), we assume that the pilot signals assigned to the different transmit antennas, are mutually orthogonal. For more details on signal design for multiple transmit antenna systems see

also [25, 26]. This assumption requires that $K \geq LM$. Consequently we define a K-dimensional temporal vector $\mathbf{g}_m = [g_m(1) \cdots g_m(K)]^{\mathrm{T}}$, whose k-th component is $g_m(k)$ (in (2)), as

$$\mathbf{g}_m = \underbrace{\sum_{i=1}^{K-LM} a_{im}^d d_{im}^d \mathbf{s}_i^d}_{Data} + \underbrace{\sum_{j=1}^{L} a_{jm}^p d_{jm}^p \mathbf{s}_{jm}^p}_{Pilot}. \tag{3}$$

In the above the first sum is the information, i.e., data bearing signal and the second corresponds to the pilot signal, corresponding to the transmit antenna m. Superscripts "d" and "p" denote values assigned to the data and pilot, respectively. d_{im}^d is the unit-variance circularly symmetric complex data symbol. The pilot symbols (d_{jm}^p, $j = 1, \cdots, L$) are predefined and known at the receiver. Without loss of generality, we assume that $|d_{jm}^p|^2 = 1$. We also assume that the amplitudes are $a_{im}^d = A$, and $a_{jm}^p = A_P$, and they are known at the receiver. Further, the amplitudes are related as $A_P = \alpha A$. Note that the amplitudes are identical across the transmit antennas (because we assumed that the transmit power is equally distributed across them).

Furthermore, $\mathbf{s}_i^d = [s_i^d(1) \cdots s_i^d(K)]^{\mathrm{T}}, (i = 1, \cdots, (K - LM))$ and $\mathbf{s}_{jm}^p = [s_{jm}^p(1) \cdots s_{jm}^p(K)]^{\mathrm{T}}$, ($j = 1, \cdots, L$, and $m = 1, \ldots M$) are waveforms, denoted as *temporal signatures*. The temporal signatures are mutually orthogonal. For example, $\mathbf{s}_i^d$ (or $\mathbf{s}_{jm}^p$) could be a canonical waveform such as a TDMA-like waveform, where $\mathbf{s}_i^d$ (or $\mathbf{s}_{jm}^p$) is the unit-pulse at the time instance i. Alternately, $\mathbf{s}_i^d$ (or $\mathbf{s}_{jm}^p$) could also be a K-dimensional CDMA sequence spanning all K sample intervals [23]. Note that the above model while being general enough is particularly suitable for MIMO implementations over CDMA systems (see [24]).

As said earlier, we assume that the pilot signals are orthogonal between the transmit antennas. The indexing and summation limits in (3) conform to this assumption, i.e, temporal signatures $\mathbf{s}_{jm}^p (j = 1, \cdots, L)$ are uniquely assigned to the transmit antenna m. In other words, transmit antenna m must not use the temporal signatures that are assigned as pilots to other antennas and assigned to data, which is consequently lowering the achievable data rates (will be revisited in the following sections). Unlike the pilot temporal signatures, the data bearing temporal signatures $\mathbf{s}_i^d$ ($i = 1, \ldots, (K - LM)$) are reused across the transmit antennas, which is an inherent property of MIMO systems, potentially resulting in high achievable data rates. It is interesting to note that the

assumptions regarding the orthogonality between the pilots (motivated by practical considerations) are also shown to be optimal in [22], maximizing the open loop capacity lower bound. Similar conclusions are drawn in [20, 26]. We rewrite the received spatial vector in (1) as

$$\mathbf{y}(k) = \mathbf{H}(k)(\mathbf{d}(k) + \mathbf{p}(k)) + \mathbf{n}(k), \quad \mathbf{d}(k) \in \mathcal{C}^M, \mathbf{p}(k) \in \mathcal{C}^M \tag{4}$$

where $\mathbf{d}(k)$ is the information, i.e., data bearing signal and $\mathbf{p}(k)$ is the pilot portion of the transmitted spatial signal, at the time instance k. The m-th component of the data vector $\mathbf{d}(k) = [d_1(k) \cdots d_M(k)]^{\mathrm{T}}$ (i.e., data signal at the transmit antenna m) is

$$d_m(k) = \sum_{i=1}^{K-LM} a_{im}^d d_{im}^d s_i^d(k). \tag{5}$$

The m-th component of the pilot vector vector $\mathbf{p}(k) = [p_1(k) \cdots p_M(k)]^{\mathrm{T}}$ (i.e., pilot signal at the transmit antenna m) is

$$p_m(k) = \sum_{j=1}^{L} a_{jm}^p d_{jm}^p s_{jm}^p(k). \tag{6}$$

Let us now describe the assumed properties of the MIMO channel $\mathbf{H}(k)$. The channel coherence time is assumed to be greater or equal to KT_{smp}. This assumption approximates the channel to be constant over at least K samples ($h_{nm}(k) \approx h_{nm}$, for $k = 1, \cdots, K$, for all m and n), i.e., approximately constant during the pilot period. In the literature, channels with the above property are known as block-fading channels [16]. Furthermore, we assume that the elements of $\mathbf{H}$ are independent identically distributed (*iid*) random variables, corresponding to highly scattering channels. In general, the MIMO propagation measurements and modeling have shown that the elements of $\mathbf{H}$ are correlated (i.e., not independent) [8–10]. The effects of correlation on the capacity of MIMO systems is studied in [27]. Assuming independence is a common practice because the information about correlation is usually not available at the receiver and/or the correlation is time varying (not stationary) and hard to estimate. Based on the above, the received temporal vector at the receiver n, whose k-th component is $y_n(k)$ (in (2)), is

$$\mathbf{r}_n = [y_n(1) \cdots y_n(K)]^{\mathrm{T}} = \sum_{m=1}^{M} h_{nm}\mathbf{g}_m + \mathbf{n}_n, \;\; \mathbf{r}_n \in \mathcal{C}^K \tag{7}$$

where $\mathbf{n}_n = [n_n(1) \cdots n_n(K)]^{\mathrm{T}}$ and $\mathrm{E}[\mathbf{n}_n \mathbf{n}_n^{\mathrm{H}}] = N_0 \, \mathbf{I}_{K \times K}$.

Note that when applying different number of transmit antennas, the total average transmitted power must stay the same, i.e., conserved. This is a common assumption in MIMO systems [6,7]. Also, the power is equally distributed across the transmit antennas. The average transmit power (from all transmit antennas) is

$$P_{av} = M\frac{\left(\sum_{i=1}^{K-LM}(a_{im}^{d})^2 \; + \; \sum_{j=1}^{L}(a_{jm}^{p})^2\right)}{K} = M\frac{((K-LM)+L\alpha^2)A^2}{K}. \tag{8}$$

Thus

$$A = \sqrt{\frac{K}{((K-LM)+\alpha^2 L)}\frac{P_{av}}{M}}. \tag{9}$$

As seen from the above, we assume that the total average transmitted energy (within the pilot period) is the same, but differently distributed between the data bearing portion of the signal and the pilot. Consequently, we observe the performance of the system with respect to the amount of transmitted energy that is allocated to the pilot (percentage wise). This percentage is denoted as μ and is given as

$$\mu = \frac{L\alpha^2}{(K-LM)+L\alpha^2}100 \;\; [\%]. \tag{10}$$

As said earlier, in this study we consider two different pilot arrangements:

1. $L = 1$ and $A_P \neq A$. The amplitude is

$$A_1 = \sqrt{\frac{K}{((K-M)+\alpha^2)}\frac{P_{av}}{M}}. \tag{11}$$

   In the remainder of the paper, the above pilot arrangement is referred to as case 1. For example, in SISO systems the above pilot arrangement is applied in CDMA wireless systems (e.g., IS-95 and WCDMA [23]). In MIMO systems, it is applied in narrowband MIMO implementations described in [13–15]. It is also applied in a wideband MIMO implementation based on 3G WCDMA [24]. Note that under certain assumptions to be pointed out in the next section, the above pilot arrangement scheme is equivalent to the scheme in [22]

2. $L \geq 1$ and $A_P = A\,(\alpha = 1)$. The amplitude is

$$A_2 = \sqrt{\frac{K}{(K-L(M-1))}\frac{P_{av}}{M}}. \tag{12}$$

In the remainder of the paper, the above pilot arrangement is referred to as case 2. Note that the above pilot arrangement is frequently used in SISO systems, e.g., wire-line modems [28] and some wireless standards (e.g., IS-136 and GSM [16]). This arrangement is typically not used in MIMO systems.

