Physical layer reliability vs ARQ in MIMO block-fading channels

Marie Zwingelstein-Colin* and Mérouane Debbah †
*University Lille Nord de France
IEMN, UMR 8520, F-59313 Valenciennes, France
†SUPELEC
Alcatel-Lucent Chair on Flexible Radio
3 rue Joliot-Curie 91192 GIF SUR YVETTE CEDEX France

***Abstract*—In today's wireless communication systems, automatic repeat request (ARQ) is implemented at the MAC layer in order to retransmit packets that have been erroneously transmitted at the physical (PHY) layer. Following a joint PHY-MAC design, information provided by the ARQ scheme can be exploited at the PHY layer in order to improve the system's performance. This paper extends the work presented in [1] in the context of SISO channels to the context of MIMO block-fading channels. Based on statistical channel knowledge at the transmitter, it provides an analysis of the natural tradeoff that exists between the PHY layer transmission rate and the number of ARQ retransmissions. We derive a very accurate analytical formulation of the optimum transmission rate (and, equivalently, the optimum PHY packet error-rate) that maximizes the goodput, as a function of the system parameters, namely the SNR, the number of antennas and the diversity order of the channel. Interestingly, we find that the PHY layer has to be made more reliable for MIMO channels than for SISO channels, and also that the MIMO ARQ system is less sensitive to a wrong choice of the rate of transmission than the SISO ARQ system.**

I. Background and motivation

Since several years, cross-layer design have drawn the research community's attention to enhance wireless systems performance. Cross-layer design can follow two approaches: firstly, the sharing of the PHY and MAC layers'knowledge with higher levels, also known as cross-layer networking [2]; secondly, the use of the MAC layer to optimize the PHY layer. This paper concentrates on the second approach, in the context of MIMO block-fading channels whith ARQ retransmission. ARQ is a well known feeddback-based technique that consists in retransmitting packets in case of unsuccesfull transmissions, which is particularly advantageous when the instantaneous channel quality is unknown to the transmitter [3]. Within ARQ, a natural tradeoff exists between the PHY layer reliability and the ARQ retransmission rate: on the one hand, if the PHY layer tries to transmit at a high rate, a lot of packets will be detected as erroneous and a lot of retransmissions will be necessary, which degrades the rate at which information is successfully transmitted - also known as the *goodput*. On the other hand, if the PHY layer is made very reliable, almost no retransmission will be necessary, but the rate of transmission ensuring such high reliability will be quite low. [1] and [4] have analyzed this tradeoff in the context of SISO block-fading channels. In particular, they have studied how the PHY transmission rate should be adjusted in order for the goodput to be maximized. The work presented in this paper can be seen as a generalization of [1] to the MIMO case. Previous work on the MIMO ARQ channel has majoritarily focused on code design: the fundamental tradeoff between diversity, throughput and maximum number of retransmissions was derived in [5] (after its derivation is the SISO case in [6]) for discrete input constellations. Extensions to multi-bit feedback and imperfect feedback have been derived in [7] and [8] respectively. In this paper, we focus on the performance analysis of a MIMO ARQ block-fading channel that, based on statistical channel knowledge, initiates rate adaptation to maximize the goodput. In particular, we derive a very accurate analytical formulation of the optimum PHY transmission rate (and, equivalently, the optimum PHY packet error-rate) that maximizes the goodput, as a function of the system parameters, namely the SNR, the number of antennas and the diversity order of the channel. Interestingly, we find that the PHY layer has to be made more reliable for MIMO channels than for SISO channels, and also that the MIMO ARQ system is less sensitive to a wrong choice of the rate of transmission than the SISO ARQ system.

II. Channel model

We consider a MIMO Rayleigh fading channel, with n_a antennas at the transmitter and n_a antennas at the receiver. The channel is assumed to change independently from one block of data symbols to the other, and to remain constant within each block (block-fading scenario). One codeword (packet) is supposed to span K successive blocks, where K is representative of the time (or frequency) selectivity of the channel (if time selectivity is considered, the value $K = 1$ corresponds to a slow-fading scenario while higher values of K englobe the fast-fading case. If frequency selectivity is considered, $K = 1$ corresponds to a flat channel, whereas higher values reflect a dispersive channel). The channel is assumed to be perfectly known at the receiver, but only statistically known at the transmitter, which is a realistic assumption, especially in a fast-fading and/or dispersive environment. We consider a simple ARQ scheme with perfect error detection at the receiver: after processing a received codeword, and in case successfull decoding is detected, a one-bit ACK signal is sent back to the receiver for aknowledgement, whereas a one-bit NAK is fed back if unsuccessfull decoding is detected. In the latter, the transmitter retransmits the packet, until positive ACK. No attempt is made by the receiver to correct errors based on (potentially) previous versions of erroneously received packet, and the cost of the ACK/NAK feedback will be neglected.

