On Upper bounds for Decentralized MIMO Receiver

Amichai Sanderovich, Shlomo Shamai (Shitz) and Yossef Steinberg

Technion, Haifa, Israel
Email: {amichi@tx,sshlomo@ee,ysteinbe@ee}.technion.ac.il

Abstract—In this paper we investigate the achievable rate of a system that includes a nomadic transmitter with a Gaussian codebook and several antennas, which is received by multiple agents, each with a single antenna, suffering independent channel gains and additive Gaussian noise. In the nomadic regime, we assume that the agents do not have any decoding ability. These agents process their channel observations and forward it to the final destination through lossless links with fixed capacities. This paper extends a previous work [1] by providing new upper bounds on the achievable rates, and demonstrates optimality of the suggested coding scheme in several asymptotic situations. For a finite 2 × 2 setting and Rayleigh fading, the upper-bounds for both fast and block fading are demonstrated.

I. INTRODUCTION

In this paper we extend on previous results [1] on a network setting in which a nomadic transmitter has several antennas and is communicating to a remote destination, where no direct link exists between the transmitter and the final destination. The final destination receives all of its inputs from several separated agents, which are connected to it through fixed lossless links, each with a given capacity. The channel between the transmitting antennas and the agents is either the ergodic fast fading channel or the block fading channel, both with independent fading. The channel state is known to the agents and the final destination, but not to the transmitter. This setting is closely related to the setting of the Multiple Input Multiple Output (MIMO) channel, which is thoroughly treated in the literature, see [2] and others. Since the transmitter is nomadic, we assume that the agents do not possess the codebook in use, and thus do not have any decoding ability [3],[1]. In this paper we extend the upper bounds given in [1], which bounded the average mutual information for the block fading channel, to upper bounds over the outage capacity of the block fading channel, and the capacity of the fast fading channel. The bounds, and the bounding technique are related to the MIMO channel, the MIMO broadcast channel [4] and to ad-hoc networks [5]. All these works deal with situations where multiple antennas are transmitting and are received in a distributed way, either by relays, destinations or any combination of the above. The presented upper bounds constitute another application of the entropy power inequality, and were developed using ideas from [6], albeit in the vector version.

This paper is organized as follows, in section II the setting is described and the basic definitions and notations are given. Section III briefly repeats the basic achievable results that were reported in [1]. The new upper bounds are given in section IV, divided into fast fading and block fading. The two bounds are demonstrated in section V by some numeric example, to be rather close to the achievable rate when using the Wyner-Ziv (WZ) compression. Concluding remarks are then given in section VI.

II. SETTING AND MODEL DEFINITION

We consider the same setting as [1]: a system with a transmitter S which has t transmitting antennas and which transmits during n channel uses. In each channel use, the transmitter sends a vector $X \in \mathbb{C}^{t \times 1}$ to the channel, where $\frac{1}{n} \sum_{k=1}^{n} E[X(k)^*X(k)] \leq P$. By restricting $E[X(k)^*X(k)] = Q$ to be diagonal, the setting is extended also to include the sum rate of multi-user setting, where each user has a single antenna. Such restriction would not limit the achievable rate, as will become evident. The transmitter uses Gaussian signalling, which enables explicit expressions in the analytic solutions. The transmitter sends one out of $2^{nR}$ messages, when $R$ denotes the communication rate, using one codebook, chosen at random out of an ensemble [3], with codewords generated independently according to $X^n \sim \prod^n P_{X(k)}$, where $P_{X(k)} = N(0,Q)$. The rate $R$ is said to be achievable, if the error probability can be made as small as desired, by increasing $n$. In addition, we have r agents $A_1, \ldots , A_r$, each receiving the scalar channel outputs:

$$Y_i(k) = h_i(k)X(k) + N_i(k), \quad i = 1, \ldots , r, \quad k = 1, \ldots , n$$ (1)