In Section 6 we will analyze the performance of these two cases because they are widely applied in different communication systems.

3 Estimation of Channel Response

Due to the orthogonality of the pilots and assumption that the elements of $\mathbf{H}$ are *iid*, it can be shown that to obtain the maximum likelihood estimate of $\mathbf{H}$ it is sufficient to estimate h_{nm} (for $m = 1, \cdots, M,\ n = 1, \cdots, N$), independently[1]. This is identical to estimating a SISO channel response between the transmit antenna m and receive antenna n. The estimation is based on averaging the projections of the received signal on $d^p_{jm}\mathbf{s}^p_{jm}$ (for $j = 1, \cdots, L$ and $m = 1, \cdots, M$) as

$$\begin{aligned}\hat{h}_{nm} &= \frac{1}{LA_P}\sum_{j=1}^{L}(d^p_{jm}\mathbf{s}^p_{jm})^{\mathrm{H}}\mathbf{r}_n \\ &= \frac{1}{L}\sum_{j=1}^{L}(h_{nm} \ + \ (d^p_{jm}\mathbf{s}^p_{jm})^{\mathrm{H}}\mathbf{n}_n/A_P) \\ &= h_{nm} \ + \ \frac{1}{LA_P}\sum_{j=1}^{L}(d^p_{jm}\mathbf{s}^p_{jm})^{\mathrm{H}}\mathbf{n}_n \qquad (13)\end{aligned}$$

where $\hat{h}_{nm}$ denotes the estimate of the channel response h_{nm}. It can be shown that for a frequency-flat AWGN channel, given the pilot signal and the assumed properties of $\mathbf{H}$, (13) is the maximum-likelihood estimate of the channel response h_{nm} [17]. The estimation error is

$$n^e_{nm} = \ \frac{1}{LA_P}\sum_{j=1}^{L}(d^p_{jm}\mathbf{s}^p_{jm})^{\mathrm{H}}\mathbf{n}_n. \qquad (14)$$

[1]Based on the above assumptions it can be shown that in this particular case the ML estimate is equal to the LMMSE estimate considered in [22].

Note that n_{nm}^e corresponds to a Gaussian random variable with distribution $\mathcal{N}(0, N_0/(L\,(\alpha A)^2))$. Thus, the channel matrix estimate $\widehat{\mathbf{H}}$ is

$$\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{H_e} \tag{15}$$

where $\mathbf{H}_e$ is the estimation error. Each component of the error matrix $\mathbf{H}_e$ is an independent and identically distributed random variable n_{nm}^e given in (14) (where n_{nm}^e is the n-th row and m-th column element of $\mathbf{H}_e$).

Having the channel response estimated, the estimate of the transmitted data that is associated with the temporal signature $\mathbf{s}_i^d$ is obtained starting from the following statistics

$$z_{ni} = \frac{1}{A}(\mathbf{s}_i^d)^{\mathrm{H}}\mathbf{r}_n \tag{16}$$

where the amplitude A is assumed to be known at the receiver. z_{ni} corresponds to the n-th component of the vector

$$\mathbf{z}_i = [z_{1i}\cdots z_{Ni}]^{\mathrm{T}} = \mathbf{H}\,\mathbf{d}_i + \frac{1}{A}\mathbf{n}_i,\ \ i = 1,\ldots,K-LM \tag{17}$$

where the m-th component of $\mathbf{d}_i = [d_{i1}^d \cdots d_{iM}^d]^{\mathrm{T}}$ is d_{im}^d (data transmitted from the antenna m and assigned to the temporal signature $\mathbf{s}_i^d$). Further, $\mathrm{E}[\mathbf{n}_i\mathbf{n}_i^{\mathrm{H}}] = N_0\,\mathbf{I}_{N\times N}$. It can be shown that $\mathbf{z}_i$ is a sufficient statistic for detecting the transmitted data. Using $\mathbf{z}_i$ a MIMO receiver would perform detection of the transmitted data. Detection of the spatially multiplexed data which is not a focus of this paper can be done for example, using the VBLAST algorithm [11, 24].

As a common practice, the detection procedure assumes that the channel response is perfectly estimated, and that $\widehat{\mathbf{H}}$ corresponds to the true channel response. Let us rewrite the expression in (17) as

$$\mathbf{z}_i = (\mathbf{H} + \mathbf{H_e})\,\mathbf{d}_i + \frac{1}{A}\mathbf{n}_i - \mathbf{H_e}\,\mathbf{d}_i = \widehat{\mathbf{H}}\,\mathbf{d}_i + \left(\frac{1}{A}\mathbf{n}_i - \mathbf{H_e}\,\mathbf{d}_i\right). \tag{18}$$

The effective noise in the detection procedure (as a spatial vector) is

$$\bar{\mathbf{n}}_i = \left(\frac{1}{A}\mathbf{n}_i - \mathbf{H_e}\,\mathbf{d}_i\right). \tag{19}$$

For the given $\widehat{\mathbf{H}}$, the covariance matrix of the effective noise vector is

$$\boldsymbol{\Upsilon} = \boldsymbol{\Upsilon}(A) = \mathrm{E}_{\bar{\mathbf{n}}_i|\widehat{\mathbf{H}}}[\bar{\mathbf{n}}_i\bar{\mathbf{n}}_i^{\mathrm{H}}] = \frac{N_0}{A^2}\mathbf{I} + \mathrm{E}_{\mathbf{H_e}|\widehat{\mathbf{H}}}[\mathbf{H_e}\mathbf{H_e}^{\mathrm{H}}] \tag{20}$$

and it is a function of the amplitude A. As said earlier $\mathbf{H_e}$ is a matrix of *iid* Gaussian random variables with distribution $\mathcal{N}(0, N_0/(L\,(\alpha A)^2))$.

It can be shown that for a Rayleigh channel, where the entries of $\mathbf{H}$ are *iid* Gaussian random variables with distribution $\mathcal{N}(0,1)$, the above covariance matrix is

$$\boldsymbol{\Upsilon} = \frac{N_0}{A^2}\mathbf{I} + M\frac{1}{1+L\,(\alpha A)^2/N_0}\mathbf{I} + \left(\frac{1}{1+L\,(\alpha A)^2/N_0}\right)^2 \widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}. \tag{21}$$

4 Estimates Available to Receiver: Open Loop Capacity

Assuming that the channel response estimate is available to the receiver, only, we determine the lower bound for the open loop ergodic capacity as follows.

$$C \geq R = \frac{K-LM}{K}\,\mathrm{E}_{\widehat{\mathbf{H}}}\left[\log_2 \det\left(\mathbf{I}_{M\times M} + \widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\boldsymbol{\Upsilon}^{-1}\right)\right]. \tag{22}$$

The term $(K-LM)/K$ is introduced because L temporal signature per each transmit antenna are allocated to the pilot. Also, the random process $\widehat{\mathbf{H}}$ has to be stationary and ergodic (this is a common requirement for fading channel and ergodic capacity [5, 29]). We assume that the channel coding will span across great number of channel blocks (i.e., we use the well known infinite channel coding time horizon, required to achieve error-free data transmission with rates approaching capacity [30]).

In the above expression, equality holds if the effective noise (given in (19)) is AWGN with respect to the transmitted signal. If the effective noise is not AWGN, then the above rates represent the worst-case scenario, i.e., the lower bound [22, 31]. In achieving the above rates, the receiver assumes that the effective noise is interference (which is independent of the transmitted data) with a Gaussian distribution and spatial covariance matrix $\boldsymbol{\Upsilon}$. In addition, in the above expression R represents an achievable rate for reliable transmission (error-free) for the specific estimation procedure assumed. Knowing the channel response perfectly or using a better channel estimation scheme (e.g., decision driven schemes) may result in higher achievable rates.

Note that the capacity lower bounds for MIMO channel estimation independently derived in [22] assume the time multiplexing of data and pilot (i.e., training) symbols. The authors also present analytical results on the optimal properties required of the training sequences, their duration and power. The signal model presented here is more general than that and the distinguishing feature of this paper is the mismatched closed loop transmission analysis presented in the next section.

