Rate-Limited Transmitter-Cooperation in Wyner's Asymmetric Interference Network

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

Michèle Wigger
Telecom ParisTech
75013 Paris, France
michele.wigger@telecom-paristech.fr

***Abstract*—We study Wyner's asymmetric interference network (soft-handoff model) when each transmitter has local, *rate-limited* side-information about the messages of the J transmitters to its left and the J transmitters to its right. We distinguish two scenarios. In Scenario A the neighbors of Transmitter k can have different, *individual*, side-information about Message M_k. In Scenario B they all have *the same* side-information about M_k. For both scenarios we derive the asymptotic multiplexing gain per-user, that is, the limiting ratio of the multiplexing gain divided by the number of users K when $K \to \infty$.**

I. Setup and Result

We consider a communication scenario with K transmitter/receiver pairs labeled $1, \ldots, K$. Each transmitter is equipped with a single transmitting antenna, and each receiver with a single receiving antenna. The symbols $x_{k,1}, \ldots, x_{k,n}$ sent at Transmitter k and the symbols $y_{k,1}, \ldots, y_{k,n}$ observed at Receiver k are assumed to be in $\mathbb{R}$. (Extending our results to a setup with complex symbols is straightforward.) We envision a network with only local, short-range interference, as e.g. in [1], [2], [3], [4], [5], [6], where far-away transmissions do not interfere. More specifically, we assume that the transmitters and the receivers are located on two parallel lines, each Receiver k opposite its corresponding Transmitter k; and the signal sent at Transmitter k is only observed at Receivers k and $(k+1)$. Thus, the time-t signal observed at Receiver k is given by:

$$Y_{k,t} = x_{k,t} + \alpha_k x_{k-1,t} + Z_{k,t}, \tag{1}$$

where $\{Z_{k,t}\}$ is a sequence of independent and identically distributed (i.i.d.) standard Gaussians; $\alpha_k \neq 0$ is a given real number; and $x_{0,t}$ is deterministically 0 for all times t.

The goal of the communication is that for each $k \in \{1, \ldots, K\}$, Transmitter k conveys a Message M_k to Receiver k. The messages $\{M_k\}_{j=1}^{K}$ are independent of each other and of the noises $\{Z_{k,t}\}$, and each M_k is uniformly distributed over the set $\mathcal{M}_k \triangleq \{1, \ldots, \lfloor e^{nR_k} \rfloor\}$. Here, n denotes the block-length and R_k the rate of Message M_k.

We impose a symmetric average block-power constraint P on the input sequences. That means, they have to satisfy

$$\frac{1}{n}\sum_{t=1}^{n} x_{k,t}^2 \leq P, \quad k \in \{1, \ldots, K\}. \tag{2}$$

The research of S. Shamai has been supported by the Israel Science Foundation (ISF)

The described network is sometimes called *Wyner's asymmetric cellular model* or *Wyner's generalized*[1] *soft-handoff model*. It has been studied (see e.g., [1], [2], [4], [5]) under various assumptions on the allowed cooperation between the transmitters and between the receivers. The works in [1], [2], [3] assume that only the receivers can cooperate but not the transmitters. In [1], [2] the receivers can fully cooperate—i.e., they can jointly decode their intended messages—whereas in [3] they can only cooperate through *clustered decoding*—i.e., each receiver can decode its desired messages based on the signals received at a certain subset of the receiving antennas. In contrast, in [4] the receivers cannot cooperate at all whereas the transmitters are allowed to cooperate through *message cognition*—i.e., each transmitter knows a certain subset of the other transmitters' messages. In [5] the transmitters can cooperate through message cognition and the receivers through clustered decoding.

Here, we consider a scenario with local cooperation between the transmitters only: as in [4], [5] each transmitter has side-information about the messages of the J transmitters to its left and the J transmitters to its right. However, here the side-information can be *imperfect* in that each transmitter k does not directly observe the messages $M_{k-J}, \ldots, M_{k-1}, M_{k+1}, \ldots, M_{k+J}$, but only side-information[2]

$$U_{k-J,k}, \ldots, U_{k-1,k},\ U_{k+1,k}, \ldots, U_{k+J,k}.$$

The transmitters learn their side-information during a first *error-free but rate-limited* communication phase which precedes the communication over the interference network. In this first communication phase each Transmitter $j \in \{1, \ldots, K\}$ sends side-information $U_{j,k}$ about its Message M_j to Transmitter $k \in \{1, \ldots, K\}$ if $1 \leq |j-k| \leq J$, where $U_{j,k}$ takes value in a discrete finite set $\mathcal{U}_{j,k}$ and is a function of M_j, i.e.,

$$U_{j,k} = \phi_{j,k}^{(n)}(M_j)$$

[1] *Generalized* refers to the fact that the cross-gains $\{\alpha_k\}$ can be different.