At block number k, the l-th received data symbol can be written:

$$\mathbf{y}_l = \sqrt{\frac{\mathsf{SNR}}{n_a}}\mathbf{H}_k\mathbf{x}_l + \mathbf{w}_l$$

where $\mathbf{x}_l \in \mathcal{C}^{n_a}$ is the l-th transmitted data symbol ($\mathbb{E}\{\mathbf{x}_l\mathbf{x}_l^H\} = 1$), $\mathbf{w}_l \sim \mathcal{CN}(0, \mathbf{I}_{n_a})$ represents the additive Gaussian noise and $\mathbf{H}_k \in \mathcal{C}^{n_a \times n_a}$ is the channel matrix at block number k whose independent elements $\sim \mathcal{CN}(0, 1)$.

We are interested in analyzing the tradeoff between the transmission rate and the packet error rate, when one wants to optimize the goodput. For this purpose, we make the assumption that a strong channel code is used, which leads to the source of error for the transmission of a packet being limited only to the outage of the mutual information. In that case, the packet error rate ϵ is simply

$$\epsilon = \mathsf{Pr}\left[I\left(K, n_a, \mathsf{SNR}\right) \leq R\right] \tag{1}$$

where R is the transmission rate in bit/symbol and $I\left(K, n_a, \mathsf{SNR}\right) = \frac{1}{K}\sum_{k=1}^{K}\log\left|\mathbf{I} + \frac{\mathsf{SNR}}{n_a}\mathbf{H}_k\mathbf{H}_k^H\right|$ is the channel mutual information. Note that no attempt is made to optimize the input covariance matrix. Only rate adaptation is performed based on channel's statistics. We denote X the random variable which represents the number of transmission attempts of each packet. In accordance with the block-fading scenario, successive blocks -and thus packets - errors are independent, so the goodput is simply the ratio of R over the average number of transmission attempts $\mathbb{E}\{X\}$:

$$G = \frac{R}{\mathbb{E}\{X\}}$$

Under error-free acknowledgment, the probability of i packet transmission is simply $\mathsf{Pr}\left[X = i\right] = \epsilon^{i-1} \times (1 - \epsilon)$, that is X is a geometrically distributed random variable of mean $\frac{1}{1-\epsilon}$. Consequently, the goodput can be expressed

$$G = R \times (1 - \epsilon) \tag{2}$$

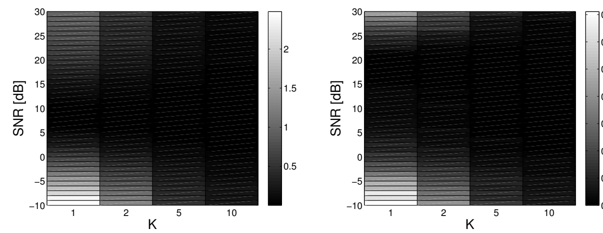
where the dependency of R with ϵ follows equation (1).

III. Gaussian approximation

The analysis of equation (2) requires first to examine the dependency of R with ϵ, and also to investigate how this dependency evolves with the system parameters n_a, K and SNR. For this purpose, we propose to approximate the mutual information $I\left(K, n_a, \mathsf{SNR}\right)$ by a Gaussian random variable with the same mean and variance, noted $\mu\left(K, n_a, \mathsf{SNR}\right)$ and $\sigma^2\left(K, n_a, \mathsf{SNR}\right)$ respectively[1]. Including this Gaussian approximation to the packet error rate equation (1) we get[2]

$$\epsilon \approx \begin{cases} \mathrm{Q}\left(\frac{\mu(K,n_a,\mathsf{SNR})-R}{\sigma(K,n_a,\mathsf{SNR})}\right) & (\mu \geq R) \\ 1 - \mathrm{Q}\left(\frac{R-\mu(K,n_a,\mathsf{SNR})}{\sigma(K,n_a,\mathsf{SNR})}\right) & (\mu \leq R) \end{cases} \tag{3}$$

(obviously, $\epsilon = \frac{1}{2}$ when $R = \mu$).



