Cognitive Wyner Networks With Clustered Decoding

Amos Lapidoth, *Fellow, IEEE*, Nathan Levy, Shlomo Shamai (Shitz), *Fellow, IEEE*, and Michèle Wigger, *Member, IEEE*,

***Abstract*— We study an interference network where equally numbered transmitters and receivers lie on two parallel lines, with each transmitter opposite its intended receiver. We consider two short-range interference models: the asymmetric network, where the signal sent by each transmitter is interfered only by the signal sent by its left neighbor (if present), and a symmetric network, where it is interfered by both its left and its right neighbors. Each transmitter is cognizant of its own message, the messages of the t_ℓ transmitters to its left, and the messages of the t_r transmitters to its right. Each receiver decodes its message based on the signals received at its own antenna, at the r_ℓ receive antennas to its left, and at the r_r receive antennas to its right. For such networks, we provide upper and lower bounds on the multiplexing gain, i.e., on the high signal-to-noise ratio asymptotic logarithmic growth of the sum-rate capacity. In some cases, our bounds coincide, e.g., for the asymmetric network. Our results exhibit an equivalence between the transmitter side-information parameters t_ℓ, t_r and the receiver side-information parameters r_ℓ, r_r in the sense that increasing/decreasing t_ℓ or t_r by a positive integer δ has the same effect on the multiplexing gain as increasing/decreasing r_ℓ or r_r by δ. Moreover—even in asymmetric networks—there is an equivalence between the left side-information parameters (t_ℓ, r_ℓ) and the right side-information parameters (t_r, r_r).**

***Index Terms*— Clustered decoding, dirty-paper coding, interference networks, successive interference cancellation, message cognition, multiplexing gain.**

I. Introduction

We CONSIDER a cellular mobile communication system (either uplink or downlink) where K cells are positioned on a linear array. We assume short-range interference where the signals sent in a given cell interfere only with the signals sent in the left-adjacent cell and/or the right-adjacent cell, depending on the position of the mobiles within the cells. Our goal is to determine the throughput of such a cellular system at high signal-to-noise ratio (SNR).

The high-SNR throughput of our system (where we assume constant non-fading channel gains) does not depend on the number of mobiles in a cell (provided this number is not zero), because in each cell there is only one base station. Therefore, we restrict attention to setups with only one mobile per cell.

We focus on two regular setups. The first setup exhibits *asymmetric interference*: the communication in a cell is only interfered by the signals sent in the cell to its left but not by the signals sent in the cell to its right (e.g., because we model the uplink and all the mobiles lie close to the right border of their cells). The second setup exhibits *symmetric interference*: the communication in a cell is interfered by the signals sent in the cells to its left as well as to its right (e.g., because the mobiles lie in the center of their cells). The symmetric setup was introduced in [1] and [2].

On a more abstract level, our communication scenario is described as follows: K transmitters wish to communicate independent messages to their K corresponding receivers, and it is assumed that these communications interfere. Moreover, the K transmitters are assumed to be located on a horizontal line, and the K receivers are assumed to lie on a parallel line, each receiver opposite its corresponding transmitter. We consider two specific networks. In the *asymmetric network*, each receiver observes a linear combination of the signals transmitted by its corresponding transmitter, the signal of the transmitter to its left, and additive white Gaussian noise (AWGN). See Figure 1. In the *symmetric network*, each receiver observes a linear combination of the signal transmitted by its corresponding transmitter, the two signals of the transmitter to its left and the transmitter to its right, and AWGN. See Figure 2. The symmetric network is also known as Wyner's linear model or the full Wyner model; the asymmetric network is known as the asymmetric Wyner model or the soft hand-off model.

In [1] and [2] the receivers were allowed to *fully* cooperate in their decoding, and thus the communication scenario was modeled as a multiple-access channel (MAC). In contrast, here we assume that each receiver has to decode its message individually, and therefore our communication scenario is modeled as an interference network. However, we still allow for *partial* cooperation between neighboring receivers where neighboring receivers can cooperate in the form of *clustered*

Manuscript received March 14, 2012; revised June 5, 2014; accepted July 13, 2014. Date of publication July 31, 2014; date of current version September 11, 2014. This work was supported in part by the Israel Science Foundation, by the European Commission within the framework of the FP7 Network of Excellence in Wireless Communications NEWCOM++ and NEWCOM#, and in part by the City of Paris through the Program Emergences. This paper was presented at the 2009 Information Theory and Applications Workshop and 2009 IEEE International Symposium on Information Theory.

A. Lapidoth is with the Department of Information Technology and Electrical Engineering, ETH Zurich, Zurich 8092, Switzerland (e-mail: lapidoth@isi.ee.ethz.ch).

N. Levy was with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel (e-mail: nhlevy@gmail.com).

S. Shamai is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

M. Wigger was with the Department of Information Technology and Electrical Engineering, ETH Zurich, Zurich 8092, Switzerland. She is now with the Department of Communications and Electronics, Telecom ParisTech, Paris 75013, France (e-mail: michele.wigger@telecom-paristech.fr).

Communicated by W. Yu, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2014.2345070

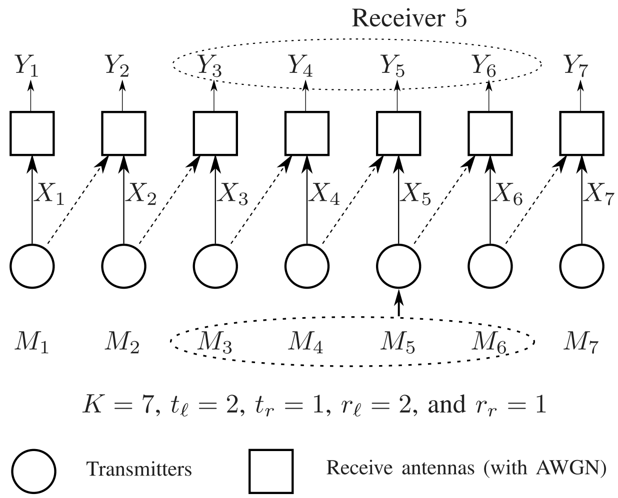

Fig. 1. Asymmetric network.

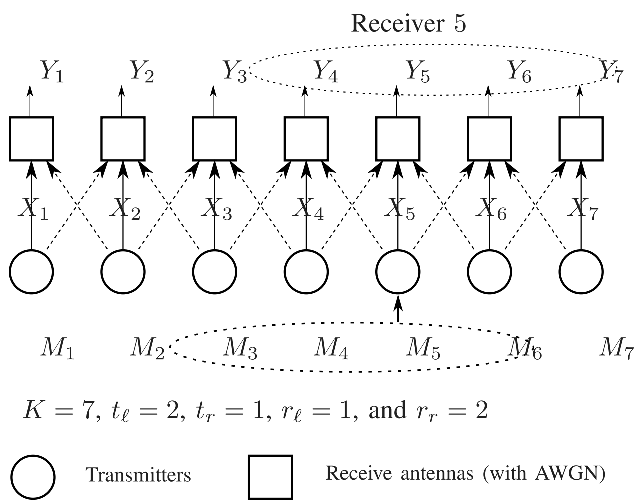

Fig. 2. Symmetric network.

local decoding. That means that, in addition to its own antenna, each receiver also has access to the antennas of some of the receivers to its left and to its right. Similarly, we also allow for (partial) cooperation between the transmitters in the form of *message cognition*. That means that, in addition to its own message, each transmitter is also cognizant of the messages of some transmitters to its left and to its right. The described scenario with message cognition and clustered local decoding may arise in the uplink as well as in the downlink of cellular mobile systems because the base stations can communicate over a backhaul and the mobiles can communicate using bluetooth connections. Thus, in an uplink scenario the transmitting mobiles can share their messages using the bluetooth links before communicating to their corresponding base stations and the receiving base stations can share their observed signals over the backhaul. In a downlink scenario the receiving mobiles can use the bluetooth connections to relay their observed signals to the mobiles in neighboring cells and the transmitting base stations can use the backhaul to share their messages.

Notice that the described model represents a combination of the cognitive model in [3] and the clustered decoding model in [4]. Also, clustered local processing is in a way a compromise between the joint (multi-cell) decoding in [1] and [2] and the single (single-cell) decoding in [3] and [5]. Clustered decoding has also been considered in [6] for fully-connected interference networks. The cognitive transmitter model considered here has been refined in [7], where the transmitters can exchange parts of their messages prior to the actual communication over rate-limited pipes, similar to [8]–[12].

Our focus in this paper is on the asymptotic behavior of the sum-rate capacity of these networks as captured by two figures of merit: the multiplexing gain and the asymptotic multiplexing gain per user. The latter is defined as the multiplexing gain of a network divided by the number of transmitter/receiver pairs K in the asymptotic regime of large K. For both networks we provide upper and lower bounds on both figures of merit.

For the asymmetric network our upper and lower bounds coincide and thus yield the exact multiplexing gain and asymptotic multiplexing gain per user. The results exhibit an equivalence between cooperation at the transmitters and cooperation at the receivers. Moreover, although the network is asymmetric, the asymptotic multiplexing gain per user also exhibits an equivalence between the transmitters' information about their right-neighbors' messages and their information about their left-neighbors' messages. Likewise, they also exhibit, an equivalence between the receivers' information about the signals observed at their right-neighbors' antennas and their information about the signals observed at their left-neighbors' antennas.

For the symmetric network our upper and lower bounds coincide only in some special cases. In these special cases the multiplexing gain—and thus also the asymptotic multiplexing gain per user—again exhibits an equivalence between cooperation at the transmitters and cooperation at the receivers. For the symmetric network, we mostly assume that the nonzero cross-gains are all equal. Our techniques extend to general cross-gains, but the statement of the results becomes cumbersome and is therefore omitted. Instead, we also consider a random model where the cross-gains are drawn from a continuous distributions. Our main results continue to hold (with probability 1) for this randomized setup.