In the following we compare the above result in equation (22) to an information theoretical result presented in [5] (page 2641, expression (3.3.55)). The result is presented for the conventional SISO case, introducing a capacity lower bound for mismatched decoding as

$$C \geq R^* = \mathrm{E}_{\hat{h}} \left[\log_2 \left(1 + \frac{\hat{h}^2 P}{\mathrm{E}_{h|\hat{h}}(|h-\hat{h}|^2)P + N_0} \right) \right] \tag{23}$$

where h and $\hat{h}$ are the SISO channel response and its estimate, respectively. The above result is quite general, not specifying the channel response estimation procedure. The bound in (22) is an extension of the information theoretical bound in (23), capturing the more practical pilot assisted channel response estimation scheme and generalizing it to the MIMO case. Consequently,

Proposition 1 *For the SISO case ($M = 1, N = 1$), the rate R in (22) and R^* in (23), are related as*

$$R = \frac{K-L}{K} R^*, \;\; for \;\; P = \frac{K}{(K-L) + \alpha^2 L} P_{av} \tag{24}$$

where $\hat{h}$ is obtained using the pilot assisted estimation.

5 Estimates Available to Transmitter and Receiver: Closed Loop Capacity

In MIMO systems, when the channel state $\mathbf{H}$ is perfectly known at the transmitter, to maximize the capacity (under constrained transmit power), the transmitter performs optimization known as *the water pouring on eigen modes*. For SISO systems the water pouring algorithm is given in [2]. In practical communication systems, the channel state $\mathbf{H}$ has to be estimated at the receiver,

and then fed to the transmitter. In the case of a time varying channel, this practical procedure results in noisy and delayed (temporally mismatched) estimates being available to the transmitter to perform the optimization.

As said earlier, the MIMO channel is time varying. Let $\mathbf{H}_{i-1}$ and $\mathbf{H}_i$ correspond to consecutive block faded channel responses. In the following, the subscripts i and $i-1$ on different variables will indicate values corresponding to the block channel periods i and $i-1$, respectively. The temporal characteristic of the channel is described using the correlation

$$\mathrm{E}\left[h_{(i-1)nm}\, h^*_{inm}\right]/\Gamma = \kappa, \tag{25}$$

where $\Gamma = \mathrm{E}[h_{inm}h^*_{inm}]$, and h_{inm} is a stationary random process (for $m = 1, \cdots, M$ and $n = 1, \cdots, N$, denoting transmit and receive antenna indices, respectively). We assume that the knowledge of the correlation κ is not known at the receiver and the transmitter. Note that the above channel is modeled as a first order discrete Markov process[2].

Adopting a practical scenario, we assume that the receiver feeds back the estimate $\widehat{\mathbf{H}}_{i-1}$. Because the ideal channel state $\mathbf{H}_i$ is not available at the transmitter, we assume that $\widehat{\mathbf{H}}_{i-1}$ is used instead to perform the water pouring transmitter optimization for the i-th block. In other words the transmitter is ignoring the fact that $\mathbf{H}_i \neq \widehat{\mathbf{H}}_{i-1}$.

The water pouring optimization is performed as follows. First, the estimate is decomposed using singular value decomposition (SVD) as $\widehat{\mathbf{H}}_{i-1} = \widehat{\mathbf{U}}_{i-1}\widehat{\mathbf{\Sigma}}_{i-1}\widehat{\mathbf{V}}^{\mathrm{H}}_{i-1}$ [32]. Then, if the data vector $\mathbf{d}(k)$ is to be transmitted (in equation (4)), the following linear transformation is performed at the transmitter

$$\bar{\mathbf{d}}(k) = \widehat{\mathbf{V}}_{i-1}\mathbf{S}_i\mathbf{d}(k), \tag{26}$$

where the matrix $\mathbf{S}_i$ is a diagonal matrix whose elements s_{ijj} $(j = 1, \cdots M)$ are determined by the water pouring algorithm per singular value of $\widehat{\mathbf{H}}_{i-1}$, i.e., the diagonal element of $\widehat{\mathbf{\Sigma}}_{i-1}$ (denoted as

[2]Note that in the case of Jake's model, $\kappa = J_0(2\pi f_d \tau)$, where f_d is the maximum Doppler frequency and τ is the time difference between $h_{(i-1)nm}$ and h_{inm}.

$\hat{\sigma}_{(i-1)jj}$, $j = 1, \cdots M$). The diagonal element of $\mathbf{S}_i$ is defined as

$$s_{ijj}^2 = \begin{cases} \frac{1}{\gamma_0} - \frac{N_0}{|\hat{\sigma}_{(i-1)jj}|^2 A^2} & \text{for } |\hat{\sigma}_{(i-1)jj}|^2 A^2 \geq \gamma_0 \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

γ_0 is a *cut-off* value, and it depends on the channel fading statistics. It is selected such that the average transmit power stays the same P_{av} [2]. Consequently, at the time instant k the received spatial vector is

$$\mathbf{y}(k) = \mathbf{H}_i \widehat{\mathbf{V}}_{i-1} \mathbf{S}_i \mathbf{d}(k) + \mathbf{H}_i \mathbf{p}(k) + \mathbf{n}(k) = \mathbf{G}\mathbf{d}(k) + \mathbf{H}_i \mathbf{p}(k) + \mathbf{n}(k) \tag{28}$$

and

$$\mathbf{G} = \mathbf{H}_i \widehat{\mathbf{V}}_{i-1} \mathbf{S}_i. \tag{29}$$

The water pouring optimization is applied on the data bearing portion of the signal $\mathbf{d}(k)$, while the pilot $\mathbf{p}(k)$ is not changed. The receiver knows that the transformation in (26) is performed at the transmitter. The receiver performs estimation of the channel response matrix as given in section 3, resulting in $\widehat{\mathbf{G}} = \widehat{\mathbf{H}}_i \widehat{\mathbf{V}}_{i-1} \mathbf{S}_i$ and the error matrix $\mathbf{G_e} = \mathbf{H}_{\mathbf{e}i} \widehat{\mathbf{V}}_{i-1} \mathbf{S}_i$. In this case, the effective noise in (19) and its covariance matrix in (20) are modified accordingly resulting in

$$\boldsymbol{\Upsilon}^{WP} = \boldsymbol{\Upsilon}^{WP}(A) = \frac{N_0}{A^2}\mathbf{I} + \mathrm{E}_{\mathbf{G_e}|\widehat{\mathbf{G}}}[\mathbf{G_e}\mathbf{G_e}^{\mathrm{H}}]. \tag{30}$$

In the above and following expressions the superscript "*WP*" denotes *water pouring*. Note that the above application of the water pouring algorithm per eigen mode is suboptimal, i.e., it is mismatched (because $\widehat{\mathbf{H}}_{i-1}$ is used instead of $\mathbf{H}_i$). Consequently, the closed loop system capacity is bounded as,

$$C^{WP} \geq R^{WP} = \frac{K - LM}{K} \mathrm{E}_{\widehat{\mathbf{G}}} \left[\log_2 \det \left(\mathbf{I}_{M \times M} + \widehat{\mathbf{G}} \widehat{\mathbf{G}}^{\mathrm{H}} (\boldsymbol{\Upsilon}^{WP})^{-1} \right) \right]. \tag{31}$$

Similar to the comments related to the result in (22), the random process $\widehat{\mathbf{G}}$ has to be stationary and ergodic. Also, the channel coding will span across infinite number of channel blocks to achieve error-free data transmission approaching the above rates. Again, the equality holds if the effective noise is AWGN with respect to the transmitted signal and if not, then the above rates represent the

worst-case scenario, i.e., the lower bound [31]. In achieving the above rates, the receiver assumes that the effective noise is interference which is independent of the transmitted data with Gaussian distribution and spatial covariance matrix $\mathbf{\Upsilon}^{WP}$. Knowing the channel response perfectly or using a better channel estimation, or prediction scheme may result in higher achievable rates. There has also been some recent work in [33] on closed loop MIMO OFDM transmission over a parametric frequency selective channel model.

6 Examples and Numerical Results

6.1 SISO Systems

To illustrate the above analysis we start with SISO systems. In the SISO case, all previously defined spatial vectors and related matrices are now single dimensional (e.g., $\mathbf{d}_i$, $\mathbf{H}$, $\widehat{\mathbf{H}}$ and $\mathbf{\Upsilon}$ are now scalars d_i, h, $\hat{h}$ and υ, respectively). In Figure 1, we present the rate R in (22) as a function of the power allocated to the pilot (equation (10)). In this example, a pilot period K is 10 and coincides with the coherence time. A frequency-flat Rayleigh fading channel is assumed. The results are shown for the pilot arrangements corresponding to both case 1 and case 2. For the ideal knowledge of the channel response we apply the ergodic capacity formula [5]. Regarding the achievable rates, from the above results we observe that case 1 is less sensitive to the pilot power allocation than case 2 (i.e., in case 2, R is dropping faster if the allocated power is different than the one that results in the maximum value). Further, case 1 is achieving higher maximum achievable rates than case 2.

