A Linear Interference Network with Local Side-Information

Amos Lapidoth
ETH Zurich
CH-8092 Zurich, Switzerland
Email: lapidoth@isi.ee.ethz.ch

Shlomo Shamai (Shitz)
Technion
Haifa 32000, Israel
Email: sshlomo@ee.technion.ac.il

Michèle A. Wigger
ETH Zurich
CH-8092 Zurich, Switzerland
Email: wigger@isi.ee.ethz.ch

***Abstract*— For an interference network where Receiver k receives the sum of the signal transmitted by Transmitter k and a scaled version of the signal transmitted by Transmitter $k-1$ corrupted by Gaussian noise we compute the pre-log of the sum-rate capacity for the case where each transmitter has side-information consisting of the messages to be sent by its J predecessors.**

I. Introduction and Main Result

We envision a wireless communication scenario where multiple transmitters wish to communicate with multiple receivers. We assume a point-to-point setting, that is, each transmitter sends a message to only one particular receiver. However, due to the wireless communication channel each receiver observes a noisy version of the sum of the signal transmitted by the corresponding transmitter and the attenuated signals transmitted by closely located transmitters. The communication network should thus be modeled as an interference network.

Also, we envision that transmitters can be located close to each other and thus transmitters might be cognizant of the messages of nearby located transmitters. Such a scenario for a example arises in cognitive radio networks [1].

In this work specifically we consider an interference network with $\mathbb{K}$ transmitters and $\mathbb{K}$ receivers and envision that the transmitters and receivers are located on a grid, which we model as the set $\mathcal{K} = \{1, \ldots, \mathbb{K}\}$. Based on its observation $\mathbf{Y}_k = (Y_{k,1}, \ldots, Y_{k,n})$ at times 1 through n, Receiver k wishes to decode Message k, which we denote by M_k. We assume that $M_1, \ldots, M_{\mathbb{K}}$ are independent and that for each $k \in \mathcal{K}$ the random variable M_k is uniformly distributed over the set

$$\mathcal{M}_k = \left\{1, \ldots, \left\lfloor e^{nR_k} \right\rfloor \right\}, \tag{1}$$

where R_k denotes the rate of transmission from Transmitter k.

The time-t signal $Y_{k,t}$ received by Receiver k is given by

$$Y_{k,t} = x_{k,t} + \alpha x_{k-1,t} + Z_{k,t}, \quad k \in \mathcal{K},\ 1 \leq t \leq n, \tag{2}$$

where $x_{k,t}$ denotes the time-t symbol produced by Transmitter k, α is some non-zero deterministic constant, and $\{Z_{k,t}\}$ are $\mathbb{K} \cdot t$ IID mean-zero variance-N random variables:

$$\{Z_{k,t}\}_{\substack{k \in \mathcal{K} \\ 1 \leq t \leq n}} \sim \text{ IID } \mathcal{N}(0, N). \tag{3}$$

The research of S. Shamai was supported by the Israel Science Foundation, and by the EU 6th framework programme, via NEWCOM Network of Excellence. The research of Michèle A. Wigger was supported by the Swiss National Science Foundation under Grant 200021-111863 /1.

We assume that in addition to its own message M_k, Transmitter k is also cognizant of the J messages $M_{k-1}, \ldots, M_{k-J}$, where, to simplify notation, we define $M_0, \ldots, M_{-(J-1)}$ to be some deterministic variables. Here J is some non-negative integer that models the amount of local side-information. The problem becomes interesting for $J \geq 1$. The time-t symbol $X_{k,t}$ produced by Transmitter k can thus depend on the messages $M_k, M_{k-1}, \ldots, M_{k-J}$:

$$X_{k,t} = X_{k,t}\big(M_k, M_{k-1}, \ldots, M_{k-J}\big). \tag{4}$$

We impose a symmetric average power constraint

$$\frac{1}{n} \sum_{t=1}^{n} X_{k,t}^2 \left(M_k, \ldots, M_{k-J}\right) \leq P, \quad k \in \mathcal{K}, \tag{5}$$

where P is some positive constant modeling the available average power at each transmitting node.

We denote by R_Σ the sum of the rates

$$R_\Sigma = \sum_{k \in \mathcal{K}} R_k \tag{6}$$

and by $C_\Sigma^{J,\mathbb{K}}(P, N)$ the supremum of all achievable sum-rates, where $(R_1, \ldots, R_{\mathbb{K}})$ is achievable if, as the block-length n tends to infinity, each decoder can decode the message intended for it with an average probability of error that decays to zero.