For large number of users, i.e. $K \gg 1$, our multiplexing-gain results are of the form $\mathcal{S}_\infty \cdot K + o(K)$, where $o(K)$ denotes a function that grows sublinearly in K. As we shall see, $\mathcal{S}_\infty \in [0.5, 1]$ is strictly monotonic in the side-information parameters t_ℓ, t_r, r_ℓ, r_r, and thus if we increase one or several of the side-information parameters, then also the factor $\mathcal{S}_\infty$ increases.[1] The results in [13]–[16] suggest that this strict monotonicity relies on the weak connectivity of the network, i.e., the fact that there are relatively few interference links. Indeed, [13]–[16] show that for *fully-connected networks*, i.e., when all the transmitted signals interfere at all received

[1]The parameter $\mathcal{S}_\infty$ is called the *asymptotic multiplexing-gain per user* and will be introduced formally in the next section.

signals, and when there is no clustering at the receivers ($r_\ell = r_r = 0$), then for the side-information pattern considered here, $\mathcal{S}_\infty = 1/2$ irrespective of t_ℓ and t_r. This result holds even in the stronger setup where for each message we can choose the set of $t_\ell + t_r + 1$ adjacent transmitters that are cognizant of this message [16]. In this stronger setup, a given transmitter may not even know the message intended for its corresponding receiver. And, indeed, sometimes (for example for the networks considered here), the multiplexing gain can be increased by assigning a given Message M_k that is intended for Receiver k to a subset of $t_\ell + t_r + 1$ transmitters that does not contain Transmitter k [16]. We will describe this in more detail after describing our results.

General interference networks with transmitter cooperation have also been studied in [17]–[20]. In particular, in [17], the authors completely characterized the set of networks and transmitter side-informations that have full multiplexing gain K or multiplexing gain $K-1$. In [20], a network is presented where adding an interference link to the network—while keeping the same transmitter side-informations—can increase the multiplexing gain.

The asymmetric network has also been studied by Liu and Erkip [21], with a focus on finite-SNR results but without transmitter cognition or clustered decoding. For general $K \geq 3$, [21] characterizes the maximum sum-rate that is achievable using a simple Han-Kobayashi scheme without time-sharing and where the inputs follow a Gaussian distribution. For $K = 3$, they show that this scheme achieves the sum-capacity in *noisy-interference and mixed-interference regimes* and it achieves the entire capacity region in a *strong interference regime*. Zhou and Yu [22] considered a cyclic version of this model where the K-th transmitted signal also interferes with the first receive signal, i.e., the interference pattern is cyclic. In [22], an expression for the Han-Kobayashi region with arbitrary (also non-Gaussian) inputs was presented. It was shown that this achievable region is within 2 bits of the K-user cyclic asymmetric network in the *weak-interference regime*. In the strong interference regime, it achieves capacity. (In their achievability proofs it suffices to consider Gaussian inputs.) For $K = 3$ the authors also present an improved Han-Kobayashi scheme involving time-sharing that achieves rates within 1.5 bits of capacity. Finally, [22] also characterizes the generalized degrees of freedom (GDoF) of the symmetric capacity assuming that all cross-gains in the network are equal. Interestingly, this result shows that the GDoF of the K-user cyclic asymmetric Wyner network with equal cross-gains has the same GDoF as the standard two-user interference channel [23].

Other related results on Wyner-type networks can be found in [24]–[34].

The lower bounds in our paper are based on coding strategies that silence some of the transmitters and thereby split the network into non-interfering subnetworks that can be treated separately. Depending on the considered setup, a different scheme is then used for the transmission in the subnetworks. In some setups, some of the messages are transmitted using an interference cancellation scheme and the others are transmitted using Costa's dirty-paper coding. (Costa's dirty paper coding can also be replaced by a simple linear beamforming scheme as in [17], see also [35]–[37].) In other setups, the messages are transmitted using one of the following elementary bricks of multi-user information theory depending on the available side-information: an optimal multi-input/multi-output (MIMO) scheme, an optimal MIMO multi-access scheme, or an optimal MIMO broadcast scheme. Introducing also Han-Kobayashi type ideas to our coding strategies might improve the performance of our schemes for finite SNR.

Our upper bounds rely on an extension of Sato's multi-access channel (MAC) bound [38] to more general interference networks with more than two transmitters and receivers and where the transmitters and the receivers have side-information (see also [3], [16], [39], and in particular [17, Lemma 1 and Theorem 3]). More specifically, we first partition the K receivers into groups $\mathcal{A}$ and $\mathcal{B}_1, \ldots, \mathcal{B}_q$, and we allow the receivers in Group $\mathcal{A}$ to cooperate. We then let a genie reveal certain linear combinations of the noise sequences to the receivers in Group $\mathcal{A}$. Finally, we request that the receivers in Group $\mathcal{A}$ jointly decode all messages $M_1, \ldots, M_K$ whereas all other receivers do not have to decode anything. We choose the linear combinations that are revealed by the genie so that, for each $i = 1, \ldots, q$, if the receivers in Group $\mathcal{A}$ can successfully decode their own messages and the messages intended for the receivers in groups $\mathcal{B}_1, \ldots, \mathcal{B}_{i-1}$, then they can also reconstruct the outputs observed at the receivers in Group $\mathcal{B}_i$. In this case they can also decode the messages intended for the receivers in Group $\mathcal{B}_i$ at least as well as the receivers in Group $\mathcal{B}_i$. This iterative argument is used to show that the capacity region of the resulting MAC is included in the capacity region of the original network. The upper bound is then concluded by upper bounding the multiplexing gain of the MAC.

We conclude this section with notation and an outline of the paper. Throughout the paper, $\mathbb{R}$, $\mathbb{N}$, and $\mathbb{N}_0$ denote the sets of real numbers, natural numbers, and nonnegative integers. Their m-fold Cartesian products are denoted $\mathbb{R}^m$, $\mathbb{N}^m$, and $\mathbb{N}_0^m$. Also, $\log(\cdot)$ denotes the natural logarithm, and $a \mod b$ denotes the remainder in the Euclidean division of a by b. Random variables are denoted by upper case letters, their realizations by lower case letters. Vectors are denoted by bold letters: random vectors by upper case bold letters and deterministic vectors by lower case bold letters. Given a sequence of random variables $X_1, \ldots, X_n$ we denote by X^n the tuple $(X_1, \ldots, X_n)$ and by $\mathbf{X}$ the n-dimensional column-vector $(X_1, \ldots, X_n)^\intercal$. For sets we use calligraphic symbols, e.g., $\mathcal{A}$. The difference of two sets $\mathcal{A}$ and $\mathcal{B}$ is denoted $\mathcal{A}\backslash\mathcal{B}$. We further use the Landau symbols, and thus $o(x)$ denotes a function that grows sublinearly in x.

The paper is organized as follows. In Section II we describe the channel model and the results for the asymmetric network and in Section III the channel model and the results for the symmetric network. In Section IV we present a *Dynamic-MAC Lemma* that we use to prove our converse results for the multiplexing-gain. In the rest of the paper we prove our results: in Section V the results for the asymmetric network; in Section VI the achievability results for the symmetric network with symmetric side-information; in Section VII the achievability results for the symmetric network with general side-information; and finally in Section VIII the converse results for the symmetric network with general side-information parameters.

II. Asymmetric Network

A. Description of the Problem

We consider K transmitter/receiver pairs that are labeled from $\{1,\ldots,K\}$. The purpose of the communication is that each transmitter $k \in \{1,\ldots,K\}$ will convey its message M_k to its intended receiver k. The messages $\{M_k\}_{j=1}^K$ are assumed to be independent with M_k being uniformly distributed over the set $\mathcal{M}_k \triangleq \{1,\ldots,\lfloor e^{nR_k}\rfloor\}$, where n denotes the block-length of transmission and R_k the rate of transmission of Transmitter k.

We assume that the transmitters and receivers are all equipped with a single antenna and that the channels are discrete-time and real-valued. Denoting by $x_{k,t}$ the time-t symbol transmitted by Transmitter k, and by $Y_{k,t}$ the time-t symbol received by Receiver k,

$$Y_{k,t} = x_{k,t} + \alpha_k x_{k-1,t} + N_{k,t}, \quad k \in \{1,\ldots,K\}, \tag{1}$$

where the K noise sequences $\{N_{1,t}\},\ldots,\{N_{K,t}\}$ are independent with each comprising independent and identically distributed standard Gaussians; and where to simplify notation we define $x_{0,t}$ to be deterministically 0 for all times t. Thus, the communication of the k-th transmitter/receiver pair is interfered only by the communication of the transmitter/receiver pair to its left; see Figure 1.

It is assumed that, in addition to its own message, each transmitter is also cognizant of the messages transmitted by the $t_\ell \geq 0$ transmitters to its left and by the $t_r \geq 0$ transmitters to its right. Thus, for every $k \in \{1,\ldots,K\}$, Transmitter k is cognizant of the messages $M_{k-t_\ell},\ldots,M_k,\ldots,M_{k+t_r}$, where $M_{-t_\ell+1},\ldots,M_0$ and $M_{K+1},\ldots,M_{K+t_r}$ are defined to be deterministically one. Thus, Transmitter k can produce its sequence of channel inputs X_k^n as

$$X_k^n = f_k^{(n)}(M_{k-t_\ell},\ldots,M_k,\ldots,M_{k+t_r}), \tag{2}$$

for some encoding function

$$f_k^{(n)}\colon \mathcal{M}_{k-t_\ell}\times\cdots\times\mathcal{M}_k\times\cdots\times\mathcal{M}_{k+t_r} \to \mathbb{R}^n. \tag{3}$$

The transmitters are assumed to have equal average power at their disposal. Denoting by P the maximal average power with which each of the transmitters can communicate, we thus require that, with probability 1,

$$\frac{1}{n}\sum_{t=1}^{n} X_{k,t}^2 \leq P, \qquad k \in \{1,\ldots,K\}. \tag{4}$$

Each receiver observes the signals received by its own antenna, the symbols received by the $r_\ell \geq 0$ receivers to its left, and the symbols received by the $r_r \geq 0$ receivers to its right. Receiver k, for $k \in \{1,\ldots,K\}$, can thus produce its guess of Message M_k based on the $t_\ell + t_r + 1$ output sequences $Y_{k-r_\ell}^n,\ldots,Y_{k+r_r}^n$, i.e., as

$$\hat{M}_k \triangleq \varphi_k^{(n)}(Y_{k-r_\ell}^n,\ldots,Y_{k+r_r}^n), \tag{5}$$

for some decoding function

$$\varphi_k^{(n)}\colon \mathbb{R}^{n(r_\ell+r_r+1)} \to \mathcal{M}_k, \tag{6}$$

where $Y_{-r_\ell+1}^n,\ldots,Y_0^n$ and $Y_{K+1}^n,\ldots,Y_{K+r_r}^n$ are defined to be deterministically 0.

The parameters $t_\ell, t_r, r_\ell, r_r \geq 0$ are given positive integers. We call t_ℓ and t_r the *transmitter side-information parameters* and r_ℓ and r_r the *receiver side-information parameters*. Similarly, we call t_ℓ and r_ℓ the *left side-information parameters* and t_r and r_r the *right side-information parameters*.