[2] Due to edge effects of our network, not all transmitters have J neighbors to their left and right. For readability, and since our focus is on large networks ($K \to \infty$) where the influence of these edge effects vanishes, we mostly ignore this issue. Wherever we abuse notation in this sense, we think of M_j and $U_{j,k}$ as being deterministically 0, for all $j \notin \{1, \ldots, K\}$.

for some $\phi_{j,k}^{(n)}\colon \mathcal{M}_j \to \mathcal{U}_{j,k}$. The side-information $U_{j,k}$ is perfectly observed at Transmitter k whenever the rate-constraint

$$\frac{1}{n} H(U_{j,k}) \leq \frac{\mu}{2}\log(P) \tag{3}$$

is satisfied for a given parameter $\mu \geq 0$.

Notice that our setup differs from the classical setup with conferencing encoders [7] in that here the communication in the first phase does not occur in rounds, and the side-information sent by Transmitter j can only depend on M_j.

We dinstinguish two scenarios for the first communication phase:

Scenario A: Each transmitter $j \in \{1,\dots,K\}$ can send *individual side-information* to the J neighbors to its left and its right. That means, all the side-informations $U_{j,j-J},\dots,U_{j,j-1},\ U_{j,j+1},\dots,U_{j,j+J}$ can be different.

Scenario B: Each transmitter $j \in \{1,\dots,K\}$ has to send *the same side-information* to the J neighbors to its left and its right. That means,

$$U_{j,j-J} = \ldots = U_{j,j-1} = U_{j,j+1} = \ldots = U_{j,j+J}.$$

In the second communication phase over the interference network each Transmitter k can compute its channel inputs $\boldsymbol{X}_k$ as a function of its message and the side-information:

$$X_k^n = f_k^{(n)}(U_{k-J,k},\dots,U_{k-1,k},M_k,U_{k+1,k},\dots,U_{k+J,k}), \tag{4}$$

where $f_k^{(n)}\colon \mathcal{U}_{k-J,k}\times\mathcal{U}_{k-1,k}\times\mathcal{M}_k\times\mathcal{U}_{k+1,k}\times\mathcal{U}_{k+J,k}\to\mathbb{R}^n$ denotes Transmitter k's encoding function.

We assume no cooperation at the receivers. Thus, each receiver k decodes its desired message M_k based only on the observed channel outputs $Y_k^n \triangleq (Y_{k,1},\dots,Y_{k,n})$, i.e., it produces a guess of the form

$$\hat{M}_k \triangleq \varphi_k^{(n)}(Y_k^n) \tag{5}$$

for some chosen decoding function $\varphi_k^{(n)}\colon \mathbb{R}^n \to \mathcal{M}_k$.

An error occurs in the communication whenever

$$(M_1,\dots,M_K) \neq (\hat{M}_1,\dots,\hat{M}_K).$$

The described setup models downlink communication from multiple base-stations to corresponding mobiles where the base-stations can partially cooperate by communicating over a back-haul.

For the described scenarios we say that a rate-tuple $(R_1,\dots,R_K)$ is achievable if, as the block-length n tends to infinity, the average probability of error decays to 0, i.e.,

$$\lim_{n\to 0} \Pr\left[(M_1,\dots,M_K) \neq (\hat{M}_1,\dots,\hat{M}_K)\right] = 0.$$