The expression

$$\eta(K, J) \triangleq \lim_{P \to \infty} \frac{C_\Sigma^{J,\mathbb{K}}(P, N)}{\frac{\mathbb{K}}{2} \log\left(1 + \frac{P}{N}\right)} \tag{7}$$

is called the sum-rate pre-log per user and yields a comparison in the limit when $P \to \infty$ of the sum-rate capacity per user in our setting to the sum-rate capacity per user that would be achievable in the absence of any interference.

Note that the described setting is a asymmetric version of Wyner's linear cellular network model [2]. It should also be emphasized that in our set-up Transmitter k knows the *messages* $M_k, M_{k-1}, \ldots, M_{k-J}$ but not necessarily the *signals* $X_{k-1,t}, \ldots, X_{k-J,t}$. Otherwise the interference at all receivers could be canceled entirely using dirty-paper coding [3].

Our Main result is that for this set-up,

$$\boxed{\eta(K, J) = 1 - \frac{\left\lfloor \frac{\mathbb{K}}{J+2} \right\rfloor}{\mathbb{K}}.} \tag{8}$$

This result illustrates that for $K \geq 4$ and when every transmitter in addition to its own message is also cognizant of the preceding $J \geq 1$ messages, the pre-log per user $\eta(K, J)$ is strictly larger than when every transmitter knows only its own message (i.e., when $J = 0$). It also shows that for $K \geq 2$ and when every transmitter in addition to its own message is also cognizant of all preceding messages (i.e., when $J = K - 1$), then $\eta(K, J) = 1$, so that in the limit $P \to \infty$ the sum-rate capacity per user approaches the sum-rate capacity per user of a channel without interference.

It is interesting to compare these results with the results for the 2-transmitters/2-receivers interference channel [4]. In the latter each transmitter needs to be cognizant of all messages for the sum-rate pre-log per user to equal 1.

In the next section we sketch an achievability proof and in the section following that a converse.

II. Achievability

The achievability of (8) is based on a coding scheme that uses Costa's "writing on dirty paper" [3] and silences some of the transmitters[1]. We sketch the proposed scheme below.

Receiver 1 suffers only from noise when decoding M_1, so in our proposed scheme Transmitter 1 uses a Gaussian codebook that achieves the capacity of the single-user Gaussian channel and Receiver 1 uses a standard (e.g., weak typicality or maximum-likelihood) decoder for the Gaussian channel. This yields the achievability of $R_1 = 1/2 \cdot \log(1 + P/N)$.

Receiver 2 experiences an interference from Transmitter 1 but, provided that $J \geq 1$, this interference is known to Transmitter 2 (who, if $J \geq 1$, knows M_1 and hence the codeword and the signal $\{X_{1,t}\}$ generated by Transmitter 1). Consequently, if $J \geq 1$, Transmitter 2 can use dirty-paper coding to allow Receiver 2 to decode M_2 as if there were no interference from Transmitter 1. This allows Transmitter 2 also to achieve $1/2 \cdot \log(1 + P/N)$.

If $J \geq 2$, then Transmitter 3 knows M_1 and hence the signal that was sent by Transmitter 1. Consequently, since it also knows M_2, it also knows the signal that was produced by Transmitter 2 in its dirty-paper coding scheme. Knowing the signal transmitted by Transmitter 2 allows it to use dirty-paper coding to allow Receiver 3 to decode Message M_3 as though there were no interference from Transmitter 2, thus also achieving $1/2 \cdot \log(1 + P/N)$. If J is not greater than 2, we silence Transmitter 3.

In general using dirty-paper coding the transmitters $1, \ldots, J + 1$ can all communicate at the rate $1/2 \cdot \log(1 + P/N)$

$$R_1 = R_2 = \cdots = R_{J+1} = 1/2 \cdot \log(1 + P/N). \quad (9)$$

Our coding scheme then silences Transmitter $J + 2$ so that

$$R_{J+2} = 0. \quad (10)$$

The pattern now repeats because Receiver $J + 3$ now suffers no interference (because Transmitter $J + 2$ is silent), so Transmitter $J+3$ can use a simple Gaussian codebook. Transmitters $(J+4)$ through $2J+3$ can now communicate at full rate using dirty-paper coding, and we silence Transmitter $2(J+2)$.