For the given SNR, we define the capacity efficiency ratio η as the ratio between the maximum rate R (with respect to the pilot power) and the ergodic capacity $C_{m\times n}$ in the case of the ideal knowledge of the channel response, i.e.,

$$\eta_{m\times n} = \frac{\max_{\mu} R}{C_{m\times n}}. \tag{32}$$

The index m and n correspond to number of transmit and receive antennas, respectively. In Figure 2, we show that the capacity efficiency ratio $\eta_{1\times 1}$ increases with the channel coherence time. From

the above results we conclude that case 1 is a more efficient scheme than case 2.

6.2 MIMO Systems

In Figure 3, we present the rate R in (22) as a function of the power allocated to the pilot (equation (10)), for different number of transmit and receive antennas. In this and the following numerical examples we consider only the pilot arrangement case 1 (treating case 2 impractical for MIMO systems). We observe the rates for the Rayleigh channel, $SNR = 12$dB and the channel coherence time length $K = 40$. Solid lines correspond to a system with the channel response estimation, and dashed lines to a system with the ideal knowledge of the channel response. Further, in Figure 4 we show the capacity efficiency ratio η for different number of transmit and receive antennas vs. different channel coherence time lengths. We observe that as the number of transmit antennas increase, the sensitivity to the channel response estimation error is more pronounced (while keeping the same number of receive antennas). For example, for the same channel coherence time length, the capacity efficiency ratio of the 4×4 system is lower than that in the case of the 3×4 system.

In Figure 5(a) we present open loop (solid lines) and closed loop (dashed lines) ergodic capacities. Idealized conditions are assumed, i.e., the ideal knowledge of the channel response is available to the transmitter and receiver and perfect temporal match $\mathbf{H}_{i-1} = \mathbf{H}_i$ (for the water pouring optimization) is assumed. Comparing the closed loop and open loop capacity, we observe that the gains are more pronounced for lower SNR (e.g, for 4×4 system at 0dB, the gain of the closed loop system is approximately 2dB, while at 12dB, it drops below 0.5dB). Further, we note that in the case of 2×4 ond 1×4 systems, the gain practically disappears. This is explained as an effect of multiple receive antennas (greater than the number of transmit antennas) providing already sufficient degree of diversity, eliminating any need for transmitter optimization. Instead of the ergodic capacities, when observing the cumulative distribution function (cdf) of the capacity, the difference is more pronounced (Figure 5(b), for $SNR = 4$dB) (see more on the "capacity versus outage" approach in [5]).

From the results in Figure 6, we observe how the temporal mismatch between successive

channel responses ($\mathbf{H}_{i-1} \neq \mathbf{H}_i$) affects the achievable rates R^{WP} in (31). As said earlier, the temporal mismatch is characterized by the correlation $\mathrm{E}\left[h_{(i-1)nm}\, h^*_{inm}\right]/\Gamma = \kappa$ (for $m = 1, \cdots, M$ and $n = 1, \cdots, N$). We observe the cases when the ideal channel response (dashed lines) and channel response estimates (solid lines) are available at the transmitter and the receiver. Solid lines correspond to the channel response estimation where the pilot power is selected to maximize the achievable rate R^{WP}. We observe the rates for the Rayleigh channel, $SNR = 4$dB and the coherence time length $K = 40$. Note that for $\kappa = 0$ (i.e., when the successive channel responses are uncorrelated), the achievable rate is lower than in the case of $\kappa = 1$ (i.e., when the successive channel responses are fully correlated). The drop in the achievable rates is not substantial, even though the water pouring algorithm is fully mismatched for $\kappa = 0$. We explain this behavior in the following. In the case of a Rayleigh channel, the matrix $\widehat{\mathbf{V}}_{i-1}\mathbf{S}_i$ usually has M degrees of freedom, and a small condition number of the corresponding covariance matrix. Consequently, even in the mismatched case, multiplying $\mathbf{H}_i$ with $\widehat{\mathbf{V}}_{i-1}\mathbf{S}_i$ preserves the degrees of freedom of the matrix $\mathbf{H}_i$ resulting in a high capacity of the composite channel $\mathbf{G}$ in (29). We expect the detrimental effects of the mismatch to be amplified in the case of Rician channels, especially those with large K-factor. This is because Rician channels result in the matrix $\widehat{\mathbf{V}}_{i-1}\mathbf{S}_i$ having a few dominant degrees of freedom there by making accurate feedback beneficial.

In Figure 7 we compare the open loop scheme to the closed loop scheme under temporal mismatch. It is observed that when the channel coherence is low (i.e., low correlation κ), it is better to not use a closed loop scheme. In the observed case (4×4, $SNR = 4$dB and coherence time $K = 40$), for the correlation coefficient $\kappa < 0.7$ the achievable rates for the closed loop scheme are lower than in the open loop case.

7 Conclusion

In this paper we have studied how the estimation error of the frequency-flat time-varying channel response affects the performance of a MIMO communication system. Using a block-fading channel model, we have connected results of information theory with practical pilot estimation for such

systems. The presented analysis may be viewed as a study of mismatched receiver and transmitter algorithms in MIMO systems. We have considered two pilot based schemes for the estimation. The first scheme uses a single pilot symbol per block with different power than the data symbol power. The second scheme uses more than one pilot symbol per block, whose power is the same as the data symbol power. We have presented how the achievable data rates depend on the percentage of the total power allocated to the pilot, background noise level and the channel coherence time length. Our results have shown that the first pilot-based approach is less sensitive to the fraction of power allocated to the pilot. Furthermore, we have observed that as the number of transmit antennas increase, the sensitivity to the channel response estimation error is more pronounced (while keeping the same number of receive antennas). The effects of the estimation error are evaluated in the case of the estimates being available at the receiver only (open loop), and in the case when the estimates are fed back to the transmitter (closed loop) allowing water pouring transmitter optimization. In the case of water pouring transmitter optimization and corresponding rates, we have not observed significant gains versus the open loop rates for the channel models considered here. Further, we observe in certain cases, it is better to use the open loop scheme as opposed to the closed loop scheme. The analysis presented here can be used to optimally allocate pilot power for various system and channel operating conditions, and to also determine the effectiveness of closed loop feedback.

Acknowledgments

The authors would like to thank Dr. Gerard Foschini and Dr. Dimitry Chizhik for their constructive comments and valuable discussions.

References

[1] G. J. Foschini and J. Salz, "Digital Communications over Fading Radio Channels," *Bell Labs Technical Journal*, pp. 429–456, February 1983.

[2] A. Goldsmith and P. Varaiya, "Increasing Spectral Efficiency Through Power Control," *ICC*, vol. 1, pp. 600–604, 1993.

[3] G. Caire, G. Taricco, and E. Biglieri, "Optimum Power Control Over Fading Channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 1468 –1489, July 1999.

[4] L. Song and N. Mandayam, "Hierarchical SIR and Rate Control on the Forward Link for CDMA Data Users Under Delay and Error Constraints," *IEEE JSAC*, vol. 19, pp. 1871 –1882, October 2001.

[5] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-Theoretic and Communications Aspects," *IEEE Transactions on Information Theory*, vol. 44, pp. 2619 –2692, October 1998.

[6] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multiple Antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[7] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, no. 6, pp. 315–335, 1998.

[8] D. Chizhik, G. J. Foschini, M. J. Gans, and R. Valenzuela, "Keyholes, Correlations, and Capacities of Multielement Transmit and Receive Antennas," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 361–368, April 2002.

[9] D. Chizhik, J. Ling, P. W. Wolniansky, R. A. Valenzuela, N. Costa, and K. Huber, "Multiple Input Multiple Output Measurements and Modeling in Manhattan," *VTC Fall*, vol. 1, pp. 107–110, 2002.

[10] P. Soma, D. S. Baum, V. Erceg, R. Krishnamoorthy, and A. Paulraj, "Analysis and Modeling of Multiple-Input Multiple-Output (MIMO) Radio Channel Based on Outdoor Measurements Conducted at 2.5 GHz for Fixed BWA Applications," *IEEE Conference ICC 2002*, pp. 272 – 276, 2002.