(a) SISO (b) MIMO 2 × 2

Fig. 1: β_1 vs SNR and K

As the number of fading coefficients increases with n_a and K, the central limit theorem tells us that the Gaussian approximation will be accurate for high n_a and K. But how high? To answer this question, we quantify the accuracy of the approximation by comparing the values of two caracteristic statistical parameters, namely

$$\beta_1\left(K, n_a, \mathsf{SNR}\right) \overset{\text{def}}{=} \frac{\left[\mu_3\left(K, n_a, \mathsf{SNR}\right)\right]^2}{\left[\mu_2\left(K, n_a, \mathsf{SNR}\right)\right]^3}$$

and

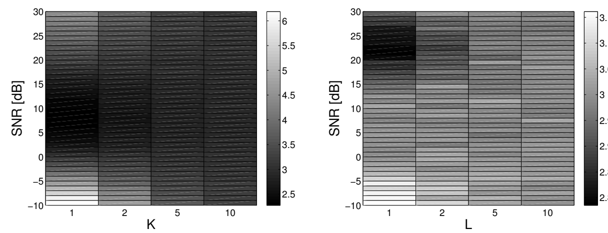
$$\beta_2\left(K, n_a, \mathsf{SNR}\right) \overset{\text{def}}{=} \frac{\mu_4\left(K, n_a, \mathsf{SNR}\right)}{\left[\mu_2\left(K, n_a, \mathsf{SNR}\right)\right]^2}$$

where μ_i denotes the central moment of order i of the distribution, for the actual and for the Gaussian distributions (it can be shown that $\beta_1 = 0$ and $\beta_2 = 3$ for a Gaussian distribution, independently of the values of its central moments). Figures 1 and 2 represents the values of β_1 and β_2 for $K \in \{1, 2, 5, 10\}$, as a function of the SNR and $n_a \in \{1, 2\}$. The closer β_1 and β_2 are to 0 and 3 respectively, the more accurate the Gaussian approximation is. In the SISO case, β_1 and β_2 are not very close to 0 and 3 respectively, especially for $K \in \{1, 2\}$ and SNR values < 0 dB or > 20 dB. The Gaussian approximation is thus not very accurate, and it can only be used as an indicator in the SISO case. Now looking at the MIMO 2×2 case, we see that the accuracy of the approximation is greatly improved, even if the selectivity of the channel is reduced to $K = 1$. The approximation can be shown to be even more accurate when $n_a = 4$.

Note: In the following, all dotted curves will correspond to approximated data according to equation (3), whereas data obtained by monte-carlo simulations will be represented as continuous curves.

[1] Obviously we have $\mu\left(K, n_a, \mathsf{SNR}\right) = \mu\left(1, n_a, \mathsf{SNR}\right)$ and $\sigma^2\left(K, n_a, \mathsf{SNR}\right) = \frac{\sigma^2(1, n_a, \mathsf{SNR})}{K}$.

[2] $\mathrm{Q}\left(x\right)$ denotes the complementary error function under Gaussian statistics.



(a) SISO (b) MIMO 2×2

Fig. 2: β_2 vs SNR and K

IV. Tradeoff analysis

In order to analyze the tradeoff between the packet error-rate ϵ and the transmission rate R, when considering the optimization of the goodput G, we first examine the relationship between ϵ and R, and then consider the dependency of G with R and with ϵ.

A. ϵ versus R

The relationship between ϵ and R is characterized by equations (1) (rigorous) and (3) (approximated). Figure (3) shows the curve ϵ versus R for SNR = 5dB and $K \in \{1, 2, 10, \infty\}$, in the SISO, MIMO 2×2 and MIMO 4×4 cases.