The last transmitter we silence is Transmitter $\gamma(J + 2)$, where

$$\gamma = \left\lfloor \frac{\mathrm{K}}{J + 2} \right\rfloor. \quad (11)$$

The remaining transmitters $\gamma(J + 2) + 1$ through K can now transmit at full rate $1/2 \cdot \log(1 + P/N)$ because Receiver $\gamma(J + 2) + 1$ suffers no interference, and hence Transmitter $\gamma(J + 2) + 1$ can use a simple Gaussian codebook, and because transmitters $\gamma(J + 2) + 2$ through K can employ dirty-paper coding.

Analyzing the sum rate we note that, since γ transmitters are silenced and the rest communicate at full rate,

$$R_\Sigma = (\mathrm{K} - \gamma) 1/2 \cdot \log(1 + P/N), \quad (12)$$

which yields the lower bound needed to establish (8).

III. Converse

The proof of the converse in the case where $\mathrm{K} \leq J + 1$ is very simple. Indeed,

$$C_\Sigma^{J,\mathrm{K}}(P, N) = \frac{\mathrm{K}}{2} \log(1 + P/N), \quad \mathrm{K} \leq J + 1 \quad (13)$$

where the achievability follows from our proposed coding scheme of the previous section and where the converse follows by revealing to each receiver ν the signals transmitted by transmitters $1, \ldots, \nu - 1$. We shall therefore focus on the more interesting case of

$$\mathrm{K} \geq J + 2. \quad (14)$$

The first step in proving the converse for this case is to establish that the sum-rate capacity of our network is upper bounded by the sum-rate capacity $\tilde{C}_\Sigma^{J,\mathrm{K}}(P, N)$ of a modified network where receivers $1(J+2)+1, 2(J+2)+1, \ldots, \gamma(J+2)+1$ do not suffer from any interference:

$$C_\Sigma^{J,\mathrm{K}}(P, N) \leq \tilde{C}_\Sigma^{J,\mathrm{K}}(P, N). \quad (15)$$

This is done by showing that the sum-rate capacity $\tilde{C}_\Sigma^{J,\mathrm{K}}(P, N)$ of the latter modified network would be identical to the sum-rate capacity of our network had these receivers been given, as side information, all preceding messages, i.e., if, for all $1 \leq \nu \leq \gamma$, Receiver $\nu(J + 2) + 1$ had been made cognizant of messages $M_1, \ldots, M_{\nu(J+2)}$.

The next step is to show that the elimination of the interference suffered by the receivers $1(J + 2) + 1, 2(J + 2) + 1, \ldots, \gamma(J + 2) + 1$ results in the modified network having a sum-rate capacity which is equal to the sum of the sum-rate capacities of $\gamma+1$ networks of the original kind where γ of these sub-networks have $J + 2$ transmitters/receivers and where one of the sub-networks has $\mathrm{K} - \gamma(J + 2)$ transmitters/receivers:

$$\tilde{C}_\Sigma^{J,\mathrm{K}}(P, N) = \gamma C_\Sigma^{J,J+2}(P, N) + C_\Sigma^{J,\mathrm{K}-\gamma(J+2)}(P, N). \quad (16)$$

The first sub-network consists of transmitters 1 through $J+2$, the second of transmitters $J+3$ through $2(J+2)$, the γ-th of

[1] But see [5] for a simpler scheme. This paper also addresses more general networks.

transmitters $(\gamma-1)(J+2)+1$ through $\gamma(J+2)$, and the final sub-network consists of the transmitters $\gamma(J+2)+1$ through $\mathbb{K}$.

The number of transmitters in this last sub-network is smaller than $J+2$, so its sum-rate capacity is given, using (13), by

$$C_\Sigma^{J,\mathbb{K}-\gamma(J+2)}(P,N)=\big(\mathbb{K}-\gamma(J+2)\big)\frac{1}{2}\log\left(1+P/N\right), \quad (17)$$

so that by (15), (16), and (17)

$$C_\Sigma^{J,\mathbb{K}}(P,N)\le \gamma C_\Sigma^{J,J+2}(P,N) + \big(\mathbb{K}-\gamma(J+2)\big)\frac{1}{2}\log(1+P/N). \quad (18)$$