[11] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified Processing for Wireless Communications at High Spectral Efficiency," *IEEE JSAC*, vol. 17, pp. 1841–1852, November 1999.

[12] D. Samardzija, P. Wolniansky, and J. Ling, "Performance Evaluation of the VBLAST Algorithm in W-CDMA Systems," *The IEEE Vehicular Technology Conference (VTC)*, vol. 2, pp. 723–727, September 2001. Atlantic City.

[13] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection Algorithm and Initial Laboratory Results using V-BLAST Space-Time Communication Architecture," *Electronics Letters*, vol. 35, pp. 14–16, January 1999.

[14] H. Zheng and D. Samardzija, "Performance Evaluation of Indoor Wireless System Using BLAST Testbed," *The IEEE Vehicular Technology Conference (VTC)*, vol. 2, pp. 905–909, October 2001. Atlantic City.

[15] D. Samardzija, C. Papadias, and R. A. Valenzuela, "Experimental Evaluation of Unsupervised Channel Deconvolution for Wireless Multiple-Transmitter/Multiple-Receiver Systems," *Electronic Letters*, pp. 1214–1215, September 2002.

[16] G. L. Stuber, *Principles of Mobile Communications*. Kluwer Academic Publishers, first ed., 1996.

[17] H. V. Poor, *An Introduction to Signal Detection and Estimation*. Springer-Verlag, second ed., 1994.

[18] A. Lapidoth and S. Shamai, "Fading Channels: How Perfect Need "Perfect Side Information" Be?," *IEEE Transaction on Information Theory*, vol. 48, pp. 1118 –1134, May 2002.

[19] S. Shamai and T. Marzetta, "Multiuser Capacity in Block Fading With No Channel State Information," *IEEE Transaction on Information Theory*, vol. 48, pp. 938 –942, April 2002.

[20] T. L. Marzetta, "Blast Training: Estimating Channel Characteristics for High-Capacity Space-Time Wireless," *37th Annual Allerton Conference on Communications, Control and Computing*, September 1999.

[21] J. Baltersee, G. Fock, and H. Meyr, "Achievable Rate of MIMO Channels With Data-Aided Channel Estimation and Perfect Interleaving," *IEEE JSAC*, vol. 19, pp. 2358–2368, December 2001.

[22] B. Hassibi and B. M. Hochwald, "How Much Training is Needed in Multiple-Antenna Wireless Links," *Technical Memorandum, Bell Laboratories, Lucent Technologies*, October 2001.

[23] H. Holma and A. Toskala, eds., *WCDMA for UMTS*. Wiley, first ed., 2000.

[24] A. Burg, E. Beck, D. Samardzija, M. Rupp, and et al., "Prototype Experience for MIMO BLAST over Third Generation Wireless System," *IEEE JSAC Special Issue on MIMO Systems and Applications*, vol. 21, April 2003.

[25] J. C. Guey, M. P. Fitz, M. R. Bell, and W. Kuo., "Signal Design for Transmitter Diversity Wireless Communication Systems over Rayleigh Fading Channels," *Proc. of IEEE Vehicular Technology Conference*, vol. 1, pp. 136–140, 1996. Atlanta.

[26] J. C. Guey, M. P. Fitz, M. R. Bell, and W. Y. Kuo, "Signal Design for Transmitter Diversity Wireless Communication Systems over Rayleigh Fading Channels," *IEEE Transactions on Communications*, vol. 47, pp. 527–537, 1999.

[27] D. Shiu, G. J. Foschini, and M. J. Gans, "Fading Correlation and Its Effects on the Capacity of Multielement Antenna Systems," *IEEE Transactions on Communications*, vol. 48, pp. 502 –513, March 2000.

[28] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 3rd ed., 1995.

[29] L. Li and A. Goldsmith, "Capacity and Optimal Resource Allocation for Fading Broadcast Channels - Part I: Ergodic Capacity," *Information Theory, IEEE Transactions*, vol. 47, pp. 1083–1102, March 2001.

[30] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.

[31] R. G. Gallager, *Information Theory and Reliable Communications*. New York: John Wiley and Sons, 1968.

[32] G. Strang, *Linear Algebra and its Applications*. Harcourt Brace Jovanovich, third ed., 1988.

[33] C. Wen, Y. Wang, and J. Chen, "Adaptive Spatio-Temporal Coding Scheme for Indoor Wireless Communication," *IEEE JSAC*, vol. 21, pp. 161–170, February 2003.

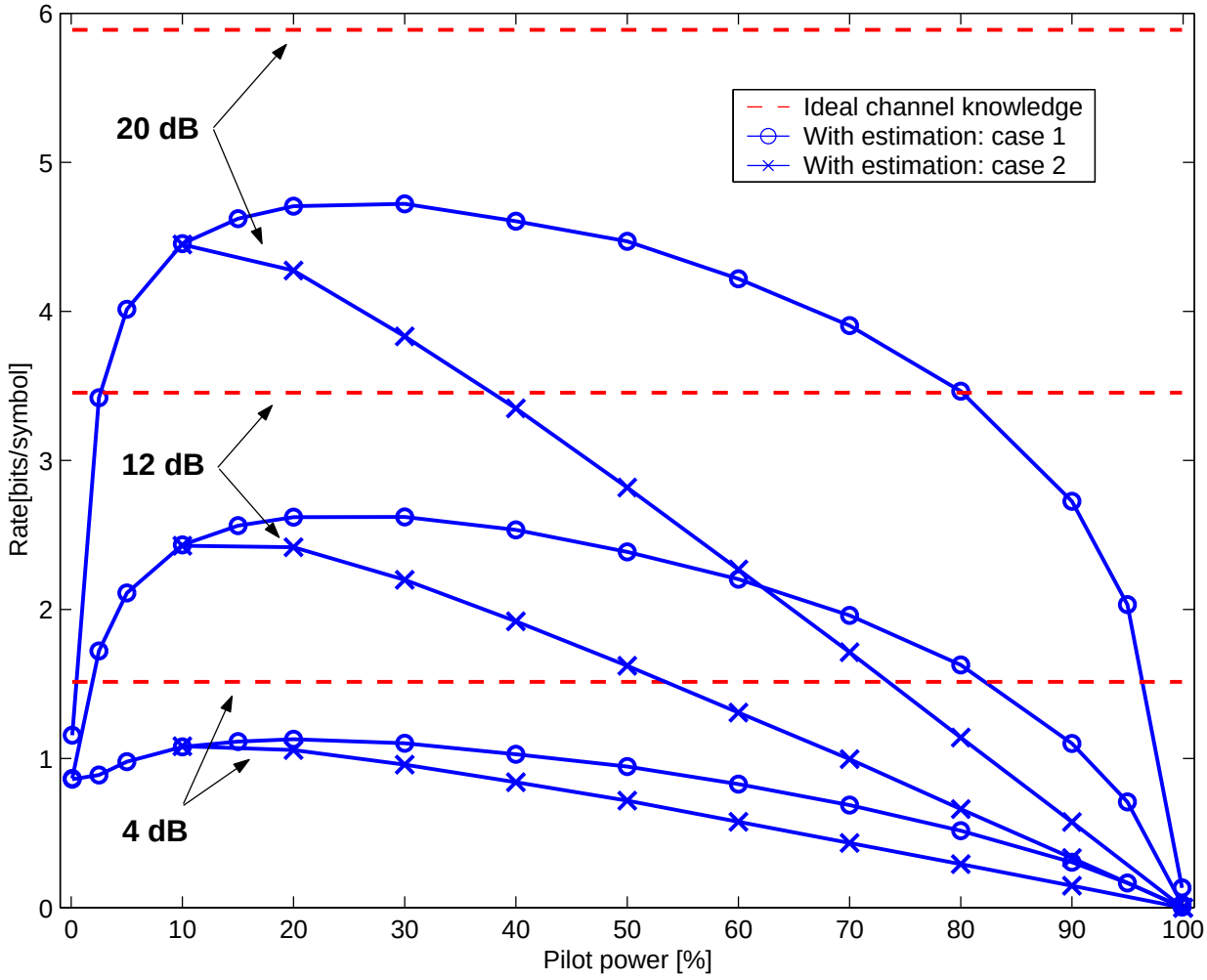


Figure 1: Achievable open loop rates vs. power allocated to the pilot, SISO system, $SNR = 4, 12, 20$dB, coherence time $K = 10$, Rayleigh channel.