For $K \to \infty$, it can be noted that ϵ behaves like $\mathrm{Q}\left(\frac{\mu - R}{\sigma} \times \infty\right) = 0$ for $R < \mu$, and like $1 - \mathrm{Q}\left(\frac{R-\mu}{\sigma} \times \infty\right) = 1$ for $R < \mu$, hence the step in the curve at absissa $R = \mu$ in figure (3), for $K = \infty$. This result can also be analyzed by noting that $\lim_{K\to\infty} I\left(K, n_a, \mathsf{SNR}\right) = \mu$ is the channel ergodic capacity, C_{erg}, and that strong channel coding ensures no error at the PHY layer as long as $R < C_{\text{erg}}$.

For $K < \infty$, we see that the increase in ϵ is monotonous, whith an inflexion point at $R = \mu$. The only difference between the SISO, MIMO 2×2 and MIMO 4×4 cases for the approximated curves reside in the absissa of the inflexion point. Otherwise, the curves look similar. For the exact curves, however, we can see that they are closer to the approximated curves in the MIMO case than in the SISO case, which is consistent with the analysis of section III.

B. G

Approximated expressions of $G(R)$ and $G(\epsilon)$ can be directly obtained by adequately mixing equations (2) and (3). The curves are plotted in figures 4 and 5 respectively, both exact and approximated, for the SISO, MIMO 2×2 and MIMO 4×4 channels. Consider first the limit case $K \to \infty$, for which we have

$$G_{K\to\infty} \overset{\text{def}}{=} \lim_{K\to\infty} G \approx \begin{cases} R\left(1-0\right) = R & (\mu \geq R) \\ R\left(1-1\right) = 0 & (\mu \leq R) \end{cases}$$

, which, again, simply reflects the fact that error free transmission is achieved at the PHY layer when the transmission rate $R < C_{\text{erg}}$. Considering the dependency of G with ϵ, we can also write: $G_{K\to\infty} = \mu\left(1-\epsilon\right)$. $G_{K\to\infty}$ is thus linearly decreasing from $\mu = C_{\text{erg}}$ when $\epsilon = 0$ to 0 when $\epsilon = 1$. In

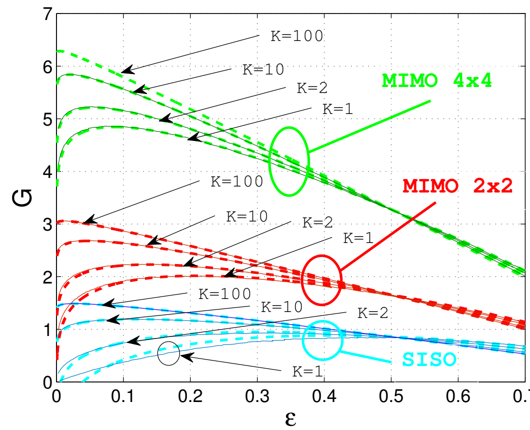


Fig. 5: G versus ϵ, SNR= 5dB.

fact, due to the stepwise nature of the function ϵ versus R when $K \to \infty$, only the two extreme points $(0, \mu)$ and $(1, 0)$ in figure 5 are achievable.

For $K < \infty$, we can see on figure 4 that $G\left(R\right) \approx R$ for small values of R. Then, as far as R increases, errors appears at the PHY layer and thus $G < R$. G achieves its maximum value G^* at $R = R^*$, and then decreases up to 0 when the PHY layer is no more reliable and consequently the number of ARQ retransmission tends to ∞. Note that all curves coincide at the point $\left(\mu, \frac{\mu}{2}\right)$ (independently of K), which is normal since at $R = \mu$ we have $\epsilon = \frac{1}{2}$. Note also that we always have $R^* < \mu$.

When comparing the behavior of G for the MIMO and the SISO channels, we observe that the distance between R^* and μ is increasing with the number of antennas. In figure 5, which plots G versus ϵ, we can also see that the value of ϵ that maximizes G is smaller in the MIMO case. Hence, the PHY layer has to be made more reliable for a MIMO channel than for a SISO channel when one attempts to optimize the goodput.