In the third step we prove that

$$\lim_{P\to\infty}\frac{C_\Sigma^{J,J+2}(P,N)}{\frac{1}{2}\log\left(1+\frac{P}{N}\right)}\le J+1, \quad (19)$$

which combined with (18) establishes

$$\lim_{P\to\infty}\frac{C_\Sigma^{J,\mathbb{K}}(P,N)}{\frac{\mathbb{K}}{2}\log\left(1+\frac{P}{N}\right)}\le \frac{\gamma(J+1)}{\mathbb{K}}+\frac{\mathbb{K}-\gamma(J+2)}{\mathbb{K}} = 1-\frac{\gamma}{\mathbb{K}} \quad (20)$$

and thus completes the proof of the converse. The proof of (19) is based on the inequality

$$\begin{aligned}
&C_\Sigma^{J,J+2}(P,N)\\
&\le (J+1)\frac{1}{2}\log\left(1+\frac{P}{N}\right)+(J+1)\log\left(\frac{1+|\alpha|}{|\alpha|}\right)\\
&\quad +J\log\left(\sqrt{2}\right)+(J-1)\max\{0,\log(|\alpha|)\}\\
&\quad +\max\left\{0,\log\left(\sqrt{2}|\alpha|\right)\right\}+\max\left\{0,\log\left(\frac{|\alpha|}{\sqrt{2}}\right)\right\} \quad (21)
\end{aligned}$$

which we next derive.

By Fano's inequality we have that if $(R_1,\ldots,R_{J+2})$ are achievable for each $1\le k\le J+2$

$$\begin{aligned}
R_k &\le \frac{1}{n}I(M_k;\mathbf{Y}_k)+\frac{\epsilon(n)}{n}\\
&=\frac{1}{n}\left[h(\mathbf{Y}_k)-h(\mathbf{Y}_k|M_k)\right]+\frac{\epsilon(n)}{n} \quad (22)
\end{aligned}$$

where

$$\lim_{n\to\infty}\frac{\epsilon(n)}{n}=0. \quad (23)$$

Consequently,

$$\sum_{k=1}^{J+2}R_k\le\frac{1}{n}\sum_{k=1}^{J+2}\left[h(\mathbf{Y}_k)-h(\mathbf{Y}_k|M_k)\right]+(J+2)\frac{\epsilon(n)}{n}. \quad (24)$$

We next upper bound the terms in (24). Denoting by $\mathbf{X}_k$ the vector $(X_{k,1},\ldots,X_{k,n})$ and similarly for $\mathbf{Z}_k$ we have

$$\begin{aligned}
\sum_{k=2}^{J+2}h(\mathbf{Y}_k)&=\sum_{k=2}^{J+2}h\left(\mathbf{X}_k+\alpha\mathbf{X}_{k-1}+\mathbf{Z}_k\right)\\
&\le (J+1)\frac{n}{2}\log\left(2\pi e\left(P\left(1+|\alpha|\right)^2+N\right)\right)\\
&\le (J+1)\frac{n}{2}\log\left(2\pi e\left(P+N\right)\right)\\
&\quad +(J+1)n\log\left(1+|\alpha|\right), \quad (25)
\end{aligned}$$

because the Gaussian distribution maximizes differential entropy subject to a second-moment constraint. Note that while a similar bound would hold also for $h(\mathbf{Y}_1)$, we prefer not to include it in the sum.

We next turn to the conditional differential entropies. The first one can be computed directly because the first receiver experiences no interference :

$$h(\mathbf{Y}_1|M_1)=h(\mathbf{X}_1+\mathbf{Z}_1|M_1)=h(\mathbf{Z}_1)=\frac{n}{2}\log(2\pi eN). \quad (26)$$

The rest of the conditional differential entropies are bounded by

$$\begin{aligned}
&h(\mathbf{Y}_{J+2}|M_{J+2})\\
&\quad\ge h(\mathbf{Y}_{J+2}|M_{J+2},\ldots,M_2)\\
&\quad= h\left(\mathbf{X}_{J+2}+\alpha\mathbf{X}_{J+1}+\mathbf{Z}_{J+2}|M_{J+2},\ldots,M_2\right)\\
&\quad= h\left(\alpha\mathbf{X}_{J+1}+\mathbf{Z}_{J+2}|M_{J+2},\ldots,M_2\right)\\
&\quad= h\left(\mathbf{X}_{J+1}+\frac{1}{\alpha}\mathbf{Z}_{J+2}\Big|M_{J+2},\ldots,M_2\right)+n\log\left(|\alpha|\right) \quad (27)
\end{aligned}$$