The closure of the set of all rate-tuples $(R_1,\dots,R_K)$ that are achievable is called the capacity region. For Scenario A we denote it by $\mathcal{C}_A(K,J,\mu;P)$ and for Scenario B by $\mathcal{C}_B(K,J,\mu;P)$. The sum-capacity is defined as the supremum of the sum-rate $\sum_{k=1}^K R_k$ over all achievable tuples $(R_1,\dots,R_K)$. It is denoted by $C_{A,\Sigma}(K,J,\mu;P)$ for Scenario A and by $C_{B,\Sigma}(K,J,\mu;P)$ for Scenario B. Our main focus in this work is on the high-SNR asymptote of the sum-capacity. That is, on the *multiplexing gains*

$$\eta_A(K,J,\mu) \triangleq \varlimsup_{P\to\infty} \frac{C_{A,\Sigma}(K,J,\mu;P)}{\frac{1}{2}\log(P)}$$

and

$$\eta_B(K,J,\mu) \triangleq \varlimsup_{P\to\infty} \frac{C_{B,\Sigma}(K,J,\mu;P)}{\frac{1}{2}\log(P)}$$

and on the *asymptotic multiplexing gains per-user*:

$$\mathcal{S}_A^\infty(J,\mu) \triangleq \varlimsup_{K\to\infty} \varlimsup_{P\to\infty} \frac{C_{A,\Sigma}(K,J,\mu;P)}{\frac{K}{2}\log(P)}$$

and

$$\mathcal{S}_B^\infty(J,\mu) \triangleq \varlimsup_{K\to\infty} \varlimsup_{P\to\infty} \frac{C_{B,\Sigma}(K,J,\mu;P)}{\frac{K}{2}\log(P)}.$$

For μ sufficiently large (e.g., $\mu=\infty$), the two scenarios are equivalent to a message-cognition setup where each transmitter knows the messages of the J transmitters to its left and to its right (i.e., to the asymmetric network in [5] specialized to $J_\ell = J_r = J$ and $i_\ell = i_r = 0$).

Theorem 1. *For Scenario A:*

$$\mathcal{S}_A^\infty(J,\mu) = \min\left\{\frac{1+2\mu}{2}, \frac{2J+1}{2J+2}\right\}. \tag{6}$$

For Scenario B:

$$\mathcal{S}_B^\infty(J,\mu) = \min\left\{\frac{1+\mu}{2}, \frac{2J+1}{2J+2}\right\}. \tag{7}$$

Proof: See Section III. ∎

Specializing Theorem 1 to $\mu=\infty$ recovers Corollary 1 in [5] specialized to $J_\ell = J_r = J$ and $i_\ell = i_r = 0$.

Notice that when μ is below a certain threshold—which depends on J and differs for the two scenarios—then the asymptotic multiplexing gain per-user grows linearly in μ (independent of J). It thus remains unchanged when J is further increased. When μ is above the corresponding threshold, then the asymptotic multiplexing gain per-user equals $\frac{2J+1}{2J+2}$ and thus remains unchanged when μ is increased.

In the linear regime $\mathcal{S}_A^\infty$ has slope μ, whereas $\mathcal{S}_B^\infty$ has slope $\mu/2$. Thus, in this linear regime, the benefit of rate-limited side-information at high powers is approximately doubled if the transmitters can send different side-information to the different adjacent transmitters.

From our proofs in Sections III-A and IV-A, one can see that the result in (6) holds also when each transmitter has to send the same side-information to all the J transmitters to its left and the same side-information to all the J transmitters to its right.

II. A Scheme for Message Cognition ($\mu = \infty$)

We describe a coding scheme for the special case of message cognition, i.e., when $\mu=\infty$. The scheme is a slight generalization of the scheme in [5]. It has parameters

$$(K, J, \kappa_0) \tag{8}$$

where κ_0 is an integer number in $\{1,\ldots,2J+2\}$.

For each set of parameters we define:

$$\gamma(K,J,\kappa_0) \triangleq \left\lceil \frac{K-\kappa_0-J-1}{2J+2} \right\rceil + 1. \tag{9}$$

As we shall see in the following:

Lemma 1. *Our scheme with message cognition achieves a multiplexing gain of*

$$K-\gamma(K,J,\kappa_0), \tag{10}$$

and thus an asymptotic multiplexing gain per-user of

$$\frac{2J+1}{2J+2}. \tag{11}$$

The scheme is based on the idea of silencing certain transmitters and using Costa's dirty-paper coding [8] to cancel known interference. We silence $\gamma(K,J,\kappa_0)$ transmitters, and let each of the remaining $K-\gamma(K,J,\kappa_0)$ transmitters k send a different message at rate $\frac{1}{2}\log(1+P)$ or (sometimes when $|\alpha_{k+1}|<1$) at rate $\frac{1}{2}\log(1+\alpha_{k+1}^2 P)$. Such a scheme achieves the desired multiplexing gain and asymptotic multiplexing gain per-user in Lemma 1.