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

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

The closure of the set of all rate-tuples $(R_1,\ldots,R_K)$ that are achievable is called the capacity region, which we denote by $\mathcal{C}^{\text{Asym}}$. To make the dependence on the number of transmitter/receiver pairs K, the side-information parameters t_ℓ, t_r, r_ℓ, r_r, and the power P explicit, we mostly write $\mathcal{C}^{\text{Asym}}(K, t_\ell, t_r, r_\ell, r_r; P)$. The sum-capacity is defined as the supremum of the sum-rate $\sum_{k=1}^K R_k$ over all achievable tuples $(R_1,\ldots,R_K)$ and is denoted by $\mathcal{C}_\Sigma^{\text{Asym}}(K, t_\ell, t_r, r_\ell, r_r; P)$. Our main focus in this work is on the high-SNR asymptote of the sum-capacity which is characterized by the *multiplexing gain*[2]:

$$\mathcal{S}^{\text{Asym}}(K, t_\ell, t_r, r_\ell, r_r) \triangleq \varlimsup_{P\to\infty} \frac{\mathcal{C}_\Sigma^{\text{Asym}}(K, t_\ell, t_r, r_\ell, r_r; P)}{\frac{1}{2}\log(P)},$$

and for large networks ($K \gg 1$) by the *asymptotic multiplexing gain per user*:

$$\mathcal{S}_\infty^{\text{Asym}}(t_\ell, t_r, r_\ell, r_r) \triangleq \varlimsup_{K\to\infty} \frac{\mathcal{S}^{\text{Asym}}(K, t_\ell, t_r, r_\ell, r_r)}{K}.$$

B. Results

Theorem 1: Irrespective of the nonzero cross-gains $\{\alpha_k\}$ *and for any* $t_\ell, t_r, r_\ell, r_r \geq 0$, *the multiplexing gain of the asymmetric model is*

$$\mathcal{S}^{\text{Asym}}(K, t_\ell, t_r, r_\ell, r_r) = K - \left\lceil \frac{K - t_\ell - r_\ell - 1}{t_\ell + t_r + r_\ell + r_r + 2} \right\rceil. \tag{7}$$

Proof: See Section V-A for the direct part and Section V-B for the converse. ■

Specializing Theorem 1 to the case $r_\ell = r_r = 0$ where each receiver has access only to its own receive antenna, recovers the result in [3].

Remark 1: Expression (7) *depends only on the sum of the left side-information parameters* $t_\ell + r_\ell$ *and on the sum of the right side-information parameters* $t_r + r_r$. *This shows an equivalence between cognition of messages at the transmitters and clustered local decoding at the receivers.*

However, the left side-information parameters (r_ℓ, t_ℓ) *do not play the same role as the right side-information parameters* (r_r, t_r). *In fact, left side-information can be more*

[2]The multiplexing gain is also referred to as the "high-SNR slope", "pre-log", or "degrees of freedom".

valuable (in terms of increasing the multiplexing gain) than right side-information.

The difference in the roles of left and right side-information is only a boundary effect: it vanishes when K tends to infinity (see Corollary 2 and Remark 2 ahead).

As a corollary to Theorem 1 we can derive the *asymptotic multiplexing gain per user*.

Corollary 2: The asymptotic multiplexing gain per user of the asymmetric network is

$$\mathcal{S}_{\infty}^{\text{Asym}}(t_\ell, t_r, r_\ell, r_r) = \frac{t_\ell + t_r + r_\ell + r_r + 1}{t_\ell + t_r + r_\ell + r_r + 2}. \tag{8}$$

Remark 2: The asymptotic multiplexing gain per-user in (8) *depends on the parameters t_ℓ, t_r, r_ℓ, and r_r only through their sum. Thus, in the considered setup the asymptotic multiplexing gain per-user only depends on the total amount of side-information at the transmitters and receivers and not on how the side-information is distributed. In particular, cognition of messages at the transmitters and clustered local decoding at the receivers are equally valuable, and—despite the asymmetry of the interference network—also left and right side-information are equally valuable.*

El Gamal, Annapureddy, and Veeravalli [16] showed that when $r_\ell = r_r = 0$ and when for each message one can freely choose the set of $t_\ell + t_r + 1$ transmitters to which this message is assigned, then the asymptotic multiplexing gain per-user is equal to $\frac{2(t_\ell+t_r+1)}{2(t_\ell+t_r+1)+1}$ and thus larger than $\mathcal{S}_{\infty}^{\text{Asym}}$ in (8). They also showed that in this modified setup, each message M_k should again be assigned to $t_\ell + t_r + 1$ adjacent transmitters, but these transmitters do not necessarily include Transmitter k.

III. Symmetric Network

A. Description of the Problem

The symmetric network is defined in the same way as the asymmetric network in Section II, except that the channel law (1) is replaced by

$$Y_{k,t} = \alpha_{k,\ell} X_{k-1,t} + X_{k,t} + \alpha_{k,r} X_{k+1,t} + N_{k,t}, \quad k \in \{1, \ldots, K\}. \tag{9}$$

Like for the asymmetric network, for each $k \in \{1, \ldots, K\}$ the symbol $X_{k,t}$ denotes Transmitter k's channel input at time t; the symbols $X_{0,t}$ and $X_{K+1,t}$ are deterministically zero; the cross-gains $\{\alpha_{k,\ell}, \alpha_{k,r}\}$ are given non-zero real numbers; and $\{N_{k,t}\}$ are i.i.d. standard Gaussians. Let H_{Net} denote the K-by-K channel matrix of the entire network: its row-j, column-i element equals 1 if $j = i$, it equals $\alpha_{j,\ell}$ if $j - i = 1$, it equals $\alpha_{j,r}$ if $j - i = -1$, and it equals 0 otherwise.

The message cognition at the transmitters is again described by the nonnegative integers t_ℓ and t_r and the encoding rules in (2), and the clustered decoding by the nonnegative integers r_ℓ and r_r and the decoding rules in (5).

The channel input sequences have to satisfy the power constraints (4).

Achievable rates, channel capacity, sum-capacity, multiplexing gain, and the asymptotic multiplexing gain per user are defined analogously to Section II. For this symmetric model and for a given positive integer $K > 0$, nonnegative integers $t_\ell, t_r, r_\ell, r_r \geq 0$, and power $P > 0$ the capacity region is denoted by $\mathcal{C}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r; P)$, the sum-capacity by $\mathcal{C}_{\Sigma}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r; P)$, the multiplexing gain by $\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r)$, and the asymptotic multiplexing gain per user by $\mathcal{S}_{\infty}^{\text{Sym}}(t_\ell, t_r, r_\ell, r_r)$.

B. Results

In the following we only state our results for the special case of equal cross-gains

$$\alpha_{k,\ell} = \alpha_{k,r} = \alpha, \qquad k \in \{1, \ldots, K\}, \tag{10}$$

that are nonzero

$$\alpha \neq 0. \tag{11}$$

Our proof techniques extend also to non-equal cross gains: an inspection of the proofs reveals that they depend on the cross-gains only through the ranks of various principal submatrices of the network's channel matrix and the fact that the cross-gains are nonzero, see the discussions in Subsections VI-G and VII-E. A formulation of our results for general cross-gains would thus involve conditions on the rank of various principal submatrices of the network's channel matrix and be very cumbersome. We therefore omit it. Instead, we notice that when all cross-gains are drawn independently according to a continuous distribution, then the various principal submatrices of the network's channel matrix are of full rank and all cross-gains are nonzero with probability 1. This makes that our main results continue to hold (with probability 1) for this randomized setup, see Remarks 3 and 6 ahead.[3]

Definition 1: For every positive integer p and real number α we define $\mathsf{H}_p(\alpha)$ to be the $p \times p$ matrix with diagonal elements all equal to 1, elements above and below the diagonal equal to α, and all other elements equal to 0.

Notice that under the assumption of all equal cross-gains α, the network's channel matrix H_{Net} satisfies

$$\mathsf{H}_{\text{Net}} = \mathsf{H}_K(\alpha). \tag{12}$$

We first present our results for *symmetric side-information* where

$$t_\ell + r_\ell = t_r + r_r, \tag{13}$$

followed by our results for general side-information parameters $r_\ell, t_\ell, r_r, t_r \geq 0$. We treat the special case with symmetric side-information separately, because for this case we have stronger results than for general side-information.

1) Symmetric Side-Information: Throughout this subsection we assume that the parameters t_ℓ, t_r, r_ℓ, r_r satisfy (13).

Theorem 3 (Symmetric Side-Information): Depending on the value of α and the parameters $K, t_\ell, t_r, r_\ell, r_r$, the multiplexing gain satisfies the following conditions.

1) *If $K \leq t_\ell + r_\ell + 1$:*

$$\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) = K - \delta_1, \tag{14}$$

where δ_1 equals 1, if $\det(\mathsf{H}_K(\alpha)) = 0$ and 0 otherwise.

[3] Such cross-gains are typically called *generic* [36], [37]. Here, we refrain from calling them so as to avoid confusion with generic subnets which we introduce in our achievability proofs.

