Achievable rates of MIMO downlink beamforming with non-perfect CSI: a comparison between quantized and analog feedback

Giuseppe Caire
University of Sourthern California
Los Angeles CA, 90089 USA

Nihar Jindal
University of Minnesota
Minneapolis MN, 55455 USA

Mari Kobayashi
Cenre Tecnologic de Telecomunicacions de Catalunya
Barcelona, Spain

Abstract— **We consider a MIMO fading broadcast channel and compare the achievable ergodic rates when the channel state information at the transmitter is provided by "analog" noisy feedback or by quantized (digital) feedback. The superiority of digital feedback is shown whenever the number of feedback channel uses per channel coefficient is larger than 1. Also, we show that by proper design of the digital feedback link, errors in the feedback have a minor effect even by using very simple uncoded modulation. Finally, we show that analog feedback achieves a fraction $1 - 2F$ of the optimal multiplexing gain even in the presence of a feedback delay, when the fading belongs to the class of "Doppler processes" with normalized maximum Doppler frequency shift $0 \le F < 1/2$.**

I. Model setup and background

We consider a multi-input multi-output (MIMO) Gaussian broadcast channel modeling the downlink of a system where the base station (transmitter) has M antennas and K user terminals (receivers) have one antenna each. A channel use of such channel is described by

$$y_k = \mathbf{h}_k^{\mathsf{H}} \mathbf{x} + z_k, \ \ k = 1, \ldots, K \tag{1}$$

where y_k is the channel output at receiver k, $z_k \sim \mathcal{CN}(0, N_0)$ is the corresponding AWGN, $\mathbf{h}_k \in \mathbb{C}^M$ is the vector of channel coefficients from the k-th receiver to the transmitter antenna array and $\mathbf{x}$ is the channel input vector. The channel input is subject to the average power constraint $\mathbb{E}[|\mathbf{x}|^2] \le P$.

We assume that the channel *state*, given by the collection of all channel vectors $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K] \in \mathbb{C}^{M \times K}$, varies in time according to a block fading model where $\mathbf{H}$ is constant over each frame of length T channel uses, and evolves from frame to frame according to an ergodic stationary jointly Gaussian process. As a special case, we have the i.i.d. block-fading channel where $\mathbf{H}$ is an independent random matrix on each frame. The entries of $\mathbf{H}$ are mutually independent and identically distributed, and the first-order distribution of $\mathbf{H}$ is Gaussian i.i.d. with elements $\sim \mathcal{CN}(0, 1)$.

A. Capacity results

If $\mathbf{H}$ is perfectly and instantaneously known to all terminals (perfect CSIT and CSIR), the capacity region of the channel (1) is obtained by MMSE-DFE beamforming and Gaussian dirty-paper coding (see details in [1–5]). Because of simplicity and robustness to non-perfect CSIT, simpler *linear precoding* schemes with standard Gaussian coding have been extensively considered. A particularly simple scheme consists of zero-forcing (ZF) beamforming, where the transmit signal is formed as $\mathbf{x} = \mathbf{V}\mathbf{u}$, such that $\mathbf{V} \in \mathbb{C}^{M \times K}$ is a zero-forcing beamforming matrix and $\mathbf{u} \in \mathbb{C}^K$ contains the symbols from K independently generated Gaussian codewords. For $K \le M$, the k-th column $\mathbf{v}_k$ of $\mathbf{V}$ is chosen to be a unit vector orthogonal to the subspace $\mathcal{S}_k = \mathrm{span}\{\mathbf{h}_j : j \ne k\}$ generated by all other users' channels. In this case, the achievable sum rate is given by

$$R^{\mathrm{ZF}} = \max_{\sum_k \mathbb{E}[\mathcal{P}_k(\mathbf{H})] \le P} \sum_{k=1}^{K} \mathbb{E}\left[\log\left(1 + \frac{|\mathbf{h}_k^{\mathsf{H}} \mathbf{v}_k|^2 \mathcal{P}_k(\mathbf{H})}{N_0}\right)\right] \tag{2}$$

where the optimal power allocation is obtained straightforwardly by waterfilling over the set of channel gains $\{|\mathbf{h}_k \mathbf{v}_k|^2 : k = 1, \ldots, K\}$. If $K > M$, ZF precoding can be applied jointly with some *user scheduling* algorithm that selects at every channel use an active user subset of size not larger than M. Schemes for user scheduling have been extensively discussed. In this paper, however, we are not concerned with the user scheduling problem and we shall consider the situation $K = M$. We are mainly interested in the high-spectral efficiency regime, where we can characterize the achievable sum rate as $\kappa \log P/N_0 + O(1)$, and κ is the "system multiplexing gain" or "pre-log factor" of the ergodic sum rate. Hence, it is well-known that restricting to the simple constant power allocation $\mathcal{P}_k = P/M$ for all $k = 1, \ldots, M$ incurs in a loss only in the $O(1)$ term. We shall restrict to this choice in the rest of this paper.