Before explaining which transmitters are silenced, we define

$$\beta \triangleq 2J+2, \tag{12}$$

$$\kappa_1 \triangleq (K-\kappa_0) \mod \beta \tag{13}$$

$$\gamma_{\mathrm{g}} \triangleq \frac{K-\kappa_0-\kappa_1}{2J+2}. \tag{14}$$

where $\mod$ denotes the modulus-operator. We silence Transmitters $\{\kappa_0+j\beta\}$, for $j\in\{0,\ldots,\gamma_{\mathrm{g}}-1\}$; moreover, if $\kappa_1>(J+1)$ we also silence Transmitter K. This splits the network into γ_{g} non-interfering subnets if $\kappa_1=0$ and into $\gamma_{\mathrm{g}}+1$ subnets if $\kappa_1>0$.

The first subnet consists of κ_0-1 active transmitting antennas and of κ_0 receiving antennas. The next $\gamma_{\mathrm{g}}-1$ subnets all have the same topology. They consist of $2J+1$ active transmitting antennas and $2J+2$ receiving antennas. We refer to these subnets as *generic* subnets. If $\kappa_1>0$ there is an additional last subnet with κ_1 or κ_1-1 active transmitting antennas, depending on whether $\kappa_1\le(J+1)$ or $\kappa_1>(J+1)$, and with κ_1 receiving antennas. We refer to subnets with less than $2J+1$ active transmitting antennas as *reduced* subnets. Notice that if $\kappa_0=2J+2$, then the first subnet is generic, otherwise it is reduced.

As we shall see, in our scheme each transmitter ignores the part of its side-information pertaining to the messages transmitted in other subnets. Likewise, each receiver ignores the outputs of antennas belonging to subnets other than its own. Therefore, we can describe our scheme for each subnet separately. For brevity we only describe our scheme for a generic subnet; the schemes for the first subnet, if it is reduced, and for the (γ_g+1)-th subnet, if it exists, are similar.

To simplify description, we assume that $K\ge\kappa_0=2J+2$, such that the first subnet is generic, and describe the scheme for the first subnet.

When $\kappa_0=2J+2$, we transmit Messages $M_1,\ldots,M_{J+1}$ and $M_{J+3},\ldots,M_{2J+2}$ in the first subnet. Messages $1,\ldots,J+1$ are transmitted as follows.

- Transmitter 1 sends its Message using a point-to-point Gaussian code.
- For each $k=2,\ldots,J+1$, Transmitter k can use its side-information to compute the interference term $\alpha_{k-1}X_{k-1}^n$. Indeed, as we shall see shortly, in our scheme the input sequence X_{k-1}^n depends only on messages $M_1,\ldots,M_{k-1}$, and these messages are known to Transmitter k, because $k-1\le J$ for all $k=2,\ldots,J+1$.
- For each $k=2,\ldots,J+1$, Transmitter k uses a dirty-paper code of power P and rate $R_k=\frac{1}{2}\log(1+P)$ to transmit its message M_k and mitigate the interference $\alpha_{k-1}X_{k-1}^n$ experienced at the antenna of Receiver k.
- Receiver 1 can decode its message based on the interference-free output Y_1^n.
- For each $k=2,\ldots,J+1$, Receiver k decodes Message M_k applying dirty-paper decoding to the outputs Y_k^n.

If $J>0$, then Messages $M_{J+3},\ldots,M_{2J+2}$ are transmitted as follows. Define $\mathcal{T}_2=\{J+2,\ldots,2J+1\}$.

- For each $k\in\mathcal{T}_2$, Transmitter k can use its side-information to compute the interference term X_{k+1}^n. As seen shortly, in our scheme the input sequence X_{k+1}^n depends only on messages $M_{k+2},\ldots,M_{2J+2}$, which are all known to Transmitter k because $(2J+2-k)\le J$ for all $k\in\mathcal{T}_2$.
- For each $k\in\mathcal{T}_2$, Transmitter k uses a dirty-paper code of power P and rate $R_{k+1}=\min\left\{\frac{1}{2}\log(1+P),\frac{1}{2}\log(1+\alpha_{k+1}^2P)\right\}$ to transmit Message M_{k+1} and mitigate the "interference" X_{k+1}^n experienced at the antenna of Receiver $(k+1)$.
- For each $k\in\mathcal{T}_2$, Receiver $(k+1)$ decodes Message M_{k+1} applying dirty-paper decoding based on the output sequence Y_{k+1}^n.