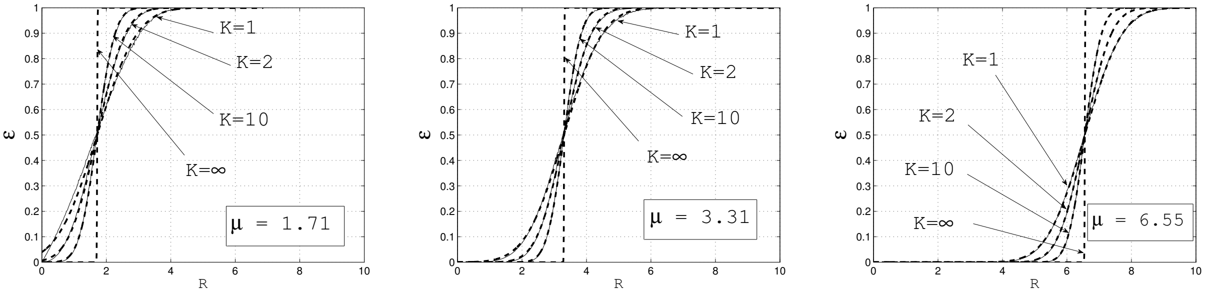
V. Goodput optimization

From a system level point of view, we have shown in section IV that it is important to make the PHY layer just as reliable as necessary in order to maximize the goodput. For this purpose, we now examine how the optimal values

$$\begin{aligned} \epsilon^* &= \arg \max_{\epsilon} \quad G \\ R^* &= \arg \max_{R} \quad G \end{aligned} \tag{4}$$

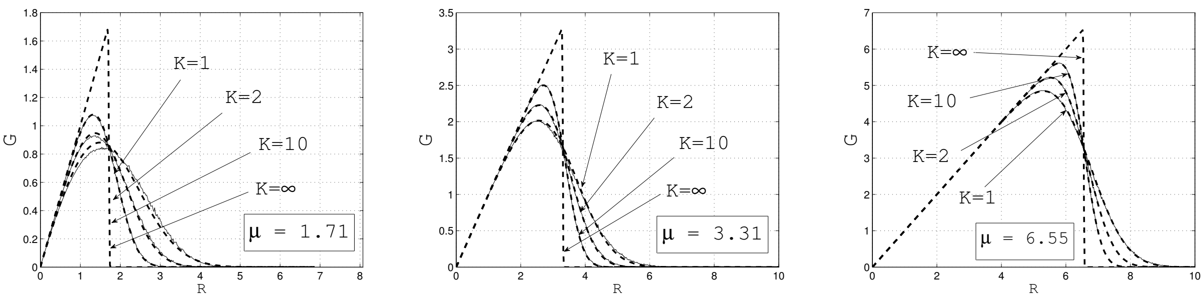
evolve with the system parameters K, n_a and SNR. Approximated values of R^* and ϵ^* are easily obtained by derivation of the approximated expressions of G:

$$G(R) \approx R\left(1 - \mathrm{Q}\left(\frac{\mu\left(K, n_a, \mathsf{SNR}\right) - R}{\sigma\left(K, n_a, \mathsf{SNR}\right)}\right)\right)$$
$$G(\epsilon) \approx \left(\mu\left(K, n_a, \mathsf{SNR}\right) - \sigma\left(K, n_a, \mathsf{SNR}\right)\mathrm{Q}^{-1}\left(\epsilon\right)\right)\left(1 - \epsilon\right)$$



(a) SISO (b) MIMO 2×2 (c) MIMO 4×4

Fig. 3: ϵ versus R, $\mathsf{SNR} = 5$ dB.



(a) SISO (b) MIMO 2×2 (c) MIMO 4×4

Fig. 4: G versus R, $\mathsf{SNR} = 5$ dB.

whith respect to R and ϵ. After some manipulation[3], we obtain that R^* is solution of

$$1 - \mathrm{Q}\left(\frac{\mu_K - R^*}{\sigma_K}\right) - \frac{R^*}{\sigma_K\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu_K - R^*}{\sigma_K}\right)^2} = 0 \qquad (5)$$

and that and ϵ^* is solution of

$$\sqrt{2\pi}\left(1-\epsilon^*\right) e^{\frac{1}{2}\left[\mathrm{Q}^{-1}(\epsilon^*)\right]^2} + \mathrm{Q}^{-1}\left(\epsilon^*\right) = \frac{\mu_K}{\sigma_K} \qquad (6)$$

where we have used $\mu_K = \mu\left(K, n_a, \mathsf{SNR}\right)$ and $\sigma_K = \sigma\left(K, n_a, \mathsf{SNR}\right)$ in order to make the equations clearer. Note than in the limit case $K \to \infty$ we have $R^* \to G^*$, $\epsilon^* \to 0$ and $G^* \to \mu\left(1, n_a, SNR\right)) = C_{\text{erg}}$.