where $h_i(k) \in \mathbb{C}^{1 \times t}$ is the vector of the channel transfer coefficients (CSI), which are either ergodic (fast fading) or block fading, and distributed independently from each other, and from any other variable, according to circularly symmetric Gaussian distribution $\mathcal{CN}(0,1)^2$. Define $H \triangleq [h_1, \ldots , h_r]$, and $H^n \triangleq \{h_1(k), \ldots , h_r(k)\}_{k=1}^n$, where we will sometimes drop the superscript $n$ to avoid cumbersome expressions. Similarly, $N_i(k) \sim \mathcal{N}(0,1)$, and are independent of each other and in time. For the sake of brevity, we drop the $k$ index from now on. The $r$ agents are connected to a remote destination with lossless links, each with capacity $C_i$ bits per channel use. The transmitter has no information regarding

\footnote{The statistical mean is denoted by $E$ and $^*$ denotes the transpose conjugate.}$^{2}$ of $\mathcal{CN}(\Xi , \Sigma)$ stands for complex Gaussian random variable with mean $\Xi$ and covariance $\Sigma$.}
For the fast fading channel, the achievable rate is [8]:

\[
R_{WZ} = E[H]\left[ \max_{\{0 \leq q|_i\} } \min_{S \subseteq \{1, \ldots, r\}} \left\{ \sum_{i \in S} |C_i - q^*_i| + \log_2 \det \left( I_{|S|} + \frac{1}{1 + P_{D_i}(q^*_i)} \right) H_S Q H_S^* \right\} \right],
\]

(2)

where \(H_S = \{h_i\}_{i \in S}\) and \(P_{D_i}(q^*_i)\) is defined by

\[
\frac{1}{1 + P_{D_i}(q^*_i)} = 1 - 2^{-q^*_i}.
\]

(3)

Since (2) is concave in \(Q\), and \(H\) is unknown to the transmitter, but is symmetrically and independently distributed, \(Q = \frac{L}{2} I_t\) is optimal.

This rate is achieved by using compression at the agents, and then forwarding the compressed signals, using a WZ scheme, to the final destination. We use compression such that the quantization noise is of power \(P_{D_i}(q^*_i)\). Equation (3) provides the connection between the power of the quantization noise and \(q^*_i\). One interpretation of \(q^*_i(k)\), from (3) is the bandwidth wasted on the compression of the additive noise by the \(i^{th}\) agent’s processing.

Remark 1: For the non-ergodic block fading channel the channel \(H^n\) is a static channel \(H^n = H\). A channel \(H\) can support reliable rate of

\[
R_{WZ}(H) = \max_{\{0 \leq q\} } \min_{S \subseteq \{1, \ldots, r\}} \left\{ \sum_{i \in S} |C_i - q^*_i| + \log_2 \det \left( I_{|S|} + \frac{1}{1 + P_{D_i}(q^*_i)} \right) H_S Q H_S^* \right\}.
\]

(4)

Because the transmitter does not have any CSI, it cannot transmit with rate \(R_{WZ}(H)\). It can however use arbitrary rate, and then with some probability, an outage event would happen. So the achievable probability of outage when the rate is \(R\) is calculated from (4) by:

\[
P (R_{WZ}(H) < R),
\]

(5)

where the probability is with respect to the randomness of \(H\). Since this rate as function of outage-probability is not concave, we need to optimize also for \(Q\), as will become evident in subsection IV-B. This is unlike the fast fading, where the optimal \(Q\) is proportional to the identity matrix, due to the concavity.

IV. UPPER BOUNDS

We first give an information theoretic upper bound for the achievable rate, based on [3]. First, define \(V_i\) to be the message sent from agent \(A_i\) after receiving \(n\) channel outputs. Notice that \(H^n\) is fully known to all agents and to the final destination, so they can use it to calculate the \(\{V_i\}\).
For any subset $S \subseteq \{1, \ldots, r\}$, the following inequalities hold:

$$
\sum_{i \in S} C_i \geq \frac{1}{n} I(Y^n_i; V_S | V_{Sc}, H^n) \geq \frac{1}{n} I(Y^n_i; V_T | H^n) - \frac{1}{n} I(Y^n_i; V_{Sc} | H^n) 
$$