2) *If* $K > t_\ell + r_\ell + 2$ *and* $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$*:*

$$K - \left\lfloor \frac{K}{t_\ell + r_\ell + 2} \right\rfloor - 1 \leq \mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \leq K - \left\lfloor \frac{K}{t_\ell + r_\ell + 2} \right\rfloor. \quad (15)$$

3) *If* $K > t_\ell + r_\ell + 2$; $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$; *and* $\det\left(\mathsf{H}_{t_\ell+r_\ell}(\alpha)\right) \neq 0$, *then*

$$\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) = K - \left\lfloor \frac{K}{t_\ell + r_\ell + 2} \right\rfloor. \quad (16)$$

(This third case is a special case of the second case. It is interesting because almost all values of α *lead to this case and because for this case we can improve the lower bound in* (15) *to meet the upper bound.)*

4) *If* $K > t_\ell + r_\ell + 2$ *and* $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) = 0$*:*

$$K - \left\lfloor \frac{K}{t_\ell + r_\ell + 1} \right\rfloor \leq \mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \leq K - 2\left\lfloor \frac{K}{2(t_\ell + r_\ell) + 3} \right\rfloor - \delta_2, \quad (17)$$

where δ_2 *equals 1 if* $(K \mod (2(t_\ell + r_\ell) + 3)) > (t_\ell + r_\ell + 1)$ *and 0 otherwise.*

Proof: The achievability results are proved in Section VI. The converse in (14) can be proved by first allowing all the transmitters to cooperate and all the receivers to cooperate, and then using the well-known expression for the capacity of the multi-antenna Gaussian point-to-point channel. The converse to (15) and (16) follows by specializing Upper bound (23) in Proposition 7 ahead to $t_\ell + r_\ell = t_r + r_r$. Similarly, the converse to (17) follows by specializing (24) ahead to $t_\ell + r_\ell = t_r + r_r$. ■

Remark 3: The proofs of the achievability and converse results in (14) *with* $\delta_1 = 0$ *and in* (16) *continue to hold for arbitrary cross-gains provided they are nonzero and that various principal submatrices of the network's channel matrix* H_{Net} *have full rank, see also the discussion in Subsection VI-G.*

When the cross-gains $\{\alpha_{k,\ell}\}$ *and* $\{\alpha_{k,r}\}$ *are drawn at random according to a continuous distribution both these properties hold with probability* 1*, and thus for this random setup with symmetric side-information parameters and arbitrary value* $K \geq 1$ *the multiplexing gain is*

$$K - \left\lfloor \frac{K}{t_\ell + r_\ell + 2} \right\rfloor. \quad (18)$$

Remark 4: We observe that when $\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)$ *and* $\mathsf{H}_{t_\ell+r_\ell}(\alpha)$ *are full rank, the multiplexing gain only depends on the sum of the side-information parameters* $(t_\ell + r_\ell)$*. Or equivalently they only depend on the sums* $(t_r + r_r)$ *or* $(t_\ell + t_r + r_\ell + r_r)$*. Thus, in these cases, message cognition at the transmitters and clustered local decoding at the receivers are equivalent in terms of increasing the multiplexing gain.*

The following corollary is obtained from Theorem 3 by letting K tend to ∞.

Corollary 4: If $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$*, then the asymptotic multiplexing gain per-user is given by*

$$\mathcal{S}_\infty^{\text{Sym}}(t_\ell, t_r, r_\ell, r_r) = \frac{t_\ell + r_\ell + 1}{t_\ell + r_\ell + 2}. \quad (19)$$

Otherwise, it satisfies

$$\frac{t_\ell + r_\ell}{t_\ell + r_\ell + 1} \leq \mathcal{S}_\infty^{\text{Sym}}(t_\ell, t_r, r_\ell, r_r) \leq \frac{t_\ell + r_\ell + \frac{1}{2}}{t_\ell + r_\ell + \frac{3}{2}}.$$

Thus, for a few values $\alpha \neq 0$ *the asymptotic multiplexing gain per-user drops.*

Remark 5: When $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$*, then to obtain the same asymptotic multiplexing-gain per-user in this symmetric network as in the asymmetric network before, we need double the "amount" of side-information* $t_\ell + t_r + r_\ell + r_r$*.*

El Gamal et al. [16] showed that also here a larger asymptotic multiplexing gain per-user is achievable when the messages are assigned to the transmitters in a different way (even when $r_\ell = r_r = 0$). In particular, if each message can be freely assigned to $t_\ell + t_r + 1$ transmitters, then an asymptotic multiplexing gain per-user of $\frac{2(t_\ell+t_r+1)}{2(t_\ell+t_r+1)+2}$ is achievable [16], which is larger than $\mathcal{S}_\infty^{\text{Sym}}$ in (19).

Example 1: Consider a symmetric network with symmetric side-information $r_\ell + t_\ell = r_r + t_r = 2$*. Let* K *be* 7*. Then, if* $\alpha \notin \{-\sqrt{2}/2, \sqrt{2}/2\}$*, by Theorem 3 the multiplexing gain is* 6*, and in contrast, if* $\alpha \in \{-\sqrt{2}/2, \sqrt{2}/2\}$ *the multiplexing gain is only* 5*.*

By Corollary 4 the asymptotic multiplexing gain per-user is $3/4$*, if* $\alpha \notin \{-\sqrt{2}/2, \sqrt{2}/2\}$*, but it is at most* $5/7$ *(which is smaller than* $3/4$*) if* $\alpha \in \{-\sqrt{2}/2, \sqrt{2}/2\}$*.*

Notice however, that even though the multiplexing gain is discontinuous at certain values of α, this does not imply that for fixed powers P also the sum-rate capacity of the network is discontinuous in α. Also, for given K the set of αs where the multiplexing gain is discontinuous is finite. This is in contrast to the fully-connected K-user interference channel where the multiplexing gain is discontinuous at all rational cross-gains.

We conclude this section with a result on the *high-SNR power-offset* which is defined as

$$\mathcal{L}_\infty^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \triangleq \varlimsup_{P\to\infty} \left(\frac{\mathcal{S}^{\text{Sym}}}{2} \log(P) - \mathcal{C}_\Sigma^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r; P) \right).$$

Proposition 5 (Symmetric Side-Information): Assume (13). Let α^* *be such that* $\det\left(\mathsf{H}_{r_\ell+t_\ell+1}(\alpha^*)\right) = 0$*. Also, let* $K = q(r_\ell + t_\ell + 2) - 1$ *for some positive integer* q*. Then, there exists a function* $c_0(\cdot)$*, bounded in the neighborhood of* α^* *such that for all* α *sufficiently close to* α^**:*

$$\mathcal{L}_\infty^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \geq -\nu \log|\alpha - \alpha^*| + c_0(\alpha^*),$$

where ν *is the multiplicity of* α^* *as a root of the polynomial* $\det\left(\mathsf{H}_{r_\ell+t_\ell+1}(X)\right)$*.*

In other words, when α approaches the critical value α^*, the power offset goes to infinity.

Proof: See Appendix D. ■

2) Results for General Parameters $t_\ell, t_r, r_\ell, r_r \geq 0$:

Proposition 6: The multiplexing gain of the symmetric network satisfies the following four lower bounds.

1) *It is lower bounded by:*

$$\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \geq K - 2\left\lfloor \frac{K}{t_\ell + t_r + r_\ell + r_r} \right\rfloor - \theta_1, \quad (20)$$

where

$$\theta_1 = \begin{cases} 0 & \text{if } \kappa_1 = 0 \\ 1 & \text{if } \kappa_1 = 1 \\ 2 & \text{if } \kappa_1 \geq 2 \end{cases}$$

for

$$\kappa_1 \triangleq (K \mod (t_\ell + t_r + r_\ell + r_r)).$$

2) *Moreover, irrespective of the right side-information t_r and r_r:*

$$\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \geq K - 2\left\lfloor \frac{K}{t_\ell + r_\ell + 1} \right\rfloor - \theta_2, \quad (21)$$

where

$$\theta_2 = \begin{cases} 0 & \text{if } \kappa_2 = 0 \\ 1 & \text{if } \kappa_2 = 1 \\ 2 & \text{if } \kappa_2 \geq 2 \end{cases}$$

for

$$\kappa_2 \triangleq (K \mod (t_\ell + r_\ell)).$$

3) *The lower bound* (21) *in 2) remains valid if on the right-hand side of* (21) *we replace the parameters t_ℓ and r_ℓ by t_r and r_r.*
4) *Finally, irrespective of the transmitter side-information t_ℓ and t_r, if the matrix $\mathsf{H}_{t_\ell + r_\ell + 1}(\alpha)$ is full rank:*

$$\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \geq K - 2\left\lfloor \frac{K}{r_\ell + r_t + 3} \right\rfloor - \theta_3, \quad (22)$$

where

$$\theta_3 = \begin{cases} 0 & \text{if } \kappa_3 = 0 \\ 1 & \text{if } \kappa_3 = 1 \\ 2 & \text{if } \kappa_3 \geq 2 \end{cases}$$

for

$$\kappa_3 \triangleq (K \mod (r_\ell + r_r + 3)).$$

Proof: See Section VII. ■

The lower bound in 2) is useful only when $t_r = r_r = 0$, the lower bound in 3) only when $t_\ell = r_\ell = 0$, and the bound in 4) only when $t_\ell + t_r \leq 2$.

Proposition 7: The multiplexing gain is upper bounded by the following three upper bounds.

1) *It is upper bounded by:*

$$\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \leq K - 2\left\lfloor \frac{K}{t_\ell + t_r + r_\ell + r_r + 4} \right\rfloor - \theta_4, \quad (23)$$

where

$$\theta_4 = \begin{cases} 0 & \text{if } \kappa_4 < \min\{t_\ell + r_\ell + 2, t_r + r_r + 2\} \\ 1 & \text{if } \kappa_4 \geq \min\{t_\ell + r_\ell + 2, t_r + r_r + 2\} \end{cases}$$

for

$$\kappa_4 \triangleq (K \mod (t_\ell + t_r + r_\ell + r_r + 4)).$$

2) *Moreover, if* $\det\left(\mathsf{H}_{r_\ell + t_\ell + 1}(\alpha)\right) = 0$*:*

$$\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r) \leq K - 2\left\lfloor \frac{K}{t_\ell + t_r + r_\ell + r_r + 3} \right\rfloor - \theta_5, \quad (24)$$

where

$$\theta_5 = \begin{cases} 0 & \text{if } \kappa_5 < t_r + r_r + 1 \\ 1 & \text{if } \kappa_5 \geq t_r + r_r + 1 \end{cases}$$

for

$$\kappa_5 \triangleq (K \mod (t_\ell + t_r + r_\ell + r_r + 3)).$$

3) *The upper bound in 2) holds also if everywhere (except for $\mathcal{S}^{\text{Sym}}(K, t_\ell, t_r, r_\ell, r_r)$) one exchanges the subscripts ℓ and r.*

Proof: See Section VIII. ■

From Propositions 6 and 7 we obtain the following corollary.

Corollary 8: Irrespective of the parameter α, the asymptotic multiplexing gain per user satisfies

$$\max\left\{\frac{r_\ell + r_r + 1}{r_\ell + r_r + 3}, \frac{t_\ell + t_r + r_\ell + r_r - 2}{t_\ell + t_r + r_\ell + r_r}\right\} \leq \mathcal{S}^{\text{Sym}}_\infty(t_\ell, t_r, r_\ell, r_r) \leq \frac{t_\ell + t_r + r_\ell + r_r + 2}{t_\ell + t_r + r_\ell + r_r + 4}.$$

Remark 6: The proofs of the achievability results in 1)–3) in Proposition 6 and the proofs of the converse result 1) in Proposition 7 continue to hold for arbitrary cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$ provided they are nonzero and that various principal submatrices of the network's channel matrix H_{Net} have full rank, see also the discussion in Subsection VII-E.