It is well-known that, under perfect CSIT and CSIR, both the optimal "Dirty-Paper" sum-rate C and the zero-forcing sum-rate R^{ZF} are equal to $M \log P/N_0 + O(1)$. On the contrary, under non-perfect CSI the rate sum may behave in a radically different way. For example, if $\mathbf{H}$ has distribution invariant under left multiplication by unitary matrices, it is known [1, 6] that under no CSIT and perfect CSIR the multiplexing gain reduces to 1, i.e., the sum rate is equal to $\log P/N_0 + O(1)$.

B. Channel state feedback models

We consider some specific CSIT and CSIR models and derive lower-bounds to the corresponding achievable ergodic rates by analyzing a *naive* beamforming scheme that computes a mismatched ZF beamforming matrix $\widehat{\mathbf{V}}$ from the CSIT. In particular, we consider an "analog" CSIT feedback scheme where the transmitter observation at frame time t is given by

$$\{\mathbf{G}(\tau) = \sqrt{\beta P}\mathbf{H}(\tau) + \mathbf{W}(\tau) : \tau = -\infty, \ldots, t - d\} \tag{3}$$

where $\{\mathbf{W}(\tau)\}$ is a spatially and spectrally white Gaussian process with elements $\sim \mathcal{CN}(0, N_0)$ and d is the feedback delay. This models the case where the channel coefficients are explicitly transmitted on the reverse link (uplink) using unquantized quadrature-amplitude modulation [7–9]. The power scaling β corresponds to the number of channel uses per channel coefficient, assuming that transmission in the feedback channel has fixed peak power P and that the channel state vector is modulated by a $\beta M \times M$ unitary spreading matrix [7]. A simplifying assumption of this work is that we consider no fading and orthogonal access in the CSIT feedback link. In addition, we assume that the SNR on the feedback channel is equivalent to the un-faded downlink SNR (P/N_0).

A different CSIT feedback approach is based on quantizing the channel vector at each receiver and transmitting back to the base station a packet of B bits, representing the corresponding quantization index. In the case of no feedback delay and no feedback errors, in [10, Theorem 1] it is shown that the gap between ZF with ideal CSI and the naive ZF scheme is given by

$$\Delta R_{\text{quant.}} \leq \log\left(1 + \frac{P}{N_0} 2^{-\frac{B}{M-1}}\right) \tag{4}$$

While this result is obtained for a particular random ensemble of channel quantization schemes referred to as *Random Vector Quantizer* (see [10] and references therein), a bound on the best possible channel vector quantizer shows that (4) is tight for large P/N_0 [10].

II. Rate gap bound for analog CSIT feedback

In the case of i.i.d. block fading and no feedback delay, the analog CSIT feedback yields the observation of $\mathbf{G} = \sqrt{\beta P}\mathbf{H} + \mathbf{W}$ at the beginning of every frame. The transmitter computes the MMSE estimate of the channel matrix, $\widehat{\mathbf{H}} \frac{\sqrt{\beta P}}{N_0 + \beta P}\mathbf{G}$. The k-th column $\widehat{\mathbf{v}}_k$ of $\widehat{\mathbf{V}}$ is a unit vector orthogonal to the subspace $\mathcal{S}_k = \mathrm{span}\{\widehat{\mathbf{h}}_j : j \neq k\}$. Notice that we can write $\mathbf{H} = \widehat{\mathbf{H}} + \mathbf{E}$, where $\widehat{\mathbf{H}}$ and $\mathbf{E}$ are mutually independent and have Gaussian i.i.d. components with mean zero and variance $\frac{\beta P}{N_0 \sigma_e^2}$ and $\sigma_e^2 = (1 + \beta P/N_0)^{-1}$, respectively. Furthermore, $\mathbf{H}$, $\widehat{\mathbf{H}}$ and $\mathbf{E}$ are identically distributed apart from different per-component variances (scaling).