and for $k=2,\ldots,J+1$ by

$$\begin{aligned}
&h(\mathbf{Y}_k|M_k)\\
&\quad\ge h\left(\mathbf{X}_k+\alpha\mathbf{X}_{k-1}+\mathbf{Z}_k\Big|M_k,\ldots,M_2\right)\\
&\quad= h\left(\mathbf{X}_k+\frac{1}{\sqrt{2}}\mathbf{U}_k+\alpha\mathbf{X}_{k-1}+\frac{1}{\sqrt{2}}\mathbf{V}_k\Big|M_k,\ldots,M_2\right)\\
&\quad\ge h\left(\alpha\mathbf{X}_{k-1}+\frac{1}{\sqrt{2}}\mathbf{V}_k\Big|\mathbf{X}_k+\frac{1}{\sqrt{2}}\mathbf{U}_k,M_k,\ldots,M_2\right)\\
&\quad= h\left(\alpha\mathbf{X}_{k-1}+\frac{1}{\sqrt{2}}\mathbf{V}_k\Big|M_k,\ldots,M_2\right)\\
&\qquad -I\left(\alpha\mathbf{X}_{k-1}+\frac{1}{\sqrt{2}}\mathbf{V}_k;\mathbf{X}_k+\frac{1}{\sqrt{2}}\mathbf{U}_k\Big|M_k,\ldots,M_2\right)\\
&\quad= h\left(\alpha\mathbf{X}_{k-1}+\frac{1}{\sqrt{2}}\mathbf{V}_k\Big|M_k,\ldots,M_2\right)\\
&\qquad -h\left(\mathbf{X}_k+\frac{1}{\sqrt{2}}\mathbf{U}_k\Big|M_k,\ldots,M_2\right)\\
&\qquad +h\left(\mathbf{X}_k+\frac{1}{\sqrt{2}}\mathbf{U}_k\Big|\alpha\mathbf{X}_{k-1}+\frac{1}{\sqrt{2}}\mathbf{V}_k,M_k,\ldots,M_2\right)\\
&\quad\ge h\left(\alpha\mathbf{X}_{k-1}+\frac{1}{\sqrt{2}}\mathbf{V}_k\Big|M_k,\ldots,M_2\right)\\
&\qquad -h\left(\mathbf{X}_k+\frac{1}{\sqrt{2}}\mathbf{U}_k\Big|M_k,\ldots,M_2\right)\\
&\qquad +\frac{n}{2}\log\left(2\pi eN\right)-n\log\left(\sqrt{2}\right)\\
&\quad= h\left(\mathbf{X}_{k-1}+\frac{1}{\sqrt{2}\alpha}\mathbf{V}_k\Big|M_k,\ldots,M_2\right)\\
&\qquad -h\left(\mathbf{X}_k+\frac{1}{\sqrt{2}}\mathbf{U}_k\Big|M_k,\ldots,M_2\right)\\
&\qquad +\frac{n}{2}\log(2\pi eN)+n\log\left(\frac{|\alpha|}{\sqrt{2}}\right) \quad (28)
\end{aligned}$$

where $\mathbf{U}_2,\ldots,\mathbf{U}_{J+1},\mathbf{V}_2,\ldots,\mathbf{V}_{J+1}$ are IID zero-mean Gaussian vectors of a covariance matrix which is the $n\times n$ identity matrix scaled by N. The first inequality holds because conditioning reduces entropy; the first equality follows by expressing the Gaussian noise-vector $\mathbf{Z}_k$ as the sum $\frac{1}{\sqrt{2}}\mathbf{U}_k + \frac{1}{\sqrt{2}}\mathbf{V}_k$; the second inequality by conditioning the differential entropy on the term $\mathbf{X}_k + \frac{1}{\sqrt{2}}\mathbf{U}_k$; the second equality by adding and subtracting the differential entropy $h\left(\alpha\mathbf{X}_{k-1} + \frac{1}{\sqrt{2}}\mathbf{V}_k|M_k,\ldots,M_2\right)$; and the third inequality by conditioning the last summand on the left hand side of the inequality on $\mathbf{X}_k$.