When the cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$ are drawn at random according to a continuous distribution both these properties hold with probability 1*, and thus results 1)–3) in Proposition 6 and result 1) in Proposition 7 also hold with probability 1. As a consequence, analogous to Corollary 8, the asymptotic multiplexing gain per-user in this randomized setup satisfies*

$$\frac{t_\ell + t_r + r_\ell + r_r - 2}{t_\ell + t_r + r_\ell + r_r} \leq \mathcal{S}^{\text{Sym}}_\infty(t_\ell, t_r, r_\ell, r_r) \leq \frac{t_\ell + t_r + r_\ell + r_r + 2}{t_\ell + t_r + r_\ell + r_r + 4}.$$

IV. Converse Proofs

For a given set of receivers $\mathcal{S} \subseteq \mathcal{K}$, let $\mathcal{R}_\mathcal{S}$ denote the set of indices $k \in \mathcal{K}$ such that Antenna k is observed by at least one of the receivers in $\mathcal{S}$.

Our converse proofs all rely on the following lemma.

Lemma 9 (Dynamic-MAC Lemma): Consider a general interference network with message cognition and clustered decoding. Let $\boldsymbol{V}_0, \ldots, \boldsymbol{V}_g$, for $g \in \mathbb{N}_0$, be a set of genie-signals and let $\mathcal{A}, \mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_q$, $q \in \mathbb{N}$, form a partition of the set of receivers $\mathcal{K}$, such that the differential entropy

$$h\big(\{\boldsymbol{N}_k\}_{k \in \mathcal{R}_\mathcal{A}} | \boldsymbol{V}_0, \ldots, \boldsymbol{V}_g\big) \tag{25}$$

is finite and bounded in P.[4] *If for any given encoding and decoding functions $f_1^{(n)}, \ldots, f_K^{(n)}$ and $\varphi_1^{(n)}, \ldots, \varphi_K^{(n)}$ there exist deterministic functions $\xi_1, \ldots, \xi_q$ on the respective domains such that for each $i \in \{1, \ldots, q\}$:*

$$\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_{\mathcal{B}_i}} = \xi_i\big(\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_{\mathcal{A}_i}}, \{M_k\}_{k \in \mathcal{A}_i}, \boldsymbol{V}_0, \ldots, \boldsymbol{V}_g\big), \tag{26}$$

where

$$\mathcal{A}_i \triangleq \mathcal{A} \cup \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_{i-1}, \tag{27}$$

then the multiplexing gain of the network is upper bounded as

$$\mathcal{S} \leq |\mathcal{R}_\mathcal{A}|. \tag{28}$$

Proof: To prove our desired upper bound we introduce a *Cognitive MAC*, whose capacity region $\mathcal{C}_{\text{MAC}}$ includes the capacity region of the original network,

$$\mathcal{C} \subseteq \mathcal{C}_{\text{MAC}}, \tag{29}$$

and whose multiplexing gain $\mathcal{S}_{\text{MAC}}$ is upper bounded as

$$\mathcal{S}_{\text{MAC}} \leq |\mathcal{R}_\mathcal{A}|. \tag{30}$$

Combining (29) and (30) establishes the desired lemma.

The Cognitive MAC is obtained from the original network by revealing the genie-information $\boldsymbol{V}_0, \ldots, \boldsymbol{V}_g$ to the receivers in Group $\mathcal{A}$ and by requiring that all the receivers that are in Group $\mathcal{A}$ jointly decode *all messages* $M_1, \ldots, M_K$, whereas all other receivers do not have to decode anything. Since the only remaining receivers in Group $\mathcal{A}$ can all cooperate in their decoding, the Cognitive MAC is indeed a MAC with only one receiver.

We now prove Inclusion (29) using a dynamic version of Sato's MAC-bound idea [38]. Specifically, we show that every coding scheme for the original network can be modified to a coding scheme for the Cognitive MAC such that whenever the original scheme is successful (i.e, all messages are decoded correctly), then so is the modified scheme. Fix a coding scheme for the original network. The transmitters of the Cognitive MAC apply the same encodings as in the original scheme. The only receiver of the Cognitive MAC, i.e., the Group $\mathcal{A}$ receiver, performs the decoding in $q+1$ rounds $0, \ldots, q$. In round $i = 0$, it decodes the messages $\{M_k\}_{k \in \mathcal{A}}$ in the same way as in the given original scheme. In rounds $i = 1, \ldots, q$,

- it attempts to reconstruct the channel outputs $\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_{\mathcal{B}_i}}$ observed by the receivers in Group $\mathcal{B}_i$ using the previously decoded messages $\{M_k\}_{k \in \mathcal{A}_i}$, the observed or previously reconstructed channel outputs $\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_{\mathcal{A}_i}}$, and the genie-information $\boldsymbol{V}_0, \ldots, \boldsymbol{V}_g$; then
- it decodes the messages $\{M_k\}_{k \in \mathcal{B}_i}$ based on its reconstructions of the outputs $\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_{\mathcal{B}_i}}$ in the same way as the receivers in Group $\mathcal{B}_i$ did in the original scheme.

By Assumption (26), the round-i reconstruction step is successful if all previous rounds' $0, \ldots, i-1$ reconstruction and decoding steps were successful. Thus, the additional reconstruction steps in the Cognitive MAC decoding do not introduce additional error events compared to the original decoding procedure, and Inclusion (29) follows.

We are left with showing that the multiplexing gain of the Cognitive MAC is upper bounded by $|\mathcal{R}_\mathcal{A}|$. Since the Group A receiver is required to decode all K messages $M_1, \ldots, M_K$, by Fano's inequality, reliable communication is possible only if

$$\begin{aligned} n \sum_{k=1}^{K} R_k &\leq I\big(\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_\mathcal{A}}, \boldsymbol{V}_0, \ldots, \boldsymbol{V}_g; M_1, \ldots, M_K\big) \\ &= I\big(\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_\mathcal{A}}; M_1, \ldots, M_K | \boldsymbol{V}_0, \ldots, \boldsymbol{V}_g\big) \\ &\leq h\big(\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_\mathcal{A}}\big) - h\big(\{\boldsymbol{N}_k\}_{k \in \mathcal{R}_\mathcal{A}} | \boldsymbol{V}_0, \ldots, \boldsymbol{V}_g\big). \end{aligned} \tag{31}$$

The multiplexing gain of $h\big(\{\boldsymbol{Y}_k\}_{k \in \mathcal{R}_\mathcal{A}}\big)$ is bounded by $|\mathcal{R}_\mathcal{A}|$. Moreover, by assumption, $h\big(\{\boldsymbol{N}_k\}_{k \in \mathcal{R}_\mathcal{A}} | \boldsymbol{V}_0, \ldots, \boldsymbol{V}_g\big)$ is finite and bounded in P. We therefore obtain from (31)

$$\overline{\lim_{P \to \infty}} \frac{\sum_{k=1}^{K} R_k}{\frac{1}{2}\log(P)} \leq |\mathcal{R}_\mathcal{A}|, \tag{32}$$

which gives the desired bound (30). ■

V. Proof of Theorem 1

Define

$$\gamma \triangleq \left\lceil \frac{K - t_\ell - r_\ell - 1}{t_\ell + t_r + r_\ell + r_r + 2} \right\rceil \tag{33}$$

$$\beta \triangleq t_\ell + t_r + r_\ell + r_r + 2, \tag{34}$$

$$\kappa \triangleq (K \mod \beta). \tag{35}$$

A. Achievability Proof of Theorem 1

We derive a lower bound by giving an appropriate coding scheme. The idea is to silence some of the transmitters, which decomposes our asymmetric network into several subnets (subnetworks), and to apply a scheme based on Costa's dirty-paper coding[5] and on successive interference cancellation in each of the subnets.

[4] For the lemma to hold, it suffices that the differential entropies grow slower than any multiple of $n\log(P)$.

[5] Alternatively, also the simpler *partial interference cancellation* scheme in [17], which is based on linear beam-forming, could be used instead of the dirty-paper coding.

1) Splitting the Network Into Subnets: We silence transmitters $j\beta$, for $j \in \{1, \ldots, \lfloor K/\beta \rfloor\}$; moreover, if $\kappa > (t_\ell + r_\ell + 1)$ we also silence Transmitter K. This splits the network into $\lceil K/\beta \rceil$ non-interfering subnets. The first $\lfloor K/\beta \rfloor$ subnets all have the same topology; they consist of $(t_\ell + t_r + r_\ell + r_r + 1)$ active transmit antennas and $(t_\ell + t_r + r_\ell + r_r + 2)$ receive antennas. We refer to these subnets as *generic* subnets. If K is not a multiple of β, there is an additional last subnet with

$$\begin{cases} \kappa \text{ active transmit antennas,} & \text{if } \kappa \leq (t_\ell + r_\ell + 1), \\ \kappa - 1 \text{ active transmit antennas,} & \text{if } \kappa > (t_\ell + r_\ell + 1), \end{cases}$$

and with κ receive antennas. We refer to such a subnet as a *reduced* subnet.

As we shall see, in our scheme each transmitter ignores its side-information about the messages pertaining to transmitters in other subnets. Likewise, each receiver ignores its side-information about the outputs of antennas belonging to receivers in other subnets. Therefore, we can describe our scheme for each subnet separately.

The scheme employed over a subnet depends on whether the scheme is generic or reduced and on the parameter $r_r \geq 0$. We describe the different schemes in the following subsections.

2) Scheme Over a Generic Subnet When $r_r > 0$: For simplicity, we assume that the parameters $K, t_\ell, t_r, r_\ell, r_r$ are such that the first subnet is generic and describe the scheme for this first subnet.

In the special case $r_\ell = 2$, $t_\ell = 2$, $t_r = 1$, and $r_r = 1$ the scheme is illustrated in Figure 3. In general, in the first subnet, we wish to transmit Messages $M_1, \ldots, M_{r_\ell + t_\ell + t_r + r_r + 1}$. Define the sets (some of which may be empty)

$$\begin{aligned} \mathcal{G}_1 &= \{1, \ldots, r_\ell + 1\} \\ \mathcal{G}_2 &= \{r_\ell + 2, \ldots, r_\ell + t_\ell + 1\} \\ \mathcal{G}_3 &= \{r_\ell + t_\ell + 2, \ldots, r_\ell + t_\ell + t_r + 1\} \\ \mathcal{G}_4 &= \{r_\ell + t_\ell + t_r + 2, \ldots, r_\ell + t_\ell + t_r + r_r + 1\}. \end{aligned}$$

Messages $\{M_k\}_{k \in \mathcal{G}_1}$ are transmitted as follows.