The signal at the k-th receiver is given by

$$y_k = (\mathbf{h}_k^{\mathsf{H}} \widehat{\mathbf{v}}_k) u_k + \sum_{j \neq k} (\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j) u_j + z_k \tag{5}$$

We assume that the frame duration is long enough such that some training scheme can be used in the downlink channel. Training allows each receiver to estimate: 1) the useful signal coefficient, $a_k = (\mathbf{h}_k^{\mathsf{H}} \widehat{\mathbf{v}}_k)$ and 2) the variance of the interference plus noise $\zeta_k = \sum_{j \neq k} (\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j) u_j + z_k$, given by $\Sigma_k = \mathbb{E}\left[|\zeta_k|^2 | \mathbf{e}_k, \widehat{\mathbf{H}}\right] = N_0 + \sum_{j \neq k} |\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j|^2 P/M$. This conditioning is due to the fact that Σ_k is estimated on each frame, and the coefficients $(\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j)$ are constant over each frame and change from frame to frame, following the block i.i.d. fading model. The maximum achievable rate of user k subject to the above assumptions is lowerbounded by assuming a Gaussian input $u_k = u_k^G \sim \mathcal{CN}(0, P/M)$, and by considering the worst-case noise plus interference distribution in every frame. Using stationarity and ergodicity, we have [1]

$$\begin{aligned} R_k &\geq \mathbb{E}\left[\inf_{\zeta_k : \mathbb{E}[|\zeta_k|^2] \leq \Sigma_k} I(u_k^G; y_k | a_k, \Sigma_k)\right] \\ &\overset{(a)}{=} \mathbb{E}\left[\log\left(1 + \frac{|a_k|^2 P}{\Sigma_k M}\right)\right] \end{aligned} \tag{6}$$

where (a) follows from [11], noticing that $a_k u_k^G$ and ζ_k are uncorrelated (even after conditioning on a_k, Σ_k).

Next, we shall bound the rate gap incurred by the naive ZF beamforming and analog feedback with respect to the ZF beamforming with ideal CSIT. Denoting by R_k^{ZF} the rate of user k with uniform (across users) and constant (in time) power allocation $\mathcal{P}_k(\mathbf{H}) = P/M$ in (2), we have

$$\begin{aligned} \Delta R_{\text{analog}} &\triangleq R_k^{\mathrm{ZF}} - R_k \\ &\leq \mathbb{E}\left[\log\left(1 + \frac{|\mathbf{h}_k^{\mathsf{H}} \mathbf{v}_k|^2 P}{N_0 M}\right)\right] - \mathbb{E}\left[\log\left(1 + \frac{|a_k|^2 P}{\Sigma_k M}\right)\right] \\ &= \mathbb{E}\left[\log\left(1 + \frac{|\mathbf{h}_k^{\mathsf{H}} \mathbf{v}_k|^2 P}{N_0 M}\right)\right] \\ &\quad -\mathbb{E}\left[\log\left(1 + \frac{\left(\sum_{j \neq k} |\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j|^2 + |a_k|^2\right) P}{N_0 M}\right)\right] \\ &\quad +\mathbb{E}\left[\log\left(1 + \sum_{j \neq k} \frac{|\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j|^2 P}{N_0 M}\right)\right] \\ &\overset{(a)}{\leq} \mathbb{E}\left[\log\left(1 + \sum_{j \neq k} \frac{|\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j|^2 P}{N_0 M}\right)\right] \\ &\overset{(b)}{\leq} \log\left(1 + \frac{P}{N_0 M} \sum_{j \neq k} \mathbb{E}[|\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j|^2]\right) \\ &\overset{(c)}{=} \log\left(1 + \frac{\sigma_e^2 P}{N_0} \frac{M-1}{M}\right), \end{aligned} \tag{7}$$

where (a) follows from the fact that $\sum_{j \neq k} |\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j|^2 + |a_k|^2$ stochastically dominates $|\mathbf{h}_k^{\mathsf{H}} \mathbf{v}_k|^2$ since $|a_k|^2$ and $|\mathbf{h}_k^{\mathsf{H}} \mathbf{v}_k|^2$ are identically distributed, (b) follows from Jensen's inequality and the final expression (c) follows by noticing that the $\widehat{\mathbf{V}}$ is a deterministic function of $\widehat{\mathbf{H}}$ and therefore it is independent of $\mathbf{E}$. Therefore, we can write $\mathbb{E}[|\mathbf{e}_k^{\mathsf{H}} \widehat{\mathbf{v}}_j|^2] = \mathbb{E}[\widehat{\mathbf{v}}_j^{\mathsf{H}} \mathbb{E}[\mathbf{e}_k \mathbf{e}_k^{\mathsf{H}}] \widehat{\mathbf{v}}_j] = \sigma_e^2 \mathbb{E}[|\widehat{\mathbf{v}}_j|^2] = \sigma_e^2$, since $\widehat{\mathbf{v}}_j$ has unit norm by construction.