Inequalities (27) and (28) imply

$$
\begin{aligned}
&h(\mathbf{X}_1+\mathbf{Z}_1) - \sum_{k=2}^{J+2} h(\mathbf{Y}_k|M_k)\\
&\quad\le h(\mathbf{X}_1+\mathbf{Z}_1)\\
&\qquad - \sum_{k=2}^{J+1}\Bigg[h\left(\mathbf{X}_{k-1} + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_k\middle|M_k,\ldots,M_2\right)\\
&\qquad -h\left(\mathbf{X}_k + \frac{1}{\sqrt{2}}\mathbf{U}_k\middle|M_k,\ldots,M_2\right)\\
&\qquad +\frac{n}{2}\log(2\pi e N) + n\log\left(\frac{|\alpha|}{\sqrt{2}}\right)\Bigg]\\
&\qquad -h\left(\mathbf{X}_{J+1} + \frac{1}{\alpha}\mathbf{Z}_{J+2}\middle|M_{J+2},\ldots,M_2\right) - n\log(|\alpha|)\\
&\quad= h(\mathbf{X}_1+\mathbf{Z}_1) - h\left(\mathbf{X}_1 + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_2\middle|M_2\right)\\
&\qquad + \sum_{\ell=2}^{J}\Bigg[h\left(\mathbf{X}_\ell + \frac{1}{\sqrt{2}}\mathbf{U}_\ell\middle|M_\ell,\ldots,M_2\right)\\
&\qquad -h\left(\mathbf{X}_\ell + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_{\ell+1}\middle|M_{\ell+1},\ldots,M_2\right)\Bigg]\\
&\qquad +h\left(\mathbf{X}_{J+1} + \frac{1}{\sqrt{2}}\mathbf{U}_{J+1}\middle|M_{J+1},\ldots,M_2\right)\\
&\qquad -h\left(\mathbf{X}_{J+1} + \frac{1}{\alpha}\mathbf{Z}_{J+2}\middle|M_{J+2},\ldots,M_2\right)\\
&\qquad -J\frac{n}{2}\log(2\pi e N) - Jn\log\left(\frac{|\alpha|}{\sqrt{2}}\right) - n\log(|\alpha|). \quad (29)
\end{aligned}
$$

We continue now to separately bound the difference in the first line, the summands, and the difference between the forth and the fifth line of the right hand side of (29). We start by establishing the following upper bound on the first difference

$$
\begin{aligned}
&h(\mathbf{X}_1+\mathbf{Z}_1) - h\left(\mathbf{X}_1 + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_2\middle|M_2\right)\\
&\qquad\qquad \le \max\left\{0, n\log\left(\sqrt{2}|\alpha|\right)\right\}. \quad (30)
\end{aligned}
$$

To prove this upper bound we first note that we can drop the conditioning on the message M_2 because it is independent of $\mathbf{X}_1$ and $\mathbf{V}_2$. For the rest of the proof we distinguish between the case when $\alpha \le \frac{1}{\sqrt{2}}$ and the case when $\alpha > \frac{1}{\sqrt{2}}$. We start with the proof for the case when $\alpha \le \frac{1}{\sqrt{2}}$ and introduce the independent Gaussian random vector $\tilde{\mathbf{Z}}_1$ which is of zero-mean and covariance matrix $N \cdot I_n$, where I_n denotes the n-dimensional identity matrix. Since both $\mathbf{V}_2$ and the pair $(\mathbf{Z}_1, \tilde{\mathbf{Z}}_1)$ are independent Gaussians and also independent of $\mathbf{X}_1$ we have the following identity

$$
h\left(\mathbf{X}_1 + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_2\right) = h\left(\mathbf{X}_1 + \mathbf{Z}_1 + \sqrt{\left(\frac{1}{2\alpha^2}-1\right)}\tilde{\mathbf{Z}}_1\right). \quad (31)
$$

Then, with (31) and the non-negativity of mutual information we obtain that

$$
\begin{aligned}
&h(\mathbf{X}_1+\mathbf{Z}_1) - h\left(\mathbf{X}_1 + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_2\right)\\
&\quad= h(\mathbf{X}_1+\mathbf{Z}_1) - h\left(\mathbf{X}_1 + \mathbf{Z}_1 + \sqrt{\left(\frac{1}{2\alpha^2}-1\right)}\tilde{\mathbf{Z}}_1\right)\\
&\quad= -I\left(\mathbf{X}_1 + \mathbf{Z}_1 + \sqrt{\left(\frac{1}{2\alpha^2}-1\right)}\tilde{\mathbf{Z}}_1; \tilde{\mathbf{Z}}_1\right) \le 0.
\end{aligned}
$$