$$
= \frac{1}{n} I(Y^n_i; X^n_i; V_T | H^n) - \frac{1}{n} I(Y^n_i; X^n_i; V_{Sc} | H^n) \geq \frac{1}{n} I(X^n_i; V_T | H^n) - \frac{1}{n} I(X^n_i; V_{Sc} | H^n) 
$$

$$
= \frac{1}{n} I(Y^n_i; X^n_i; V_T | H^n) - \frac{1}{n} I(Y^n_i; X^n_i; V_{Sc} | H^n) + \frac{1}{n} I(Y^n_i; V_T | H^n) - \frac{1}{n} I(X^n_i; V_{Sc} | H^n) 
$$

$$
= \frac{1}{n} I(X^n_i; V_T | H^n) - \frac{1}{n} I(X^n_i; V_{Sc} | H^n) + \sum_{i=1}^{r} q_i - \sum_{i \in S^c} q_i 
$$

$$
\geq \frac{1}{n} I(X^n_i; V_T | H^n) - \frac{1}{n} I(X^n_i; V_{Sc} | H^n) + \sum_{i \in S} q_i(11) 
$$

$$
\geq R(1 - P_e) - \frac{1}{n} I(X^n_i; V_{Sc} | H^n) + \sum_{i \in S} q_i(12) 
$$

where (8) is because $V_i$ is a function of $Y^n_i$ and $H^n$, so we have the Markov chain $V_i \rightarrow \{Y_i, H^n\} \rightarrow X$ and (12) is due to Fano’s inequality. Finally, $q_i$ is defined by $q_i \triangleq \frac{1}{n} I(Y^n_i; V_i | X^n_i, H^n)$.

Consequently, the achievable rate for reliable communication is upper bounded by:

$$
R \leq \min_{S \subseteq \{1, \ldots, r\}} \left\{ \sum_{i \in S} [C_i - q_i] + \frac{1}{n} I(X^n_i; V_{Sc} | H^n) + \frac{1}{n} \right\}. 
$$

Note that the upper bound applies for the case where the transmitter is nomadic, with Gaussian codebook, and the agents can use any signal processing. So we upper bound the rate by optimizing over the signal processing of the agents (which is not necessarily Gaussian) and the covariance matrix of the transmitter.

A. Upper bounding the rate for fast fading $H$

Lemma 1: If the transmitter is nomadic, so the agents have no decoding ability, and the transmitter uses Gaussian codebooks, the following inequality holds for any $S \subseteq \{1, \ldots, r\}$:

$$
\frac{1}{n} I(X^n_i; V_S | H^n = \theta) \leq m \log_2 \left( \prod_{k=1}^{n} |I_{|S|} + \Lambda_S(k)|^{\frac{1}{n}} - \prod_{k=1}^{n} |W_S(k)|^{\frac{1}{n}} \right) 
$$

(14)

where $\Lambda_S(k) \triangleq H_S(k)QH_S^*(k)$, $W_S(k) \triangleq \begin{cases} QH_S(k)^* \text{diag} (2^{-q_i(\theta)})_{i \in S} H_S(k) & |S| > t \\ \text{diag} (2^{-q_i(\theta)})_{i \in S} H_S(k)QH_S^*(k)^* & |S| \leq t, \end{cases}$ $q_i(\theta) \triangleq \frac{1}{n} I(Y^n_i; V_i | X^n_i, H^n = \theta)$ and $m \triangleq \min\{t, |S|\}$.

The proof appears in [8] and is based on the entropy power inequality, as done for the Gaussian CEO problem in [6]. Since $H_QH^*$ is distributed the same as $HU^*QUH^*$, when $U$ is unitary matrix, $Q$ can be restricted to be diagonal in (14).