We briefly discuss the impact of imperfect CSIR.[2] Assume that the user terminals have an estimate $\widehat{a}_k$ of the useful signal coefficient a_k (e.g., obtained from downlink training symbols) such that $a_k = \widehat{a}_k + f_k$, where $\mathbb{E}[|f_k|^2] = \sigma_f^2$ and $\mathbb{E}[\widehat{a}_k^* f_k] = 0$.[3] We still assume that the interference plus noise variance Σ_k is accurately known on each frame. By repeating the same lower bounding argument of (6) we arrive at

$$R_k \geq \mathbb{E}\left[\log\left(1 + \frac{|\widehat{a}_k|^2 P}{\sigma_f^2 P + \Sigma_k M}\right)\right] \tag{8}$$

[1] With some abuse of notation, the term in the second line of (6) have the following meaning:

$$\mathbb{E}\left[\inf_{\zeta_k : \mathbb{E}[|\zeta_k|^2] \leq \Sigma_k} I(u_k^G; y_k | a_k, \Sigma_k)\right] \equiv \int \inf_{\zeta_k : \mathbb{E}[|\zeta_k|^2] \leq \sigma} I(u_k^G; y_k | a_k, \Sigma_k = \sigma) dF(\sigma)$$

where $F(\sigma)$ denotes the cdf of Σ_k.

[2] Notice that in this case model (3) still applies but the noise $\mathbf{W}$ accounts for the observation noise at the receivers plus the CSIT feedback noise.

[3] This condition holds when $\widehat{a}_k$ is the MMSE estimate of a_k given the observation.

By comparing (6) and (8) we notice that the effect of non-perfect CSIR is a) replacing a_k with $\widehat{a}_k$ and b) increasing the interference power by the term $\sigma_f^2 P$. We shall see in the next section that the multiplexing gain of the analog CSIT feedback scheme depends critically on the behavior of the term $\sigma_e^2 P$ for $P \to \infty$. Typically, we have that $\sigma_f^2 P = O(\sigma_e^2 P)$. Hence, the multiplexing gain for imperfect CSIT/perfect CSIR and for imperfect CSIT/imperfect CSIR are identical. This is in sharp contrast to point to point channels, in which perfect vs. imperfect CSIR may lead to very different high-SNR behaviors [11–13].

III. Comparison with quantized CSIT feedback

For the sake of simplicity we restrict to the case of perfect CSIR. Replacing the estimation error variance $\sigma_e^2 = (1 + \beta P/N_0)^{-1}$ in (7), we obtain

$$\begin{aligned} \Delta R_{\text{analog}} &\leq \log\left(1 + \frac{P/N_0}{1+\beta P/N_0}\right) \\ &\leq \log\left(1 + \frac{1}{\beta}\right) \end{aligned} \tag{9}$$

where we have upperbounded $(M-1)/M$ by 1 and where the last line is approached for large SNR.

Let us now consider digital feedback over the same channel. The rate gap obtained in [10, Theorem 1] and reported in (4) is further upperbounded by $\log(1+(P/N_0)\cdot 2^{-\frac{B}{M}})$. Let us assume (very unrealistically) that the digital feedback link can operate error-free and at capacity, i.e., it can reliably transmit $\log(1 + P/N_0)$ bits per symbol. For the same number of feedback channel periods, βM, the number of feedback bits per mobile is $B = \beta M \log_2(1 + P/N_0)$. Replacing this into the rate gap bound, we obtain:

$$\Delta R_{\text{quant.}} \leq \log\left(1 + \frac{P/N_0}{(1+P/N_0)^{\beta}}\right). \tag{10}$$

If $\beta = 1$ the quantized and analog feedback achieve essentially the same rate gap of at most 1 b/s/Hz. However, if $\beta > 1$, unlike the analog feedback case, the rate gap of the quantized feedback vanishes for $P/N_0 \to \infty$.

We conclude that for $\beta > 1$ the quantized feedback is far superior to the analog scheme. This conclusion finds an appealing interpretation in the context of sending an analog source via a noisy channel with minimal end-to-end distortion. In our case, the source is the Gaussian channel vector $\mathbf{h}_k$ and the noisy channel is the feedback AWGN channel with SNR P/N_0 that we have postulated in our model. It is well-known (see [14] and references therein) that when the source block length (M in our case) is constrained to be equal to the channel code block length ($\beta = 1$ in our case), then an optimal strategy to send a Gaussian source over a Gaussian channel with minimal end-to-end quadratic distortion consists of scaling the source symbols and sending them uncoded and unquantized through the channel. Hence, the fact that analog feedback cannot be essentially outperformed for $\beta = 1$ is expected. However, it is also well-known that if we are allowed to use a channel coding block length larger than the source block length ($\beta > 1$ in our case), the analog strategy is strictly suboptimal because the distortion with analog transmission scales as $1/\beta$ whereas it decreases exponentially with β (i.e., along the vector quantizer R-D curve) for digital transmission.