To treat the second case, i.e., when $\alpha > \frac{1}{\sqrt{2}}$, we use that

$$
h(\mathbf{X}_1+\mathbf{Z}_1) = h\left(\mathbf{X}_1 + \sqrt{\left(1-\frac{1}{2\alpha^2}\right)}\mathbf{U}_2 + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_2\right)
$$

and then obtain the following upper bound

$$
\begin{aligned}
&h(\mathbf{X}_1+\mathbf{Z}_1) - h\left(\mathbf{X}_1 + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_2\right)\\
&\quad= I\left(\mathbf{X}_1 + \sqrt{\left(1-\frac{1}{2\alpha^2}\right)}\mathbf{U}_2 + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_2; \mathbf{U}_2\right)\\
&\quad\le I\left(\mathbf{X}_1 + \sqrt{\left(1-\frac{1}{2\alpha^2}\right)}\mathbf{U}_2 + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_2; \mathbf{U}_2\middle|\mathbf{X}_1\right)\\
&\quad= n\log\left(\sqrt{2}|\alpha|\right)
\end{aligned}
$$

where in the inequality we used that $\mathbf{X}_1$ is independent of $\mathbf{U}_2$ and conditioning reduces differential entropy. The upper bound (30) then follows by combining the results of the two cases.

Next we upper bound the summands and the difference between the forth and fifth line of the right hand side of (29). Similarly to (30) we find that for $2 \le \ell \le J$

$$
\begin{aligned}
&h\left(\mathbf{X}_\ell + \frac{1}{\sqrt{2}}\mathbf{U}_\ell\middle|M_\ell,\ldots,M_2\right)\\
&\quad -h\left(\mathbf{X}_\ell + \frac{1}{\sqrt{2}\alpha}\mathbf{V}_{\ell+1}\middle|M_{\ell+1},\ldots,M_2\right)\\
&\qquad\qquad \le \max\{0, n\log(|\alpha|)\} \quad (32)
\end{aligned}
$$

and

$$
\begin{aligned}
&h\left(\mathbf{X}_{J+1} + \frac{1}{\sqrt{2}}\mathbf{U}_{J+1}\middle|M_{J+1},\ldots,M_2\right)\\
&\quad -h\left(X_{J+1} + \frac{1}{\alpha}\mathbf{Z}_{J+2}\middle|M_{J+2},\ldots,M_2\right)\\
&\qquad\qquad \le \max\left\{0, n\log\left(\frac{|\alpha|}{\sqrt{2}}\right)\right\}. \quad (33)
\end{aligned}
$$

Then, from (24) using (25), (26), (29), (30), (32), and (33)

$$\sum_{k=1}^{J+2} R_k \leq (J+1)\frac{1}{2}\log\left(1+\frac{P}{N}\right) + (J+1)\log\left(\frac{1+|\alpha|}{|\alpha|}\right) + J\log\left(\sqrt{2}\right) + (J-1)\max\left\{0, \log\left(|\alpha|\right)\right\} + \max\left\{0, \log\left(\sqrt{2}|\alpha|\right)\right\} + \max\left\{0, \log\left(\frac{|\alpha|}{\sqrt{2}}\right)\right\} + (J+2)\frac{\epsilon(n)}{n}, \tag{34}$$

from which (21) follows by (23) by letting n tend to infinity.

ACKNOWLEDGMENT

We would like to thank Tobias Koch and Ivana Maric for helpful discussions.

REFERENCES

[1] S. Srinivasa and Syed A. Jafar, "The throughput potential of cognitive radio - a theoretical perspective," To appear in the IEEE Communications Magazine 2007.

[2] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel", *IEEE Trans. on Inform. Theory,* vol. 40, no. 6, pp. 1713–27, Nov. 1994.

[3] Max H. M. Costa, "Writing on dirty paper," *IEEE Trans. on Inform. Theory,* vol. 29, no. 3, pp. 439–441, May 1983.

[4] N. Devroye and M. Sharif, "The multiplexing gain of MIMO X-channels with partial transmit side information," in Proc. of ISIT 2007, Nice, France, June 24–29, 2007.

[5] A. Lapidoth, S. Shamai (Shitz), and M. A. Wigger, "On Cognitive Interference Networks", submitted to ITW 2007, Lake Tahoe.