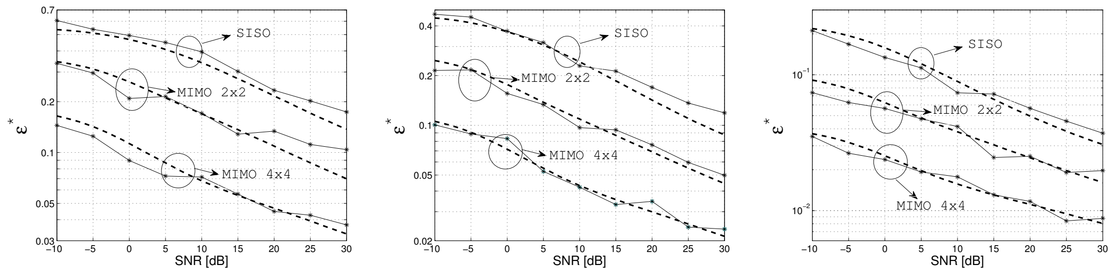
Figure 6 presents the curve ϵ^* versus SNR for $n_a \in \{1, 2, 4\}$ and $K \in \{1, 2, 10\}$. We see that ϵ^* is decreasing with SNR and K. We also see that the value of ϵ^* is quite large; for example, at 0 dB, $\epsilon^* > 0.1\ \forall K \leq 10$. Furthermore, ϵ^* is much smaller in the MIMO case than in the SISO case, and it tends to be less much smaller as SNR and K increases. Considering now G^* and R^* - which are both increasing functions of SNR, K and n_a -, we present on figure 7 the ratios $\frac{G^*}{R^*}$ and $\frac{G^*}{C_{\text{erg}}}$ versus SNR for $n_a \in \{1, 2, 4\}$ and $K \in \{1, 2, 10, 100\}$. We see that these ratios are also increasing functions of SNR, n_a and K, which means that the rate of increase of G^* is higher than those of R^* and C_{erg}. This is due to the fact that there are two effects that contributes to the increase of $G^* = R^*\left(1-\epsilon^*\right)$: the increase of R^* and the decrease of ϵ^*. We conclude the discussion on figure 7 by noting that, as for given SNR and K the two ratios are increasing with n_a, the ARQ scheme under cross-layer control is more efficient in the MIMO case than in the SISO case.

In order to evaluate the impact on the goodput G of a non optimal choice of ϵ, we present in figure 8 the ratio $G_{x\epsilon}/G^*$ as a function of the SNR, where $G_{x\epsilon} \stackrel{\text{def}}{=} G|_{\epsilon = x \times \epsilon^*}$, for $x = \frac{1}{5}$, $n_a \in \{1, 2, 4\}$ and $K \in \{1, 2, 10, 100\}$. We see that the ratio is increasing with n_a and K, which means that MIMO and highly selective channels are more robust to a wrong choice of ϵ than SISO and/or flat, slow-fading channels.

VI. Conclusion

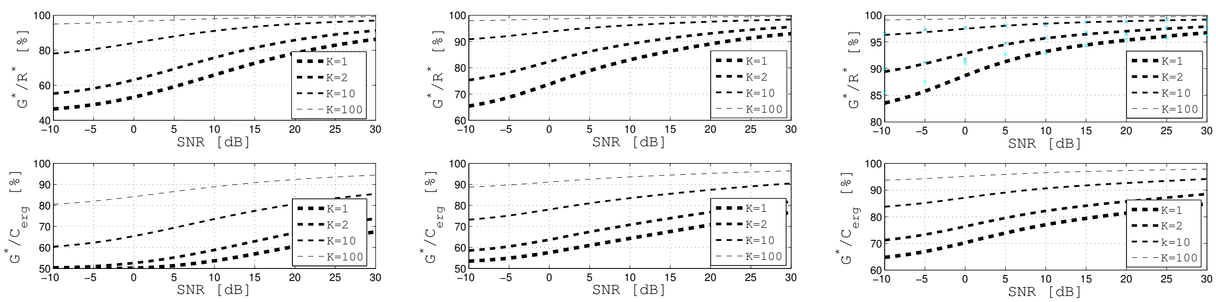
In this paper, we have examined the properties of the natural tradeoff that exists in MIMO block-fading channels between the packet-error rate at the PHY layer and the number of ARQ retransmissions, when rate adaptation is performed - based on statistical information about the channel - to optimize the goodput. The general findings are that the optimal packet error-rate at the PHY layer is smaller for MIMO channels than for SISO channels - and that it diminishes with the number of antennas -, and that MIMO channels are less sensitive to a mismatch in the packet error-rate than SISO channels. We also show that the ARQ scheme under cross-layer control