IV. Effects of CSIT feedback errors

We wish to investigate the impact on the above conclusions of the optimistic assumption that the quantized feedback channel can operate error-free and arbitrarily close to capacity. This assumption is particularly unrealistic because the feedback block coding length is very small (βM). In addition, the sensitivity to feedback delay (see Section V) is likely to require relatively simple codes to be used.

We shall consider a very simple CSIT feedback scheme that certainly represents a lower bound on the best quantized feedback strategy. The user terminals perform quantization using *Random Vector Quantizers* [10], and transmit the feedback bits using simple uncoded QAM. Furthermore, no intelligent mapping of the quantization bits onto the QAM symbols is used, Therefore, if even a single feedback bit from user k is erroneously received, the corresponding k-th CSIT vector is completely independent of the actual k-th channel vector or its quantization (this is because all of the quantization vectors in the codebook are random and are randomly assigned to the QAM symbols). Furthermore, since only uncoded QAM symbols are sent, error detection is not possible: the base station computes the beamforming matrix $\widehat{\mathbf{V}}$ based on the decoded feedback messages, even if they are incorrect.

We again use βM symbol periods to transmit the feedback bits. There is a non-trivial tradeoff between quantization and channel errors. In order to maintain a bounded gap, we must scale feedback at least as $(M-1)\log_2(1 + P/N_0)$, which we approximate as $M \log_2 P/N_0$ for simplicity. Therefore, let us consider using $B = \alpha M \log_2 P/N_0$ for $\alpha \geq 1$. We send these B bits in βM symbol periods, and thus we send $\frac{\alpha}{\beta}\log_2(P/N_0)$ bits per QAM feedback symbol.

From [15], using the fact that the QAM constellation size is equal to $L = (P/N_0)^{\frac{\alpha}{\beta}}$, we have the following upper bound to the symbol error probability for QAM modulation:

$$P_s \leq 2\exp\left(-\frac{3}{2}\left(\frac{P}{N_0}\right)^{1-\alpha/\beta}\right) \tag{11}$$

For $\alpha = \beta$ (which means trying to signal at capacity with uncoded modulation!) P_s does not decreases with SNR and the system performance is very poor. However, for $\alpha/\beta < 1$, which corresponds to transmitting at a constant fraction of capacity, then $P_s \to 0$ as $P/N_0 \to \infty$. The upper bound on the error probability of the whole quantized vector (transmitted in βM symbols) is given by $P_{e,fb} = 1-(1-P_s)^{\beta M}$. A lower bound on the achievable ergodic rate is obtained by assuming that when a feedback error occurs for user k its SINR is zero (because user k receives a large amount of interference due to the useless quantization vector available to the transmitter), while if the feedback error does not occur its rate is given $R_k^{\text{ZF}} - \Delta R_{\text{quant.}}$, that is, the rate of ideal ZF decreased by the (upper bound to) the rate gap. It follows that the ergodic rate of user k in the presence of quantized feedback with errors is upperbounded by

$$R_k \geq (1-P_s)^{\beta M}\left(R_k^{\text{ZF}} - \log\left(1 + (P/N_0)^{1-\alpha}\right)\right) \tag{12}$$

Choosing $1 < \alpha < \beta$ we achieve both vanishing P_s and vanishing $\Delta R_{\text{quant.}}$ as $P/N_0 \to \infty$. Thus, even under this very simple CSIT feedback scheme the optimal ZF performance can be eventually approached for sufficiently high SNR.

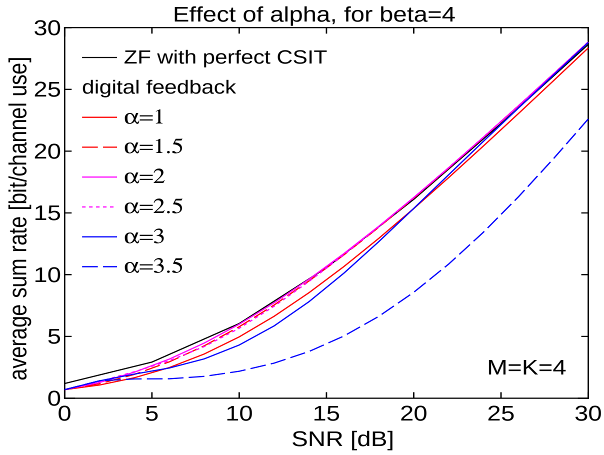

Fig. 1. Quantized feedback with QAM modulation.