- For each $k \in \mathcal{G}_1$ we construct a single-user Gaussian code $\mathcal{C}_k$ of power P, blocklength n, and rate $R_k = \frac{1}{2}\log(1 + P)$.[6] The code $\mathcal{C}_k$ is revealed to Transmitter k and to Receivers $k, \ldots, r_\ell + 1$.
- Each Transmitter $k \in \mathcal{G}_1$ ignores the side-information about other transmitters' messages and codes for a Gaussian single-user channel. That is, it picks the codeword from codebook $\mathcal{C}_k$ that corresponds to its message M_k and sends this codeword over the channel.
- Decoding is performed using *successive interference cancellation*, starting by decoding Message M_1 based on the outputs of the first antenna Y_1^n.
  Specifically, each Receiver $k \in \mathcal{G}_1$ decodes as follows. Let $\hat{X}_0^n$ be an all-zero sequence of length n. Receiver k initializes j to 1 and while $j \leq k$:
  - It computes the difference
  $$Y_j^n - \alpha_j \hat{X}_{j-1}^n, \tag{36}$$
  and decodes Message M_j based on this difference using an optimal ML-decoder. Let $\hat{M}_j$ denote the decoded message.[7]
  - It picks the codeword $x_j^n(\hat{M}_j)$ from codebook $\mathcal{C}_j$ that corresponds to the guess $\hat{M}_j$ and produces this codeword as its reconstruction of the input $\hat{X}_j^n$:
  $$\hat{X}_j^n = x_j^n(\hat{M}_j). \tag{37}$$
  - It increases the index j by 1.
- Notice that each Receiver $k \in \mathcal{G}_1$ has access to the output signals $Y_1^n, \ldots, Y_k^n$ because $k \leq r_\ell + 1$, and thus the described decoding can indeed be applied.
- For each $k \in \mathcal{G}_1$, if Message M_{k-1} was decoded correctly, i.e., $\hat{M}_{k-1} = M_{k-1}$, we have
  $$Y_k^n - \alpha_k \hat{X}_{k-1}^n = X_k^n + N_k^n. \tag{38}$$
  Thus, in this case, Message M_k is decoded based on the interference-free outputs $X_k^n + N_k^n$, and, by construction of the code $\mathcal{C}_k$, the average probability of error
  $$\Pr[\hat{M}_k = M_k] \to 0 \quad \text{as} \quad n \to \infty. \tag{39}$$

If $t_\ell \geq 1$, Messages $\{M_k\}_{k \in \mathcal{G}_2}$ are transmitted as follows.

- For each $k \in \mathcal{G}_2$, we construct a dirty-paper code $\mathcal{C}_k$ that is of power P, blocklength n, and rate $R_k = \frac{1}{2}\log(1+P)$, and that is designed for noise variance 1 and interference variance $\alpha_k^2 P$ (which is the variance of $\alpha_k X_{k-1}$). The code $\mathcal{C}_k$ is revealed to Transmitters $k, \ldots, r_\ell + t_\ell + 1$ and to Receiver k.
- Each Transmitter $k \in \mathcal{G}_2$ computes the interference term $\alpha_k X_{k-1}^n$ and uses the dirty-paper code $\mathcal{C}_k$ to encode its message M_k and mitigate this interference $\alpha_k X_{k-1}^n$. It then sends the resulting sequence over the channel.
- Each Receiver $k \in \mathcal{G}_2$ ignores all the side-information about other receivers' outputs. It decodes its desired message M_k solely based on its own outputs
  $$Y_k^n = X_k^n + \alpha_k X_{k-1}^n + N_k^n \tag{40}$$
  applying dirty-paper decoding with code $\mathcal{C}_k$.
- Transmitter $k \in \mathcal{G}_2$ can compute $\alpha_k X_{k-1}^n$ because in our scheme X_{k-1}^n depends only on messages $M_{r_\ell+1}, \ldots, M_{k-1}$, and these messages are known to Transmitter k because $(k - (r_\ell + 1)) \leq t_\ell$ for all $k \in \mathcal{G}_2$.
- By construction, the sequence X_k^n, which encodes Message M_k, can perfectly mitigate the interference $\alpha_k X_{k-1}^n$, and the average probability of error
  $$\Pr[\hat{M}_k = M_k] \to 0 \quad \text{as} \quad n \to \infty. \tag{41}$$

If $t_r \geq 1$, Messages $\{M_k\}_{k \in \mathcal{G}_3}$ are transmitted as follows.

- For each $k \in \mathcal{G}_3$, we construct a dirty-paper code $\mathcal{C}_k$ of power $\alpha_{k+1}^2 P$ (the power of $\alpha_{k+1} X_k$), blocklength n, and

[6] In order to satisfy the block-power constraint imposed on the input sequences, the power of these Gaussian codebooks should be chosen slightly smaller than P. Similarly, for the probability of error tending to 0 as $n \to \infty$ the rate R_k should be slightly smaller than $1/2\log(1 + P)$. However, these are technicalities which we ignore for readability.

[7] Notice that all receivers $k = j, \ldots, r_\ell + 1$ decode Message M_j in the same way, and thus they produce the same estimate $\hat{M}_j$.

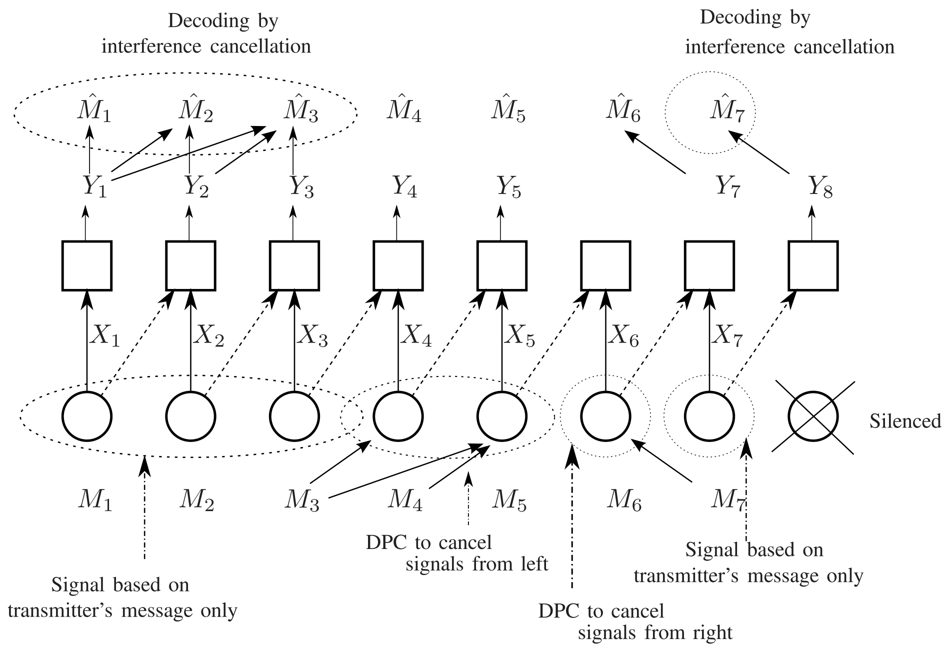

Fig. 3. Scheme in a generic subnet for parameters $t_\ell = 2$, $t_r = 1$, $r_\ell = 2$, and $r_r = 1$.

rate $R_k = \frac{1}{2}\log(1+\alpha_{k+1}^2 P)$, and that is designed for noise variance 1 and interference variance P (the variance of X_{k+1}^n). The code $\mathcal{C}_k$ is revealed to Transmitters $r_\ell + t_\ell + 2, \ldots, k$ and to Receiver k.

- Each Receiver $k \in \mathcal{G}_3$ decodes its desired message M_k based on the outputs of the antenna to its right

$$Y_{k+1}^n = X_{k+1}^n + \alpha_{k+1} X_k^n + N_{k+1}^n, \tag{42}$$

to which it has access because $r_r \geq 1$. The exact decoding procedure is explained shortly.

- Each Transmitter $k \in \mathcal{G}_3$ computes the "interference" sequence X_{k+1}^n and applies the dirty-paper code $\mathcal{C}_k$ to encode Message M_k and mitigate this "interference" X_{k+1}^n. Denoting the produced sequence by $\tilde{X}_k^n$, Transmitter k sends

$$X_k^n = \frac{1}{\alpha_{k+1}} \tilde{X}_k^n. \tag{43}$$

(The scaling by $1/\alpha_{k+1}$ in (43) reverses the amplification by α_{k+1} the sequence X_k^n experiences on its path to Receiver $(k+1)$, see (42).)

- Each Receiver $k \in \mathcal{G}_3$ applies the dirty-paper decoding of code $\mathcal{C}_k$ to the outputs

$$Y_{k+1}^n = \alpha_{k+1} X_k^n + X_{k+1}^n + N_{k+1}^n \tag{44}$$
$$= \tilde{X}_k^n + X_{k+1}^n + N_{k+1}^n. \tag{45}$$

- Notice that Transmitter $k \in \mathcal{G}_3$ can compute the "interference" X_{k+1}^n non-causally, because this latter only depends on messages $M_{k+1}, \ldots, M_{r_\ell+t_\ell+t_r+2}$ which are known to Transmitter k.
  Also, by construction of the code $\mathcal{C}_k$, the sequence $\tilde{X}_k^n$ is average block-power constrained to $\alpha_{k+1}^2 P$ and thus, by (43), the transmitted sequence X_k^n is average block-power constrained to P.
- By construction, the sequence $\tilde{X}_k^n$, which encodes Message M_k, can perfectly mitigate the "interference" X_{k+1}^n, and the average probability of error

$$\Pr[\hat{M}_k = M_k] \to 0 \quad \text{as} \quad n \to \infty. \tag{46}$$

Messages $\{M_k\}_{k \in \mathcal{G}_4}$ are transmitted as follows.