This scheme achieves multiplexing gain $2J+1$ over a generic subnet.

Similar schemes achieve multiplexing gain κ_0-1 over the first subnet, and multiplexing gain κ_1 or κ_1-1 over the last subnet, depending on whether $\kappa_1\le J+1$ or $\kappa_1>J+1$. This establishes Lemma 1. We notice the following:

Remark 1. *In our scheme, each message is transmitted at a rate not exceeding* $\frac{1}{2}\log(1+P)$.

For positive integers k,ℓ and a set of integers $\mathcal{S}$, we write $k\in_\ell\mathcal{S}$ when

$$(k \mod \ell)=(s \mod \ell),\ \text{for some } s\in\mathcal{S}.$$

Remark 2. *For the scheme to work it suffices that each Message M_k is known to Transmitter k and to:*

- *the J transmitters to the* left *of Transmitter k, if*
$$k\in_{(2J+2)}\{J+3+\kappa_0,\ldots,2J+2+\kappa_0\}, \tag{15}$$
- *the J transmitters to the* right *of Transmitter k, if*
$$k\in_{(2J+2)}\{1+\kappa_0,\ldots,J+\kappa_0\}. \tag{16}$$

III. Achievability of Theorem 1

A. Scenario A

Our scheme is based on a rate-splitting/time-sharing approach. We assume that each Message M_k can be represented as a sequence of independent submessages $(M_k^{(1)}, \ldots, M_k^{(2J+3)})$, where each $M_k^{(i)}$ is uniformly distributed over $\{1, \ldots, \lfloor e^{nR_k^{(i)}} \rfloor\}$. Notice that the sum of the rates $R_k^{(1)} + \ldots + R_k^{(2J+3)}$ tends to R_k as $n \to \infty$.

We now describe the first communication phase. In our scheme each transmitter sends the same side-information to the J transmitters to its left and the same side-information to the J transmitters to its right. Thus, for each $k \in \{1, \ldots, K\}$:

$$U_{k,\text{left}} \triangleq U_{k,k-J} = \ldots = U_{k,k-1} \tag{17}$$

and

$$U_{k,\text{right}} \triangleq U_{k,k+1} = \ldots = U_{k,k+J}. \tag{18}$$

Transmitter k chooses $U_{k,\text{left}}$ to losslessly describe all its submessages $M_k^{(i)}$ that have superscript $i \in \{1, \ldots, 2J+2\}$ satisfying

$$k \in_{(2J+2)} \{J+3+i, \ldots, 2J+2+i\}. \tag{19}$$

Therefore, in our scheme

$$\frac{1}{n} H(U_{k,k-J}) = \ldots = \frac{1}{n} H(U_{k,k-1}) \leq \sum_{\substack{i \in \{1,\ldots,2J+2\}: \\ \text{satisfying (19)}}} R_k^{(i)}. \tag{20}$$

Similarly, Transmitter k chooses $U_{k,\text{right}}$ to losslessly describe all its submessages $M_k^{(i)}$ that have superscript $i \in \{1, \ldots, 2J+2\}$ satisfying:

$$k \in_{(2J+2)} \{1+i, \ldots, J+i\}. \tag{21}$$

Therefore,

$$\frac{1}{n} H(U_{k,k+1}) = \ldots = \frac{1}{n} H(U_{k,k+J}) \leq \sum_{\substack{i \in \{1,\ldots,2J+2\}: \\ \text{satisfying (21)}}} R_k^{(i)}. \tag{22}$$

As we shall see, in our scheme, for all $i \in \{1, \ldots, 2J+2\}$:

$$R_k^{(i)} \leq \frac{\mu}{J} \cdot \frac{1}{2} \log(1+P). \tag{23}$$

Combining (19)–(23), we thus conclude that the communication during this first phase respects the rate-constraint (3), and is hence error-free.