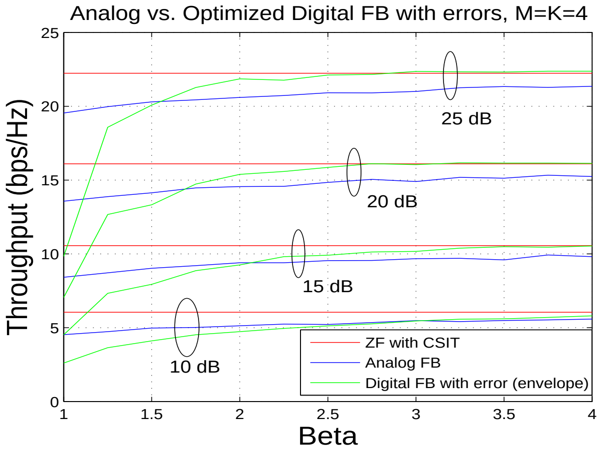

Fig. 2. Quantized vs. analog feedback.

Fig. 1 shows the ergodic rate achieved by ZF beamforming with quantized CSIT and QAM feedback trasmission for $M = K = 4$, independent Rayleigh fading, $\beta = 4$ and different values of α. It is noticed that by proper design of the feedback parameters the performance can be made very close to the ideal CSIT case.

Fig. 2 shows the rate achieved by ideal CSIT, by analog feedback and quantized digital feedback with QAM modulation at different SNRs as a function of the feedback rate β channel uses per channel coefficient. The result for digital feedback is optimized with respect to α. We notice that for β sufficiently large the digital feedback eventually outperforms the analog feedback. The cross-over point decreases as SNR increases.

V. Effects of CSIT feedback delay

We consider now the case of analog feedback when each entry of $\mathbf{H}$ evolves independently (in the block-fading way described above) according to the same complex circularly symmetric Gaussian stationary ergodic random process, denoted by $\{h(t)\}$, with mean zero, variance 1 and power spectral density (Doppler spectrum) denoted by $S_h(\xi)$, $\xi \in [-1/2, 1/2]$.

Because of stationarity, without loss of generality we can focus on $t = 0$. We are interested in the linear MMSE estimation of $h(t)$ from the observation $\{g(\tau) : \tau = -\infty, t - d\}$ where, following the analog feedback model (3), we let $g(\tau) = h(\tau) + w(\tau)$, with $w(\tau)$ i.i.d. $\sim \mathcal{CN}(0, \delta)$ and $\delta = \frac{N_0}{\beta P}$. In particular, we consider the case of 1-step prediction ($d = 1$) and the case of filtering ($d = 0$). From classical Wiener filering theory [16], we have that the prediction error is given by

$$\epsilon_1(\delta) = \exp\left(\int_{-1/2}^{1/2} \log(\delta + S_h(\xi))d\xi\right) - \delta \tag{13}$$

and that the filtering MMSE is given by

$$\epsilon_0(\delta) = \frac{\delta\epsilon_1(\delta)}{\delta + \epsilon_1(\delta)} \tag{14}$$

that is, the filtering error is the harmonic mean between the observation noise variance δ and the prediction error $\epsilon_1(\delta)$. Notice that if $\epsilon_1(\delta) = 1$ (e.g., the channel process is i.i.d. so that the observation of the past is useless), then $\epsilon_0(\delta) = \delta/(1 + \delta)$, which is the same expression that we have used in the Section II for the case of i.i.d. block-fading and delay-free feedback.

We shall discuss the rate gap bound (7) letting $\sigma_e^2 = \epsilon_d(N_0/(\beta P))$ for $d = 0, 1$, under different assumptions on the fading process $\{h(t)\}$. We distinguish two cases: Doppler process and regular process. We say that $\{h(t)\}$ is a Doppler process if $S_h(\xi)$ is strictly band-limited in $[-F, F]$, where $F < 1/2$ is the maximum Doppler frequency shift, given by $F = \frac{v f_c}{c} T_f$, where v is the mobile terminal speed (m/s), f_c is the carrier frequency (Hz), c is light speed (m/s) and T_f is the frame duration (s). Furthermore, a Doppler process must satisfy $\int_{-F}^{F} \log S_h(\xi)d\xi > -\infty$. This condition holds for most (if not all) channel models usually adopted in the wireless mobile communication literature, where typically within the support $[-F, F]$ the Doppler spectrum has no spectral nulls (see [17] and references therein). Following [13], we say that $\{h(t)\}$ is a regular process if $\epsilon_1(0) > 0$. In particular, a process satisfying the Paley-Wiener condition [16] $\int_{-1/2}^{1/2} \log S_h(\xi)d\xi > -\infty$ is regular.