- For each $k \in \mathcal{G}_4$, we construct a single-user Gaussian codebook $\mathcal{C}_k$ of power $\alpha_{k+1}^2 P$, blocklength n, and rate $R_k = \frac{1}{2}\log(1+\alpha_{k+1}^2 P)$. The codebook $\mathcal{C}_k$ is revealed to Transmitter k and to Receivers $k, \ldots, r_\ell + t_\ell + t_r + r_r + 1$.
- Each Transmitter $k \in \mathcal{G}_4$ ignores the side-information about other transmitters' messages and codes for a Gaussian single-user channel. That is, it picks the codeword from code $\mathcal{C}_k$ that corresponds to its message M_k and sends this codeword over the channel.
- Decoding is performed using successive interference cancellation, starting by decoding Message $M_{r_\ell+t_\ell+t_r+r_r+1}$ based on the outputs of the last antenna $Y_{r_\ell+t_\ell+t_r+r_r+2}^n$.
  Specifically, Receiver $k \in \mathcal{G}_4$ decodes its desired Message M_k as follows. Let $\hat{X}_{r_\ell+t_\ell+t_r+r_r+3}^n$ be an all-zero sequence of length n.
  Receiver k initializes j to $r_\ell + t_\ell + t_r + r_r + 1$, and while $j \geq k$:
  - It computes the difference

$$Y_{j+1}^n - \hat{X}_{j+1}^n, \tag{47}$$

  and decodes Message M_j based on this difference using an optimal ML-decoder.
  Let $\hat{M}_j$ denote the resulting guess of Message M_j.
  - It reconstructs the input sequence X_j^n by picking the codeword $x_j^n(\hat{M}_j)$ from codebook $\mathcal{C}_j$ that corresponds to Message $\hat{M}_j$:

$$\hat{X}_j^n = x_j^n(\hat{M}_j). \tag{48}$$

  - It decreases j by 1.

- Notice that Receiver $k \in \mathcal{G}_4$ has access to the output signals $Y_k^n, \ldots, Y_{r_\ell+t_\ell+t_r+r_r+2}^n$ because $k \geq r_\ell+t_\ell+t_r+2$.
- For each $k \in \mathcal{G}_4$, if the previous message M_{k-1} has been decoded correctly, i.e, $\hat{M}_{k-1} = M_{k-1}$, we have

$$Y_{k+1}^n - \hat{X}_{k+1}^n = \alpha_{k+1} X_k^n + N_{k+1}^n. \tag{49}$$

Thus, in this case, Message M_k is decoded based on the interference-free outputs $\alpha_{k+1} X_k^n + N_{k+1}^n$, and, by construction of the code $\mathcal{C}_k$, the average probability of error

$$\Pr[\hat{M}_k = M_k] \to 0 \quad \text{as} \quad n \to \infty. \tag{50}$$

To summarize, in the described scheme we sent messages $M_1, \ldots, M_{r_\ell+t_\ell+t_r+r_r+1}$ with vanishingly small average probability of error, see (39), (41), (46), and (50), and at rates

$$R_1 = \cdots = R_{r_\ell+t_\ell+1} = \frac{1}{2}\log(1+P) \tag{51}$$

$$R_{r_\ell+t_\ell+2} = \cdots = R_{r_\ell+t_\ell+t_r+r_r+1} = \frac{1}{2}\log(1+\alpha_{k+1}^2 P). \tag{52}$$

Conclusion 1: Irrespective of the nonzero cross-gains $\{\alpha_k\}$ and for any $t_\ell, t_r, r_\ell \geq 0$ and $r_r > 0$, our scheme achieves a multiplexing gain of $(r_\ell + t_\ell + t_r + r_r + 1)$ over a generic subnet. It uses all $(r_\ell+t_\ell+t_r+r_r+1)$ active transmit antennas of the subnet and all $(r_\ell + t_\ell + t_r + r_r + 2)$ receive antennas.

3) Scheme Over a Generic Subnet When $r_r = 0$: We again assume that the first subnet is generic and focus on this first subnet. When $r_r = 0$ we transmit Messages $M_1, \ldots, M_{r_\ell+t_\ell+1}$ and $M_{r_\ell+t_\ell+3}, \ldots, M_{r_\ell+t_\ell+t_r+2}$ over the first subnet.

Messages $\{M_k\}_{k\in(\mathcal{G}_1\cup\mathcal{G}_2)}$ are transmitted in the same way as in the previous section V-A2. If $t_r > 0$, the set $\mathcal{G}_3$ is nonempty. In this case, Messages $\{M_{k+1}\}_{k\in\mathcal{G}_3}$ are transmitted in a similar way as Messages $\{M_k\}_{k\in\mathcal{G}_3}$ in the previous section V-A2, except that now each Transmitter $k \in \mathcal{G}_3$ sends Message M_{k+1} (as opposed to Message M_k) and accordingly, each output sequence Y_{k+1}^n is used by Receiver $k+1$ to decode Message M_{k+1} (as opposed to Receiver k decoding Message M_k based on Y_{k+1}^n). More specifically:

- For each $k \in \mathcal{G}_3$, we construct a dirty-paper code $\mathcal{C}_{k+1}$ that is of power $\alpha_{k+1}^2 P$ (the power of $\alpha_{k+1} X_k$), block-length n, and rate $R_{k+1} = \frac{1}{2}\log(1+\alpha_{k+1}^2 P)$, and that is designed for noise variance 1 and interference variance P (the variance of X_{k+1}). The code $\mathcal{C}_{k+1}$ is revealed to Transmitters $r_\ell + t_\ell + 2, \ldots, k$ and to Receiver $k+1$.
- Transmitter $k \in \mathcal{G}_3$ applies the dirty-paper code $\mathcal{C}_{k+1}$ to encode Message M_{k+1} and mitigate the "interference" X_{k+1}^n. Denoting the sequence produced by the dirty-paper code by $\tilde{X}_k^n$, Transmitter k sends

$$X_k^n = \frac{1}{\alpha_{k+1}} \tilde{X}_k^n. \tag{53}$$

- Each Receiver $k+1$, for $k \in \mathcal{G}_3$, ignores its side-information about outputs observed at other antennas. It decodes its desired Message M_{k+1} solely based on the outputs at its own antenna

$$Y_{k+1}^n = \alpha_{k+1} X_k^n + X_{k+1}^n + N_{k+1}^n \tag{54}$$

$$= \tilde{X}_k^n + X_{k+1}^n + N_{k+1}^n \tag{55}$$

using the dirty-paper decoding of code $\mathcal{C}_{k+1}$.

- Notice that Transmitter $k \in \mathcal{G}_3$ can compute the "interference" sequence X_{k+1}^n because this latter only depends on messages $M_{k+2}, \ldots, M_{r_\ell+t_\ell+t_r+2}$ which are known to Transmitter k.
- By construction, the sequence $\tilde{X}_k^n$, which encodes Message M_{k+1}, can completely mitigate the "interference" X_{k+1}^n, and the average probability of error

$$\Pr[\hat{M}_{k+1} \neq M_{k+1}] \to 0 \quad \text{as} \quad n \to \infty. \tag{56}$$

To summarize, in the described scheme we transmit Messages $M_1, \ldots, M_{r_\ell+t_\ell+1}$ and $M_{r_\ell+t_\ell+3}, \ldots, M_{r_\ell+t_\ell+t_r+2}$ with vanishingly small average probability of error, see (39), (41), and (56), and at rates

$$R_1 = \cdots = R_{r_\ell+t_\ell+1} = \frac{1}{2}\log(1+P) \tag{57}$$

$$R_{r_\ell+t_\ell+3} = \cdots = R_{r_\ell+t_\ell+t_r+2} = \frac{1}{2}\log(1+\alpha_{k+1}^2 P). \tag{58}$$

Conclusion 2: Irrespective of the nonzero cross-gains $\{\alpha_k\}$ and for any $t_\ell, t_r, r_\ell \geq 0$, our scheme for $r_r = 0$ achieves a multiplexing gain of $(r_\ell + t_\ell + t_r + 1)$ over a generic subnet. If $t_r \geq 1$, it uses all $(r_\ell+t_\ell+t_r+r_r+1)$ active transmit antennas and all $(r_\ell+t_\ell+t_r+2)$ receive antennas of the subnet. If $t_r = 0$ it uses all $(r_\ell + t_\ell + 1)$ active transmit antennas; but it only uses the first $(r_\ell + t_\ell + 1)$ receive antennas and ignores the last antenna of the subnet.

4) Scheme Over a Reduced Subnet: Let

$$r'_\ell \triangleq \min\left[(\kappa-1), r_\ell\right] \tag{59a}$$

$$t'_\ell \triangleq \min\left[(\kappa - r_\ell - 1)_+, t_\ell\right] \tag{59b}$$

$$t'_r \triangleq \min\left[(\kappa - r_\ell - t_\ell - 2)_+, t_r\right] \tag{59c}$$

$$r'_r \triangleq \min\left[(\kappa - r_\ell - t_\ell - t_r - 2)_+, r_r\right] \tag{59d}$$

where $(x)_+$ is defined as $\max\{x, 0\}$. In a reduced subnet we apply one of the two schemes described for the generic subnet but now with reduced side-information parameters $r'_\ell, t'_\ell, t'_r, r'_r$. If $r'_r > 0$, we apply the scheme in Subsection V-A2 otherwise we apply the scheme in Subsection V-A3. Notice that, by definition, $r'_\ell \leq r_\ell$, $t'_\ell \leq t_\ell$, $t_r \leq t'_r$, and $r'_r \leq r_r$, and thus the transmitters and receivers have enough side-information to apply the described schemes with these parameters.

When $\kappa \leq (t_\ell + r_\ell + 1)$, then the reduced subnet consists of an equal number κ of active transmit and receive antennas because the last transmit antenna has not been silenced. In this case, also $t'_r = r'_r = 0$ and by Conclusion 2, the scheme in Subsection V-A3 achieves multiplexing gain κ over such a subnet.

When $\kappa > (t_\ell + r_\ell + 1)$, the subnet consists of $\kappa - 1$ active transmit antennas and κ receive antennas. By Conclusions 1 and 2, one of the schemes in Subsections V-A2 or V-A3 achieves multiplexing gain $\kappa - 1$ over such a subnet.

To summarize, we achieve a multiplexing gain of

$$\begin{cases} \kappa, & \text{if } \kappa \leq t_\ell + r_\ell + 1 \\ \kappa - 1, & \text{if } \kappa > t_\ell + r_\ell + 1 \end{cases} \tag{60}$$

over a reduced subnet of size κ.

5) Performance Analysis Over the Entire Network: Over each of the first $\lfloor K/\beta \rfloor$ generic subnets we achieve a multiplexing gain of $\beta - 1$ and, if it exists, then over the last reduced subnet we achieve a multiplexing gain of either κ or $\kappa - 1$, see (60). Over the entire network we thus achieve a multiplexing gain of

$$K - \gamma = \begin{cases} K - \lfloor K/\beta \rfloor, & \text{if } \kappa \leq t_\ell + r_\ell + 1 \\ K - \lfloor K/\beta \rfloor - 1, & \text{if } \kappa > t_\ell + r_\ell + 1. \end{cases} \tag{61}$$

This proves the desired lower bound.