We now describe the scheme in the second phase when communicating over the interference network. We use a time-sharing scheme where we split the block-length n into $(2J+3)$ subblocks. The first $J+2$ subblocks are of length

$$N_1 \triangleq \min\left\{ \left\lfloor \frac{\mu n}{J} \right\rfloor, \left\lfloor \frac{n}{J+2} \right\rfloor \right\} \tag{24}$$

and the last subblock is of length $N_2 \triangleq n - (2J+2)N_1$. In subblock $i \in \{1, \ldots, 2J+2\}$ we send messages $M_1^{(i)}, \ldots, M_K^{(i)}$ using the scheme in Section II for parameters (K, J, i). Notice that we can apply these schemes for message cognition, because the transmitters have exchanged the required submessages during the first phase. This can be seen by comparing Remark 2 with Equations (17), (18), (19), and (21). Moreover, by Remark 1 and because each of the considered subblocks is of length N_1, the rates $R_k^{(i)}$ for $k \in \{1, \ldots, K\}$ and $i \in \{1, \ldots, 2J+2\}$ satisfy (23) as required.

In the last subblock we send messages $M_1^{(2J+3)}, \ldots, M_K^{(2J+3)}$ using the scheme in Section II for parameters $(K, 0, 2)$. This does not require any side-information at the transmitters.

We analyze the asymptotic multiplexing gain per-user achieved by our scheme. By Lemma 1 our scheme achieves an asymptotic multiplexing gain per-user of $\frac{2J+1}{2J+2}$ in each of the first $(2J+2)$ subblocks and an asymptotic multiplexing gain per-user of $1/2$ in the last subblock. Thus, on average we achieve an asymptotic multiplexing gain per-user of

$$\min\left\{\mu \frac{2J+2}{J}, 1\right\} \frac{2J+1}{2J+2} + \left(1 - \min\left\{\mu \frac{2J+2}{J}, 1\right\}\right) \frac{1}{2}$$
$$= \min\left\{ \frac{1}{2} + \mu, \frac{2J+1}{2J+2} \right\},$$

which proves the achievability part of (6).

B. Scenario B

We apply the same scheme as for Scenario A, except that we replace the definition of N_1 in (24) by

$$N_1 \triangleq \min\left\{ \left\lfloor \frac{\mu n}{2J} \right\rfloor, \left\lfloor \frac{n}{J+2} \right\rfloor \right\}$$

and here we choose $U_{k,\text{left}} = U_{k,\text{right}}$ to losslessly describe all the submessages $M_k^{(i)}$ that have superscript $i \in \{1, \ldots, 2J+2\}$ satisfying either (19) or (21). Details omitted.

IV. Converse to Theorem 1

A. Scenario A

We prove the converse to (6), i.e.,

$$\mathcal{S}_A^{\infty}(J, \mu) \leq \min\left\{ \frac{1+2\mu}{2}, \frac{2J+1}{2J+2} \right\}. \tag{25}$$

Since with message cognition we cannot do worse than with rate-limited side-information, from [5] (Corollary 1 specialized to $i_\ell = i_t = 0$ and $J_\ell = J_r = J$) we have:

$$\mathcal{S}_A^{\infty}(J, \mu) \leq \frac{2J+1}{2J+2}. \tag{26}$$

In the following we prove

$$\mathcal{S}_A^{\infty}(J, \mu) \leq \frac{1+2\mu}{2}. \tag{27}$$

The proof of (27) is similar to a converse proof in [4]. We first let a genie reveal some genie-information $\boldsymbol{G}_k$ (specified shortly) to Receiver k, for $k \in \{1, \ldots, K\}$. Denoting the capacity and the asymptotic multiplexing gain per-user of the resulting genie-aided network by $\mathcal{C}_{A,\text{Genie}}(K, J, \mu; P)$ and $\eta_{A,\text{Genie}}(J, \mu)$, we obviously have

$$\mathcal{C}_A(K, J, \mu; P) \subseteq \mathcal{C}_{A,\text{Genie}}(K, J, \mu; P), \tag{28}$$

because the receivers can always ignore the genie-information. In the following we show that

$$\mathcal{S}^{\infty}_{A,\text{Genie}}(J,\mu) \le \frac{1+2\mu}{2}, \tag{29}$$

which combined with (28) establishes Inequality (27).