A Doppler process satisfying our assumptions has prediction error

$$\epsilon_1(\delta) = \delta^{1-2F} \exp\left(\int_{-F}^{F} \log(\delta + S_h(\xi))d\xi\right) - \delta \tag{15}$$

No feedback delay ($d = 0$). In this case

$$P\sigma_e^2 = \frac{N_0}{\beta} \frac{\epsilon_1\left(\frac{N_0}{\beta P}\right)}{\frac{N_0}{\beta P} + \epsilon_1\left(\frac{N_0}{\beta P}\right)} \tag{16}$$

Hence, $\lim_{P\to\infty} P\sigma_e^2 = \frac{N_0}{\beta}$ for both Doppler and regular processes. For the latter, this is clear from the fact that $\epsilon_1(0) > 0$. For the former, this follows from (15). Applying Jensen's inequality and the fact that $\int S_h(\xi)d\xi = 1$, we arrive at the upper bound

$$\epsilon_1\left(\frac{N_0}{\beta P}\right) \leq \left(\frac{N_0}{\beta P}\right)^{1-2F} \left[\left(\frac{1}{2F} + \left(\frac{N_0}{\beta P}\right)\right)^{2F} - \left(\frac{N_0}{\beta P}\right)^{2F}\right] \tag{17}$$

Using the fact that $\log$ is increasing, we arrive at the lower bound

$$\epsilon_1\left(\frac{N_0}{\beta P}\right) \geq \left(\frac{N_0}{\beta P}\right)^{1-2F} \left[\exp\left(\int_{-F}^{F} \log S_h(\xi)d\xi\right) - \left(\frac{N_0}{\beta P}\right)^{2F}\right] \tag{18}$$

These bounds yield that $\epsilon_1(N_0/\beta P) = \kappa P^{-(1-2F)} + O(1/P)$ for some constant κ. Hence, $\epsilon_1 = O(P^{-(1-2F)})$ while $\delta = O(1/P)$, and the limits holds.

We conclude that in the case of no feedback delay the estimation error is essentially dominated by the instantaneous observation and not much improvement can be expected by taking into account the channel memory. On the other hand, this means also that the fading time-correlation has little impact on the performance provided that the feedback is fast enough.

Feedback delay ($d = 1$). In this case, the behavior of Doppler versus regular processes is radically different. For Doppler processes, using (17) and (18), we have that $P\sigma_e^2 = P\epsilon_1(N_0/\beta P) = \kappa P^{2F} + O(1)$. It follows that the achievable rate sum is lowerbounded by

$$\sum_{k=1}^{M} R_k \geq M(1-2F)\log P + O(1) \tag{19}$$

or, equivalently, that the multiplexing gain of the system is $M(1-2F)$.

For regular processes, on the contrary, we have that $P\sigma_e^2 \geq P\epsilon_1(0) = O(P)$. Hence, the rate gap grows like $\log P$ and the achieved multiplexing gain is zero. [4]

In conclusions, the most noteworthy result of this analysis is that under common fading models (Doppler processes), the analog feedback scheme achieves a potentially high multiplexing gain even with realistic, noisy and delayed feedback. Notice for example that with mobile speed $v = 50$ km/h, $f_c = 2$ GHz, and frame duration 1 ms, we have $F = 0.0926$. With $M = 4$ antennas we achieve a yet respectable pre-log factor equal to 3.26 instead of 4.

Figs. 3 and 4 show the achievable ergodic rates for the Jakes' "J_0" correlation (strictly band-limited) and the Gauss-Markov AR-1 correlation (regular process) for different first-lag correlation values. For the AR-1 process with $d = 1$ the system becomes interference limited. On the contrary, the performance under Jakes' model degrades gracefully as the user mobility (Doppler bandwidth) increases.