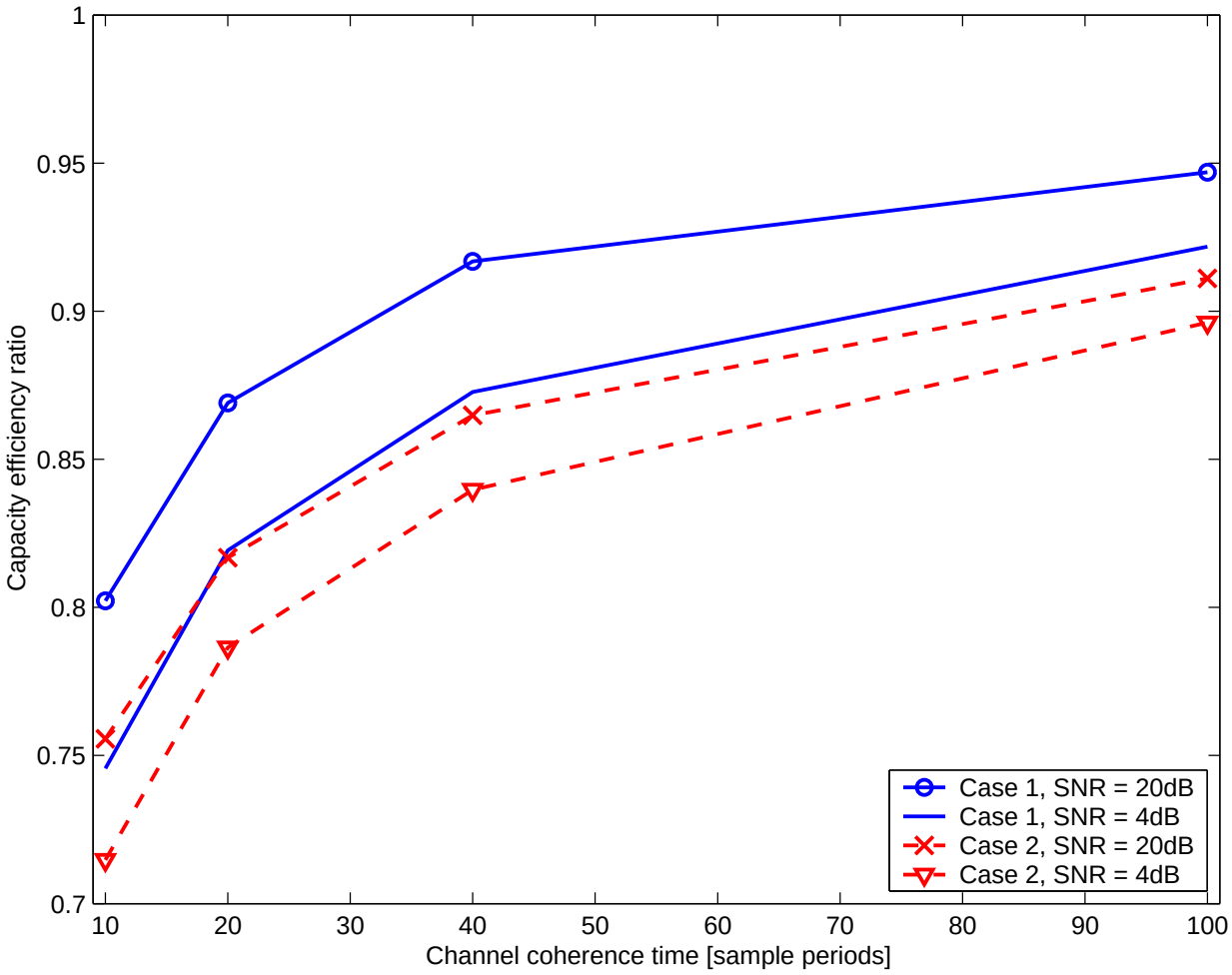


Figure 2: Capacity efficiency ratio vs. channel coherence time ($K = 10, 20, 40, 100$), SISO system, $SNR = 4, 20$dB, Rayleigh channel.

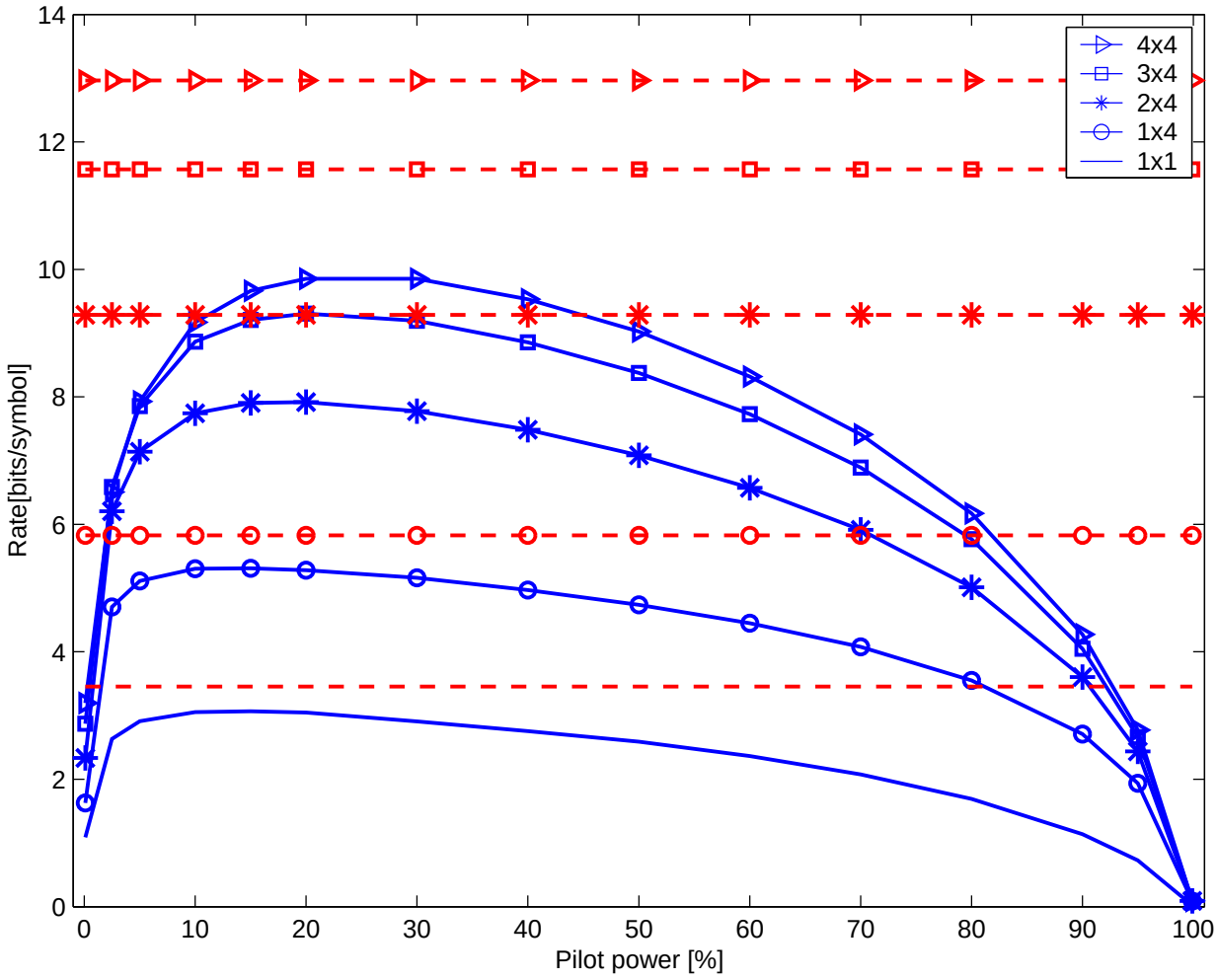


Figure 3: Achievable open loop rates vs. power allocated to the pilot, MIMO system, $SNR = 12$dB, coherence time $K = 40$, Rayleigh channel, solid line corresponds to a system with the channel response estimation, and dashed line to the case of the ideal channel response knowledge.