[3]The derivative of $\mathrm{Q}^{-1}(x)$ is $\left(\mathrm{Q}^{-1}\right)'(x) = -\sqrt{2\pi}\exp\left(\frac{\left[\mathrm{Q}^{-1}(x)\right]^2}{2}\right)$.



(a) $K = 1$ (b) $K = 2$ (c) $K = 10$

Fig. 6: ϵ^* as a function of the SNR.



(a) SISO (b) MIMO 2×2 (c) MIMO 4×4

Fig. 7: $\frac{G^*}{R^*}$ and $\frac{G^*}{C_{\text{erg}}}$ versus SNR

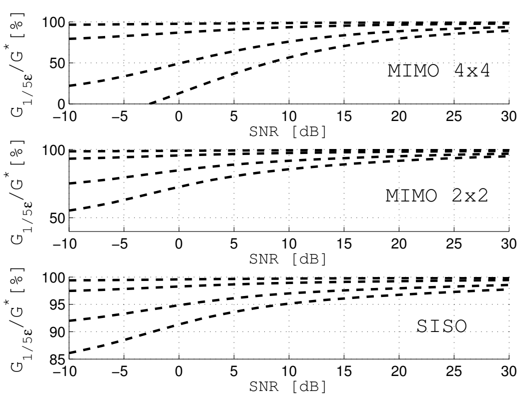


Fig. 8: $\frac{G_{\frac{1}{5}\epsilon}}{G^*}$ versus SNR for $n_a \in 1, 2, 4$ and $K \in \{1, 2, 10, 100\}$ (top: $K = 100$, bottom: $K = 1$).

is more efficient in the MIMO case, in the sense that the achievable goodput is closer to the ergodic capacity than in the SISO case. Some potentially interresting extensions of this work are considering a more interesting scenario where, first, the feedback channel introduces some errors, and, second a more sophisticated ARQ scenario is envisaged. More over, it is interesting to extend this work to the case of multiple communicating pairs that interfer to each other, such as the MIMO multiple access channel.

REFERENCES

[1] P. Wu and N. Jindal, "Coding versus ARQ in fading channels: how reliable should the phy be?" in *Proc. of IEEE GLOBECOM*, 2009.

[2] R. T. S. Shakkottai, S. and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Communications Magazine*, vol. 41(10), pp. 74–80, 2003.

[3] D. Chase, "Class of algorithms for decoding block codes with channel measurment information," *IEEE Transactions on Information Theory*, vol. 18 (1), pp. 170–182, 1972.

[4] P. Wu and N. Jindal, "Performance of hybrid-ARQ in block fading channels: A fixed outage probability analysis," *IEEE Transactions on Communications*, vol. 58 (4), pp. 1129–1141, 2010.

[5] A. R. Chuang, A. Guillen i Fabregas and I. L. K. Collings, "Optimal throughput-diversity-delay tradeoff in MIMO ARQ block-fading channels," *IEEE Transactions on Information Theory*, vol. 9, pp. 3968–3986, 2008.

[6] C. G. Gamal, H. E. and M. E. Damen, "The MIMO ARQ channel: diversity-multiplexing-delay tradeoff," *IEEE Transactions on Information Theory*, vol. 54 (9), pp. 3968–3986, 2006.

[7] R. L. K. G. i. F. A. Nguyen, K. D. and N. Letzepis, "MIMO ARQ systems with multi-level feedback," in *Proc. of ISIT*. IEEE, June 2009, pp. 254–258.

[8] G. i. F. A. Asyhari, A. T., "Coding for the MIMO ARQ block-fading channel with imperfect feedback and CSIR," in *Proc. of IEEE Information Theory Workshop (ITW)*, 2010.