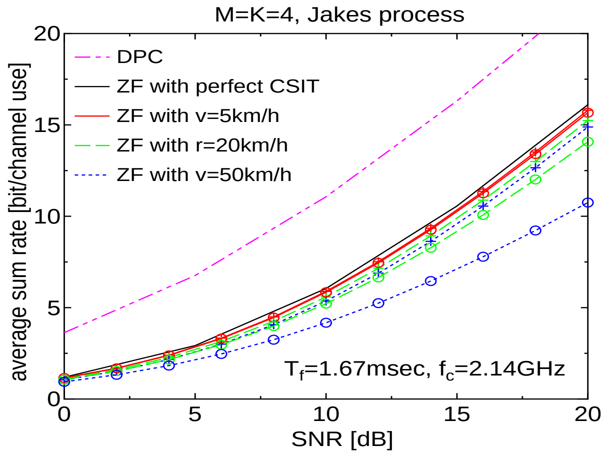

Fig. 3. Rates with feedback delay and Jakes' correlation.

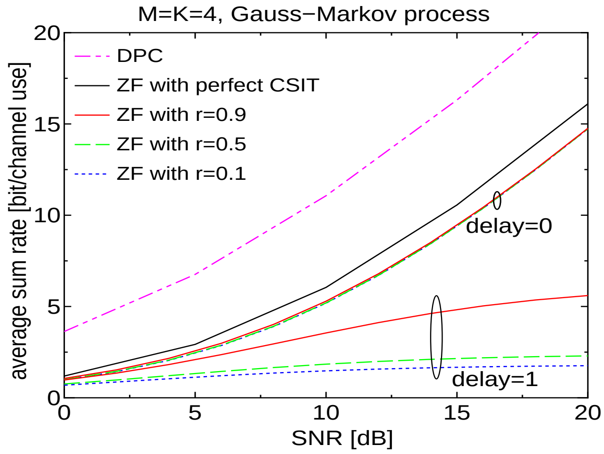

Fig. 4. Rates with Gauss-Markov AR-1 correlation.

References

[1] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. on Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[2] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. on Inform. Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.

[3] P. Viswanath and D.N.C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. on Inform. Theory*, vol. 49, no. 8, pp. 1912–1921, 2003.

[4] W. Yu and J.M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. on Inform. Theory*, vol. 50, no. 9, pp. 1875–1892, 2004.

[5] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*.

[6] S.A. Jafar and A.J. Goldsmith, "Isotropic Fading Vector Broadcast Channels: The Scalar Upper Bound and Loss in Degrees of Freedom," *IEEE Trans. on Inform. Theory*, vol. 51, no. 3, 2005.

[4] It is interesting to notice here the parallel with the results of [13] on the high-SNR capacity of the single-user scalar ergodic stationary fading channel with no CSIR and no CSIT, where it is shown that for a class of *non-regular* processes that includes the Doppler processes defined here, the high-SNR capacity grows like $\mathcal{L}\log P$, where $\mathcal{L}$ is the Lebsegue measure of the set $\{\xi \in [-1/2, 1/2] : S_h(\xi) = 0\}$. In our case, it is clear that $\mathcal{L} = 1 - 2F$.

[7] T.L.Marzetta and B.M.Hochwald, "Fast Transfer of Channel State Information in Wireless Systems," *Submitted to "IEEE Transactions on Signal Processing"*, June 2004.

[8] M. Kobayashi and G. Caire, "Joint Beamforming and Scheduling for a Multi-Antenna Downlink with Imperfect Transmitter Channel Knowledge," *to appear on IEEE J. Select. Areas Commun.*, June 2007.

[9] T.A. Thomas, K.L. Baum, and P. Sartori, "Obtaining channel knowledge for closed-loop multi-stream broadband MIMO-OFDM communications using direct channel feedback," in *IEEE Global Telecommunications Conference, 2005. GLOBECOM '05*, November 2005, vol. 6.

[10] N.Jindal, "MIMO Broadcast Channels with Finite Rate Feedback," *To appear IEEE Trans. Information Theory*, November 2006.

[11] B. Hassibi and B.M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. on Inform. Theory*, vol. 48, no. 7, pp. 1804–1824, 2002.

[12] M. Médard, "Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel," *IEEE Trans. on Inform. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[13] A. Lapidoth, "On the Asymptotic Capacity of Stationary Gaussian Fading Channels," *IEEE Trans. on Inform. Theory*, vol. 51, no. 2, pp. 437, 2005.

[14] M.Gastpar, B.Rimoldi, and M.Vetterli, "To Code, or not to code : Lossy source-channel communication revisited," *IEEE Trans. on Inform. Theory*, vol. 49, May 2003.

[15] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

[16] H.V. Poor, *An introduction to signal detection and estimation*, Springer-Verlag New York, Inc. New York, NY, USA, 1994.

[17] E. Biglieri, J. Proakis, S. Shamai, and D. di Elettronica, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. on Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.