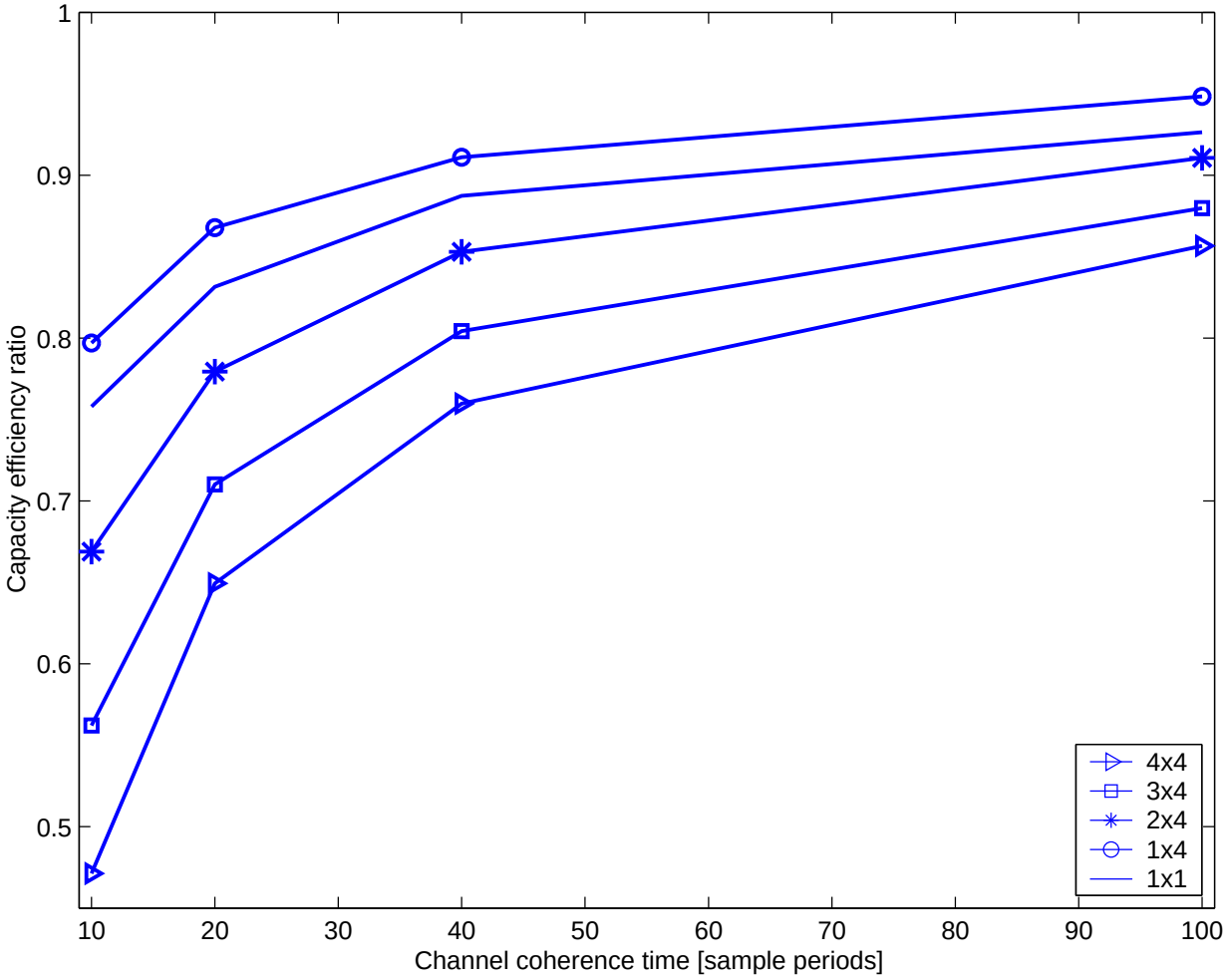
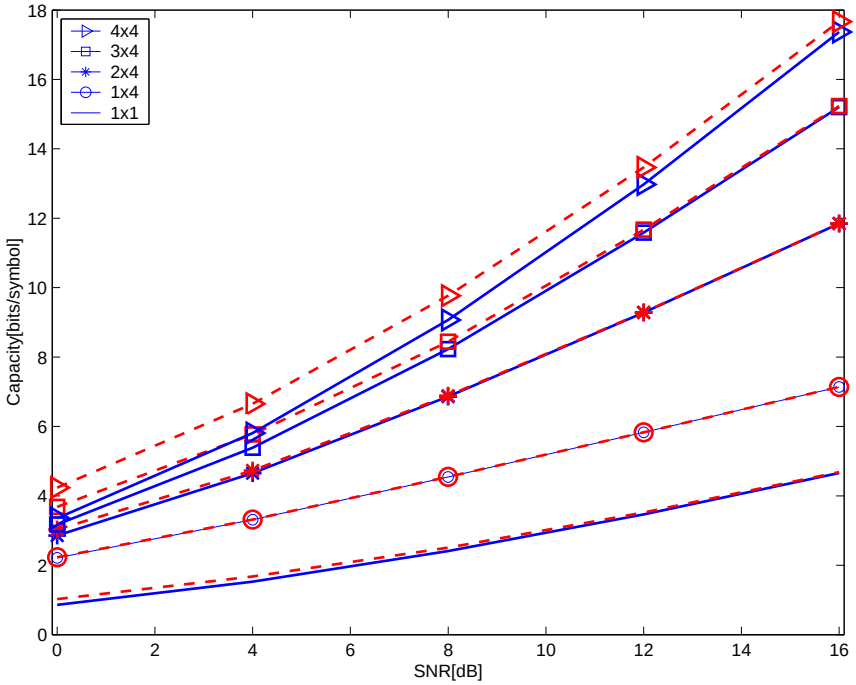


Figure 4: Capacity efficiency ratio vs. channel coherence time ($K = 10, 20, 40, 100$), MIMO system, $SNR = 12$dB, Rayleigh channel.



(a) Ergodic capacity vs. SNR, MIMO system, ideal knowledge of the channel response, Rayleigh channel, solid line corresponds to open loop capacity, and dashed line to closed loop capacity (perfect temporal match $\mathbf{H}_{i-1} = \mathbf{H}_i$ is assumed).

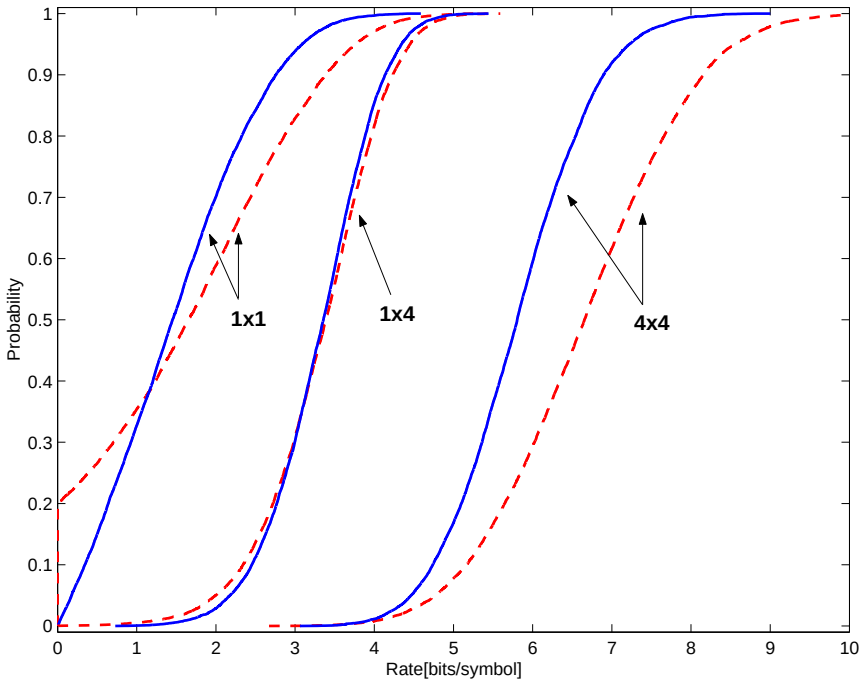


Figure 5: (b) CDF of capacity, MIMO system, $SNR = 4$dB, ideal knowledge of the channel response, Rayleigh channel, solid line corresponds to open loop capacity, and dashed line to closed loop capacity (perfect temporal match $\mathbf{H}_{i-1} = \mathbf{H}_i$ is assumed).