For $i \in \{1,\ldots,\lfloor\frac{K}{2}\rfloor\}$ we define the genie-information $\boldsymbol{G}_{2i}$ as the set containing

$$M_{2i-2},\ \{U_{j,2i-2}\}_{j=2i-2-J}^{2i-3},\ U_{2i-1,2i-2},\ \{U_{j,2i-2}\}_{j=2i+1}^{2i+J-2}, \tag{30}$$

and

$$\{U_{j,2i}\}_{j=2i-J}^{2i-1}, \{U_{j,2i}\}_{j=2i+1}^{2i+J},\ (Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n); \tag{31}$$

and $\boldsymbol{G}_{2i-1}$ as the set containing

$$M_{2i-2},\ \{U_{j,2i-2}\}_{j=2i-2-J}^{2i-3},\ \{U_{j,2i-2}\}_{j=2i}^{2i+J-2}, U_{2i,2i-1}, \tag{32}$$

and

$$\{U_{j,2i}\}_{j=2i-J}^{2i-2}, \{U_{j,2i}\}_{j=2i+1}^{2i+J},\ (Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n); \tag{33}$$

If K is odd, then we additionally define $\boldsymbol{G}_K$ as the empty set.

Remark 3. *From the tuple $(\boldsymbol{G}_{2i}, M_{2i}, Y_{2i}^n)$ it is possible to compute the outputs Y_{2i-1}^n. This holds because from $(\boldsymbol{G}_{2i}, M_{2i})$ one can compute the inputs X_{2i-2}^n and X_{2i}^n, because $\boldsymbol{G}_{2i}$ contains $(Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n)$, and because:*

$$Y_{2i-1}^n = \frac{1}{\alpha_{2i}}(Y_{2i}^n - X_{2i}^n) + \alpha_{2i-1}X_{2i}^n + \left(Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n\right). \tag{34}$$

Remark 4. *The genie-informations $\boldsymbol{G}_{2i-1}$ and $\boldsymbol{G}_{2i}$ differ only in that $U_{2i,2i-1}$ and $U_{2i,2i-2}$ are contained in $\boldsymbol{G}_{2i-1}$ but not in $\boldsymbol{G}_{2i}$ and that $U_{2i-1,2i}$ and $U_{2i-1,2i-2}$ are contained in $\boldsymbol{G}_{2i}$ but not in $\boldsymbol{G}_{2i-1}$.*

We show that whenever we have reliable communication over the genie-aided network, then for $i \in \{1,\ldots,\lfloor\frac{K}{2}\rfloor\}$:

$$\overline{\lim_{P\to\infty}} \frac{R_{2i-1}+R_{2i}}{1/2\log(P)} \le 1+2\mu. \tag{35}$$

From (35), the desired lower bound (29) is easily obtained.

To establish (35), we notice that by Fano's inequality, whenever we have reliable communication:

$$\begin{aligned}&(R_{2i-1}+R_{2i}) - \epsilon(n)\\ &\qquad \le \frac{1}{n} I(Y_{2i-1}^n; M_{2i-1}|\boldsymbol{G}_{2i-1}) + \frac{1}{n} I(Y_{2i}^n; M_{2i}|\boldsymbol{G}_{2i})\end{aligned} \tag{36}$$