This upper bound is tight when the channel is $H = (1, \ldots, 1)^T$, which is similar to the Gaussian CEO problem. Observe that unlike the achievable rate, which is a concave function of $Q$, so that $Q \propto I$ is optimal, the upper bound is not concave in $Q$, thus we need to maximize the upper bound also over $Q$ such that $Q \in \mathcal{P}$, where

$$
P = \{Q : Q_{ii} = 0 \text{ for } i \neq j, Q_{ii} \geq 0, \ \text{trace}(Q) \leq P \}. 
$$

Corollary 1: Note that in the limit of $n \to \infty$, due to the ergodic fading process:

$$
\lim_{n \to \infty} \frac{1}{n} I(X^n_i; V_S | H^n = \theta) \leq F(S, q_S),
$$

where

$$
F(S, q_S) \triangleq m \log_2 \left( 2^{\frac{1}{n} E_{H_i} \log_2 |I + \Lambda_S| - \frac{1}{n} E_{H_i} \log_2 |S| \right) 
$$

(18)

and we use the notation $q_i \triangleq q_i(\theta)$ and $q_S \triangleq (q_i)_{i \in S}$, and $\Lambda_S = \Lambda_S(1), W_S = W_S(1)$. Consequently, (17) can be averaged over the channels:

$$
\lim_{n \to \infty} \frac{1}{n} I(X^n_i; V_S | H^n) \leq F(S, q_S).
$$

The dependence of $F$ from (18) on $q_i$, stems from the definition of $q_i$, as the bandwidth used for the noise compression, and is essential for the bound, as it is used for connecting the bandwidth for the signal compression to the achievable rate. Using the upper bound from equation (13) with the corollary above, the following upper bounds any achievable rate:

Proposition 1: The achievable rate of the nomadic transmitter, with input covariance $Q$, which is received by agents through fast fading channel is upper bounded by:

$$
R \leq \max_{Q \in \mathcal{P}, 0 \leq q_i \leq C_i} \min_{S \subseteq \{1, \ldots, r\}} \left\{ F(S, q_S) + \sum_{i \in S} [C_i - q_i] \right\}. 
$$

(20)

Remark 2: When $C_i = C$ for $i = 1, \ldots, r$, since $H$ is identically distributed, then the argument which is maximized over $\{q_i\}_{i=1}^{r}$ in (20), is symmetric in $\{q_i\}_{i=1}^{r}$. Since this argument is also concave in $\{q_i\}_{i=1}^{r}$, for $C_i = C$, equation (20) is maximized by $q_i = q^*$. So that for the symmetric case:

$$
R \leq \max_{Q \in \mathcal{P}, 0 \leq q^* \leq C} \min_{S \subseteq \{1, \ldots, r\}} \left\{ F(S, q^*) + |S|(C - q^*) \right\}. 
$$

(21)

Following remark 2, we outline a special case where this bound is tight.

Corollary 2: The WZ approach is optimal for infinite transmission power, $Q = \frac{P}{t} I$, and $C_i = C$, $i = 1, \ldots, r$. Here we take $P \to \infty$ and fix $t$ and $r$.

We justify it for $r \leq t$, where basically the same holds also for $t < r$. Taking $P \to \infty$ and optimizing over $q_{WZ}$ (where
\[ q'_i = q_{WZ}, \ i = 1, \ldots, r \] in equation (2) instead of over \( \{q'_i\} \), results with:

\[
\frac{1}{n} I(X^n; V_S|H^n) = m \log_2 (P) + E_H \log_2 \left( \frac{1}{l} \frac{H_S H_S^T}{l} \right) + m E_H \log_2 \left( 1 - 2^{-q_{WZ}} \right) + o(P), \tag{22}
\]

where \( o(P) \to 0 \) when \( P \to \infty \). On the other hand, considering remark 2, we know that \( q_i = q_{UB} \) is optimal. So we get

\[
F(S, q_{UB}) = m \log_2 (P) + E_H \log_2 \left( \frac{1}{l} \frac{H_S H_S^T}{l} \right) + m \log_2 \left( 2^{o(P)} - 2^{-q_{UB}} \right). \tag{23}
\]

which is identical, in the limit, to (22), justifying the corollary.