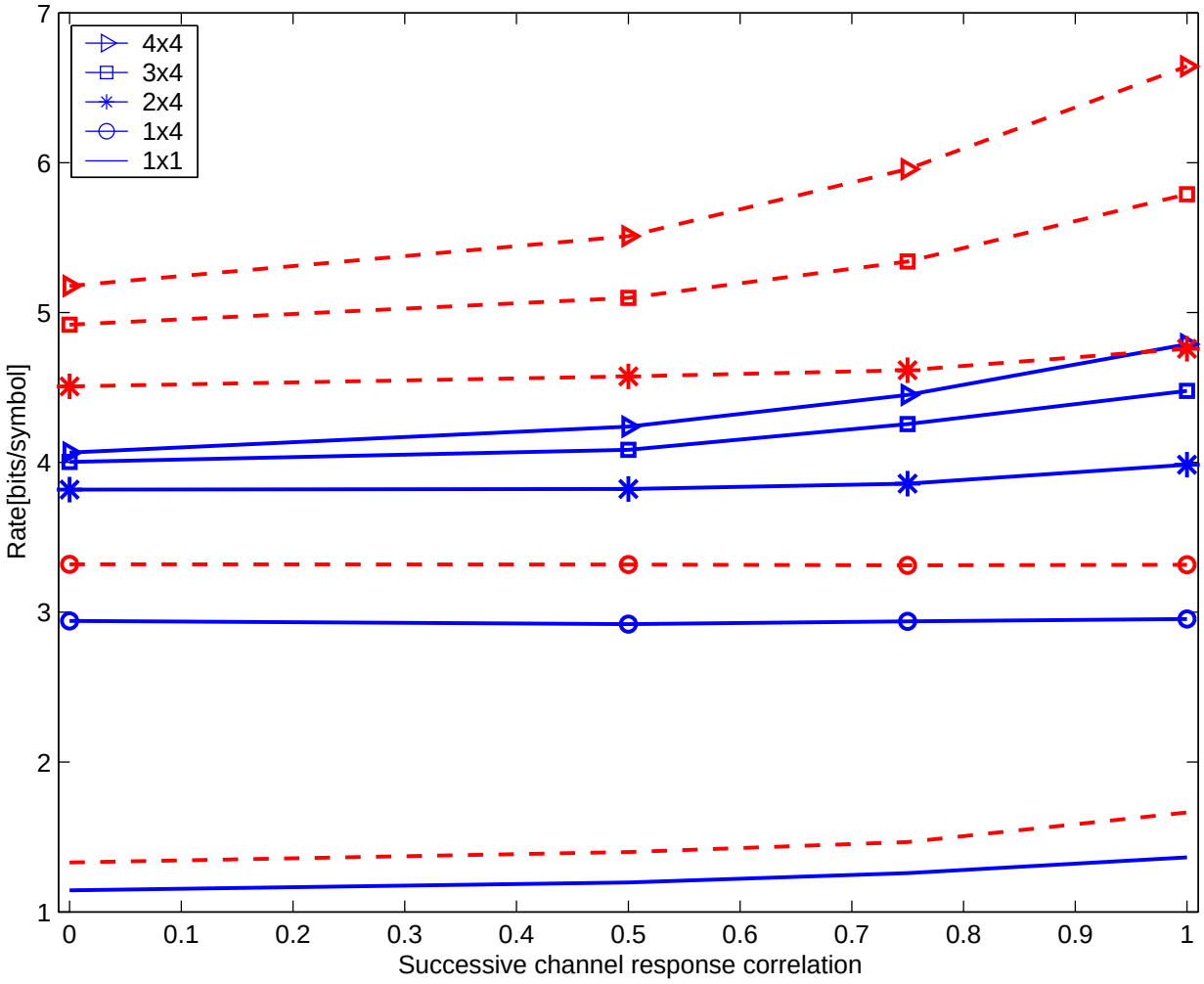


Figure 6: Achievable closed loop rates vs. correlation between successive channel responses, MIMO system, $SNR = 4$dB, coherence time $K = 40$, Rayleigh channel, solid line corresponds to a system with the channel response estimation, and dashed line to the case of the ideal channel response available at the transmitter and the receiver (but with the temporal mismatched $\mathbf{H}_{i-1} \neq \mathbf{H}_i$).

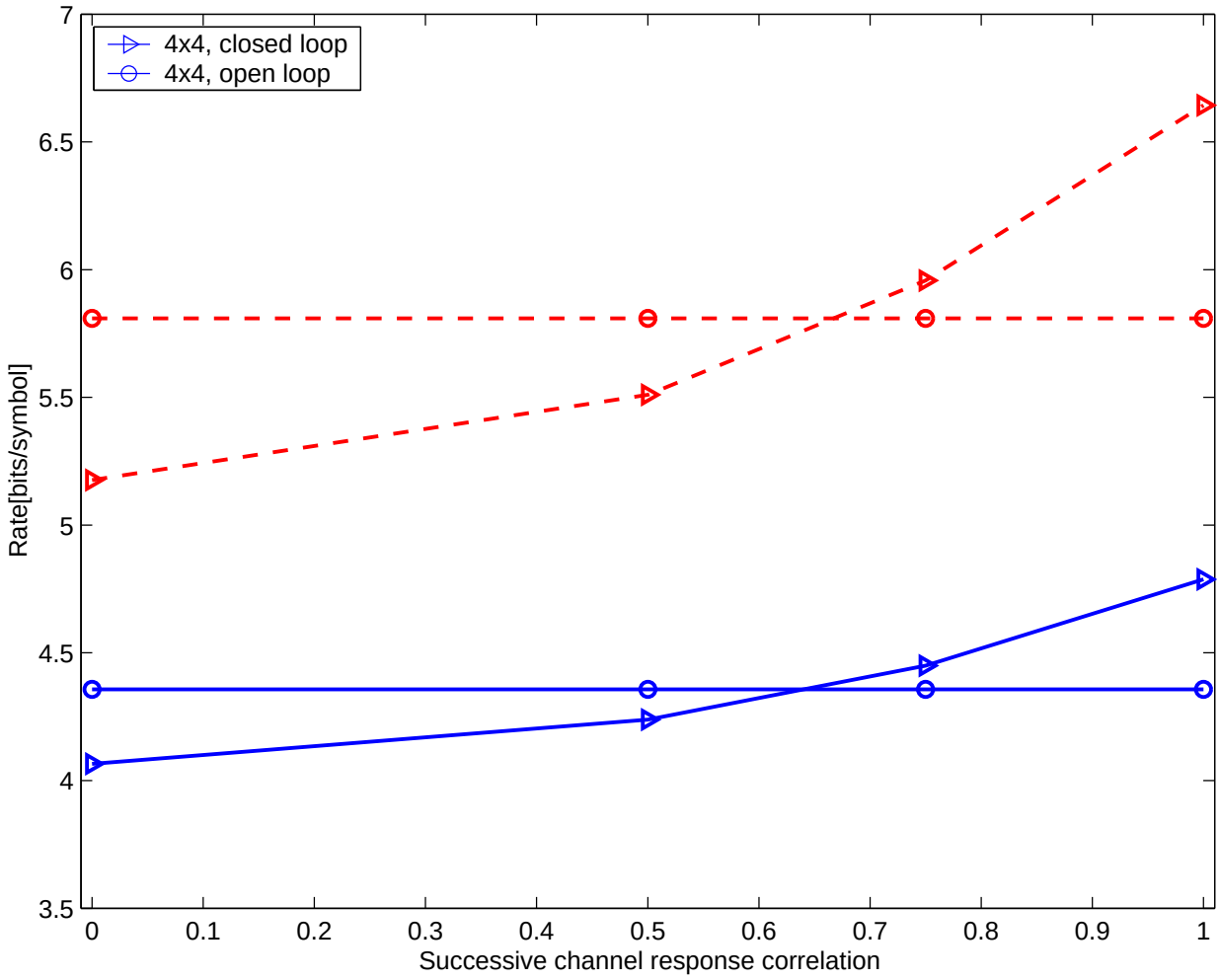


Figure 7: Achievable closed loop and open loop rates vs. correlation between successive channel responses, MIMO system 4×4, $SNR = 4$dB, coherence time $K = 40$, Rayleigh channel, solid line corresponds to a system with the channel response estimation, and dashed line to the case of the ideal channel response available at the transmitter and the receiver (but with the temporal mismatched $\mathbf{H}_{i-1} \neq \mathbf{H}_i$).