where $\epsilon(n)$ tends to 0 as $n \to \infty$. Moreover,

$$\begin{aligned}
&\frac{1}{n} I(Y_{2i-1}^n; M_{2i-1}|\boldsymbol{G}_{2i-1}) + \frac{1}{n} I(Y_{2i}^n; M_{2i}|\boldsymbol{G}_{2i})\\
&\le \frac{1}{n} h(Y_{2i-1}^n|\boldsymbol{G}_{2i-1}) - \frac{1}{n} h(Z_{2i-1}^n|\boldsymbol{G}_{2i-1})\\
&\quad + \frac{1}{n} h(Y_{2i}^n) - \frac{1}{n} h(Y_{2i}^n|M_{2i}, \boldsymbol{G}_{2i}) && (37)\\
&= \frac{1}{n} h(Y_{2i}^n) - \frac{1}{n} h\Big(Z_{2i-1}^n | Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n\Big)\\
&\quad + \frac{1}{n} h(Y_{2i-1}^n|\boldsymbol{G}_{2i-1}) - \frac{1}{n} h(Y_{2i-1}^n|M_{2i}, \boldsymbol{G}_{2i}) + \log|\alpha_{2i}| && (38)\\
&= \frac{1}{n} h(Y_{2i}^n) - \frac{1}{n} h\Big(Z_{2i-1}^n | Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n\Big)\\
&\quad + \frac{1}{n} h(Y_{2i-1}^n|\boldsymbol{G}_{2i-1}) - \frac{1}{n} h(Y_{2i-1}^n|U_{2i,2i-1}, U_{2i,2i-2}, \boldsymbol{G}_{2i})\\
&\quad + \log|\alpha_{2i}| && (39)\\
&= \frac{1}{n} h(Y_{2i}^n) - \frac{1}{n} h\Big(Z_{2i-1}^n | Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n\Big)\\
&\quad + \frac{1}{n} I(Y_{2i-1}^n; U_{2i-1,2i}, U_{2i-1,2i-2}|\boldsymbol{G}_{2i-1}) + \log|\alpha_{2i}| && (40)\\
&\le \frac{1}{n} h(Y_{2i}^n) - \frac{1}{n} h\Big(Z_{2i-1}^n | Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n\Big)\\
&\quad + \frac{1}{n} H(U_{2i-1,2i}) + \frac{1}{n} H(U_{2i-1,2i-2}) + \log|\alpha_{2i}| && (41)\\
&\le \frac{1}{2}\log\left(1+(1+|\alpha_{2i}|)^2 P\right) + \frac{1}{2}\log(1+\alpha_{2i}^2)\\
&\quad + \frac{2\mu}{2}\log(1+P) + \log|\alpha_{2i}|, && (42)
\end{aligned}$$

where Inequality (37) follows because conditioning cannot increase differential entropy; Equality (38) follows by rearranging terms, by the fact that Z_{2i-1}^n depends on $\boldsymbol{G}_{2i}$ only through $(Z_{2i-1}^n - \frac{1}{\alpha_{2i}} Z_{2i}^n)$, and by Remark 3 and Equation (34); Equality (39) follows because Y_{2i-1}^n depends on M_{2i} only through $U_{2i,2i-1}$ and $U_{2i,2i-2}$; Equality (40) follows by Remark 4; Inequality (41) follows because $U_{2i-1,2i}$ and $U_{2i-1,2i-2}$ are discrete and thus $H(U_{2i-1,2i}, U_{2i-1,2i-2}|\boldsymbol{G}_{2i-1}, Y_{2i-1}^n) \ge 0$, and because conditioning cannot increase entropy; and finally Inequality (42) follows from the Max-Entropy Theorem and from (3). Combining Inequalities (42) and (36) establishes (35) and concludes the proof.

B. Scenario B

The converse for Scenario B is similar to the converse for Scenario A. In addition, it also requires noting that here:

$$H(U_{2i-1,2i-2}, U_{2i-1,2i}) \le \frac{\mu}{2}\log(1+P)$$

because $U_{2i-1,2i-2} = U_{2i-1,2i}$. The details are omitted.

References

[1] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, pp. 1713–1727, Nov. 1994.

[2] S. V. Hanly and P. A. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommun. Syst.*, vol. 1, pp. 1–42, 1993.

[3] N. Levy and S. Shamai (Shitz), "Clustered local decoding for Wyner-type cellular models," *IEEE Trans. on IT*, vol. 55, no. 11, pp. 4967–4985, 2009.

[4] A. Lapidoth, S. Shamai (Shitz), and M. Wigger, "A linear interference network with local side-information," in *Proc. of 2007 IEEE Intern. Symposium on Inf. Theory (ISIT'07)*, (Nice, France), June 24–29, 2007.

[5] A. Lapidoth, N. Levy, S. Shamai (Shitz), and M. Wigger, "A cognitive network with clustered decoding," in *ISIT 2009)*, (Seoul, Korea), June 28–July 3 2009.

[6] D. Gesbert, S. Hanly, H. Huang, S. Shamai (Shitz), O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, pp. 1380–1408, Dec. 2010.

[7] F. M. J. Willems, "The discrete memoryless multiple access channel with partially cooperating encoders," *IEEE Trans. on IT*, Nov. 1983.

[8] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, pp. 439–441, May 1983.