Since taking \( q_{WZ} = q_{UB} \) is optimal for \( P \to \infty \), optimization over \( H \), in (2) is superfluous. This means that for large \( P \), the WZ parameters are independent of \( H \), which in turns means that the \( i \)-th agent needs to know only its own \( h_i \). The tightness, however, is maintained as long as \( C_i = C \) for \( i = [1, \ldots, r] \), where \( C \) can take any value. Observe that the CSI \((h_i)\) is still required at \( i \)-th agent, for the determination of the codebook of \( U \) (see [3]).

The upper bound of proposition 1 is not tight because the upper bound in lemma 1 was obtained using the vector version of the entropy power inequality. This inequality is known to be tight only for proportional covariance matrices, which is not our case. Thus the entropy power inequality introduces a gap that prevents the bound from being tight. This gap can be mitigated by taking into account smaller matrices. The following proposition improves upon proposition 1 by optimizing over sub-matrices of \( S \).

**Proposition 2:** The achievable rate of the nomadic transmitter, with total power \( P \), which is received by agents through fast fading channel, is upper bounded by:

\[
R_u \triangleq \max_{Q \in P, 0 \leq q \leq C_i} \left\{ \min_{\bigcup_{j=1}^r Z_j \subseteq \{1, \ldots, r\}, \ i \neq j : Z_j \cap Z_i = \emptyset} \left\{ \sum_{j=1}^r F(Z_j, q_{Z_j}) + \sum_{i \in \cup_{j=1}^r Z_j} [C_i - q_i] \right\} \right\} \tag{24}
\]

where \( F(Z_j, q_{Z_j}) \) is defined as before, in equation (17).

The proof is very simple, considering for every group of disjoint subsets \((\{Z_j\}_{j=1}^r) : Z_j \cap Z_i = \emptyset \) when \( i \neq j \) that cover \( \bigcup_{j=1}^r Z_j = S \) we can write:

\[
I(X^n; V_S|H^n) \leq \sum_{j=1}^r I(X^n; V_{Z_j}|H^n), \tag{25}
\]

which is due to the Markov chain \( V_j - X - V_i \) when \( i \neq j \), and then using the upper bound of proposition 1 again, for every element.

For the symmetric case, where \( C_i = C \) for \( i = [1, \ldots, r] \),

\[ R_u \triangleq rC + r \max_{Q \in P} \left\{ \min_{1 \leq j \leq r} \left\{ \sum_{j=1}^r F_j Z_j, q_{Z_j} \right\} \right\} \tag{26}
\]

By solving the above optimization problem for \( \{k_j\}_{j=1}^r \) and then solving for \( q^* \) by explicitly writing \( F(j, q^*) \) we can simplify (26) to

**Proposition 3:** The achievable rate of nomadic transmitter in the symmetric case, \( C_i = C, \ i = 1, \ldots, r \), is upper bounded by

\[
R_u = rC + r \max_{Q \in P} \left\{ \min_{1 \leq j \leq r} \left\{ \frac{1}{j} E_H \log_2 |I_j + H_j Q H_j^*| \right\} \right\} - \log_2 \left( 2^C + 2 \frac{1}{j} E_H \log_2 |H_j Q H_j^*| \right). \tag{27}
\]

where \( H_j \) is the fading coefficients seen by any subset of \( j \) agents (since the channel is ergodic, it does not matter which subset).

The improvement over the bound from proposition 1, is seen in the next corollary, where the inequality (25) is in fact an equality, and a conclusive result is obtained.

**Corollary 3:** The WZ approach is optimal for \( Q = \frac{1}{t} I \) and \( t \to \infty \) while \( r \) is fixed.

Since \( H Q H^* \) is proportional to the identity matrix (when \( t \to \infty \)), the received signals are independent, which is equivalent to \( r \) parallel links. Namely, when \( t \to \infty \) while \( r \) is fixed we
get

\[
\lim_{t \to \infty} \frac{1}{t} \|HH^* - I_r\| = 0.
\]  

So that the capacity is [9]:

\[
\lim_{t \to \infty} R = \max_{\{q_i\}} \left\{ \min_{S} \left\{ \sum_{i \in S^c} |C_i - q_i| + \sum_{i \in S} \log_2(1 + P(1 - 2^{-q_i})) \right\} \right\}.
\]  

(29)

B. Block fading upper bound

In this subsection, we will consider the case of \( H \) distributed independently, but once per block, such that \( H^n = H \). The achievable rate vs outage is given in remark 1. In the sequel of this subsection, we will upper bound this rate, given some permissible outage probability.

For the upper bound, we again use:

\[
R(H = \Theta) \leq \max_{\{q_i\}} \left\{ \min_{S} \left\{ \frac{1}{n} I(V_S^n; X^n|H = \Theta) + \sum_{i \in S^c} |C_i - q_i| \right\} \right\}.
\]  

(30)

For \( I(V_S^n; X^n|H = \Theta) \), we use the upper bound of lemma 1. Since \( H^n = H \), we get:

\[
G(S, q) \triangleq m \log_2 \left( \left| |S| + \Lambda_S | \right|^{\frac{1}{n}} - |W_S|^{\frac{1}{n}} \right)
\]  

(31)

\[
\frac{1}{n} I(X^n; V_S^n|H = \Theta) \leq G(S, q_S)
\]  

(32)

where \( \Lambda_S = H_S Q H_S^* \), as before and \( W_S \) is defined by \( W_S(1) \) from equation (15). Combining (30) and (32) and noticing that \( H \) is a random variable, we get the following proposition:

**Proposition 4**: An upper bound on the achievable rate \( R \), for given outage probability \( \epsilon \) is the maximal \( R \) which fulfills:

\[
P \left( \max_{\{0 \leq W \leq C_1\}} \left\{ \min_{S \subseteq \{1, \ldots, r\}} \left\{ G(S, q) + \sum_{i \in S^c} |C_i - q_i| \right\} \right\} < R \right) \leq \epsilon.
\]  

(33)

Actually, we can improve upon (33), the same way it was done in proposition 2:

\[
P \left( \max_{\{0 \leq q_i \leq C_1\}} \left\{ \min_{\bigcup_{j=1}^r Z_j \subseteq \{1, \ldots, r\}} \left\{ \sum_{j=1}^r G(Z_j, q_{Z_j}) + \sum_{i \in \bigcup_{j=1}^r Z_j} |C_i - q_{Z_j}| \right\} \right\} < R \right) \leq \epsilon.
\]  

(34)

but since the problem is not symmetric (due to the non-ergodic \( H \)), we can not further simplify it, as in proposition 3. However, the limiting behavior of (28) is true also for the block fading case. Thus the optimality of the WZ approach when \( t \to \infty \) is assured in the block fading case as well.

V. Numerical example

The achievable rates and the upper bounds for both fast fading and block fading channels, were calculated for a \( 2 \times 2 \) system, with \( C_1 = C_2 = 2 \), for several signal to noise ratios (SNR, \( P \) in dB), and the results are presented in figure 2. For the fast fading, the upper bound is obtained by averaging over 5000 channel realizations (the expectation over \( H \) in (27)). It is seen there that the upper bound is convex, and that it is close to the achievable rate, when using WZ compression. For the low and high SNR, the bound is tighter, as expected.

For block fading channel, the upper bound from (34) is depicted along with the achievable rate (5), for outage probability of \( \epsilon = 10^{-2} \). The probability was calculated using Monte Carlo simulations over \( H \). It is seen there that the bound is again very tight for the low SNR region, and the gap becomes higher, with larger SNR, although it remains rather small, no more than 1 dB throughout the figure.

VI. Conclusion

In this paper we upper bound the achievable rates in a nomadic distributed MIMO system. We give upper bound for both fast and block fading channels. The bounds are based on the vector version of the entropy power inequality, and are shown to be close to the optimal achievable rates, even for rather small number of antennas (\( 2 \times 2 \)). The bounds are also used to demonstrate optimality in two extreme cases, when \( P \to \infty \) and when \( t \to \infty \).

Acknowledgment

This research was supported by the EU 6th framework program via the NEWCOM network of excellence and the REMON consortium for 4G wireless communications.

References