Remark 7: In the described scheme a subset of γ messages is completely ignored and not sent over the network. Using time-sharing we can obtain a fair scheme that sends all messages at almost equal rates and achieves a multiplexing gain of at least $K - \gamma - 1$. More specifically, the idea is to time-share β schemes where in the i-th scheme, $i \in \{1, \ldots, \beta\}$, we silence transmitters $\{i + j\beta\}_{j \in \left\{1, \ldots, \left\lfloor \frac{K-i}{\beta} \right\rfloor\right\}}$, and if $(K \mod \beta) \geq (i + t_\ell + r_\ell + 1)$, then we also silence the last transmitter K. This splits the network into γ or $\gamma + 1$ subnets: a possibly reduced first subnet, $\gamma - 2$ or $\gamma - 1$ generic subnets, and a possibly reduced last subnet. In each of the subnets, depending on whether it is generic or reduced, one of the schemes described above is used.

B. Converse to Theorem 1

Apply the Dynamic-MAC Lemma 9 to the following choices:

- $q = 1$;
- $g = \gamma - 1$;
- $\mathcal{A} \triangleq \bigcup_{m=0}^{g} \mathcal{A}(m)$, where for $m = 0, \ldots, g-1$,

$$\mathcal{A}(m) \triangleq \{m\beta + r_\ell + 2, \ldots, (m+1)\beta - r_r\}$$

  and

$$\mathcal{A}(g) \triangleq \{g\beta + r_\ell + 2, \ldots, K\}.$$

- $\mathcal{B}_1 \triangleq \mathcal{K} \setminus \mathcal{A}$;
- genie-information

$$\boldsymbol{V}_0 \triangleq \boldsymbol{N}_1 + \sum_{\nu=1}^{r_\ell+t_\ell+1} \left(\prod_{j=1}^{\nu} \frac{-1}{\alpha_{1+j}} \right) \boldsymbol{N}_{1+\nu}, \tag{62}$$

  and, for $m \in \{1, \ldots, g\}$:

$$\begin{aligned} \boldsymbol{V}_m \triangleq\ & \boldsymbol{N}_{1+m\beta} \\ & + \sum_{\nu=1}^{r_\ell+t_\ell+1} \left(\prod_{j=1}^{\nu} \frac{-1}{\alpha_{m\beta+1+j}} \right) \boldsymbol{N}_{1+m\beta+\nu} \\ & + \sum_{\nu=1}^{t_r+r_r} \left(\prod_{j=1}^{\nu} (-\alpha_{m\beta+1+j-\nu}) \right) \boldsymbol{N}_{1+m\beta-\nu} \end{aligned} \tag{63}$$

Notice that by our choice of $\mathcal{A}$, the set difference

$$\mathcal{K} \backslash \mathcal{R}_{\mathcal{A}} = \{1 + m\beta\}_{m=0}^{g}. \tag{64}$$

Since for each $m = 0, \ldots, g$ the genie-information $\boldsymbol{V}_m$ contains an additive noise term $N_{1+m\beta}$, which is not present in all other genie-informations $\{\boldsymbol{V}_{m'}\}_{m' \neq m}$, (64) and the independence of the noises imply that the differential entropy in (25) is finite. Moreover, the differential entropy does not depend on P because neither does the genie-information. In the following, we show that also the second assumption (26) of Lemma 9 is satisfied and that thus we can apply the lemma for the described choice. This then proves the desired converse because, by (64), $|\mathcal{R}_{\mathcal{A}}| = K - g - 1 = K - \gamma$.

By (64), the set $\{M_k\}_{k \in \mathcal{A}}$ includes all messages $\{M_{r_\ell+2+\nu+m\beta}\}_{\substack{0 \leq \nu \leq t_\ell + t_r \\ 0 \leq m \leq \gamma - 1}}$, where out of range indices should be ignored. From $\{M_k\}_{k \in \mathcal{A}}$ it is thus possible to reconstruct the input sequences $\{\boldsymbol{X}_{t_\ell+r_\ell+2+m\beta}\}_{m=0}^{g}$:

$$\boldsymbol{X}_{r_\ell+t_\ell+2+m\beta} = f^{(n)}_{r_\ell+t_\ell+2+m\beta}(M_{r_\ell+2+m\beta}, \ldots, M_{r_\ell+t_\ell+t_r+2+m\beta}).$$

Using these reconstructed sequences, the output sequences observed at the receivers in Group $\mathcal{A}$, and the genie-information $\{\boldsymbol{V}_m\}_{m=0}^{g}$, it is then possible to reconstruct all channel outputs not observed by the receivers in Group $\mathcal{A}$, (64):

$$\begin{aligned} \boldsymbol{Y}_1 = & - \sum_{\nu=1}^{r_\ell+t_\ell+1} \left(\prod_{j=1}^{\nu} \frac{-1}{\alpha_{m\beta+1+j}} \right) \boldsymbol{Y}_{1+\nu} \\ & + \left(\prod_{j=1}^{t_\ell+r_\ell+1} \frac{-1}{\alpha_{m\beta+1+j}} \right) \boldsymbol{X}_{r_\ell+t_\ell+2} + \boldsymbol{V}_0 \end{aligned}$$

and, for $m \in \{1, \ldots, g\}$:

$$\begin{aligned} \boldsymbol{Y}_{1+m\beta} = & - \sum_{\nu=1}^{r_\ell+t_\ell+1} \left(\prod_{j=1}^{\nu} \frac{-1}{\alpha_{m\beta+1+j}} \right) \boldsymbol{Y}_{1+m\beta+\nu} \\ & - \sum_{\nu=1}^{t_r+r_r} \left(\prod_{j=1}^{\nu} (-\alpha_{m\beta+1+j-\nu}) \right) \boldsymbol{Y}_{1+m\beta-\nu} \\ & + \left(\prod_{j=1}^{t_\ell+r_\ell+1} \frac{-1}{\alpha_{m\beta+1+j}} \right) \boldsymbol{X}_{r_\ell+t_\ell+2+m\beta} \\ & - \left(\prod_{j=0}^{r_r+t_r} (-\alpha_{m\beta+1+j-r_r-t_r}) \right) \boldsymbol{X}_{r_\ell+t_\ell+2+(m-1)\beta} \\ & + \boldsymbol{V}_m. \end{aligned}$$

This establishes that Assumption (26) holds, and concludes the proof.

VI. Achievability Proof of Theorem 3

For each of the four lower bounds 1)–4) in Theorem 3, i.e., Inequalities (14)–(17), we present a scheme achieving this lower bound. The four schemes are similar: they all rely on the idea of switching off some of the transmitter/receiver pairs, and on using the strategy in Subsection VI-A ahead over the resulting subnets. (Here, by silencing transmitter/receiver pairs we intend that we silence the antennas at the transmitters and ignore the corresponding antennas at the receivers.) This splits the networks into non-interfering subnets. In each scheme we silence a different set of transmitter/receiver pairs. As we will see we do this in a way that splits the network into subnets that have at most $t_\ell + r_\ell + 1$ active transmitter/receiver pairs.

We first describe the strategy used to communicate over the subnets (Section VI-A). Then, we present the set of transmitter/receiver pairs that needs to be silenced in each of the four schemes, so that they achieve the lower bounds in 1)–4) (Sections VI-C–VI-F).

A. Strategy Used in the Subnets

Consider a subnet with κ transmitter/receiver pairs, where $\kappa \leq t_\ell + r_\ell + 1$. Notice that this subnet's channel matrix is given by $\mathsf{H}_\kappa(\alpha)$. We first present a coding strategy that achieves full multiplexing gain of $\operatorname{rank}(\mathsf{H}_\kappa(\alpha))$ over the subnet when

$$\kappa = t_\ell + r_\ell + 1. \tag{65}$$

Then we describe how to modify this strategy to achieve full multiplexing gain of $\operatorname{rank}(\mathsf{H}_\kappa(\alpha))$ over the subnet when $\kappa < t_\ell + r_\ell + 1$.

Depending on which of the following three cases applies, we use a different scheme to communicate over the subnet.

1.) If the transmitters and the receivers have the same amount of side-information:

$$r_\ell + r_r = t_\ell + t_r \tag{66}$$

we use a Multi-Input/Multi-Output (MIMO) point-to-point scheme.

2.) If the transmitters have more side-information than the receivers:

$$r_\ell + r_r < t_\ell + t_r \tag{67}$$

we use a MIMO broadcast scheme.

3.) If the receivers have more side-information than the transmitters:

$$r_\ell + r_r > t_\ell + t_r \tag{68}$$

we use a MIMO multi-access scheme.

We first describe the MIMO point-to-point scheme for case 1.). In this case (13) and (66) imply that

$$t_\ell = r_r \quad \text{and} \quad t_r = r_\ell. \tag{69}$$

Therefore, since $\kappa = r_\ell + t_\ell + 1$, (65), all κ transmitters are cognizant of Message M_{t_r+1} and Receiver $(t_r + 1)$ has access to all κ antennas in the subnet. Thus, all the transmitters can act as a single transmitter that transmits Message M_{t_r+1} to Receiver $(t_r + 1)$ which can decode the Message based on all the antennas in the subnet. Using an optimal MIMO point-to-point scheme for this transmission achieves a multiplexing gain of $\operatorname{rank}(\mathsf{H}_\kappa(\alpha))$ over the subnet.

We next describe the MIMO broadcast scheme for case 2.). Notice that (13) and (67) imply that

$$r_\ell < t_r. \tag{70}$$

By (13) and (65), all the transmitters are cognizant of Messages $M_{r_\ell+1}, \ldots, M_{t_r+1}$ and Receivers $(r_\ell + 1), \ldots, (t_r + 1)$ *jointly* have access to all the κ antennas in the subnet. Thus, all the transmitters in the subnet can act as a big common transmitter that transmits Messages $M_{r_\ell+1}, \ldots, M_{t_r+1}$

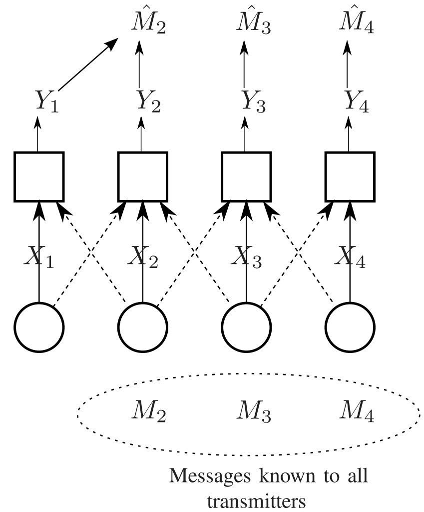

Fig. 4. Broadcast scheme employed in a subnet for parameters $\kappa = 4$, $t_\ell = 2$, $t_r = 3$, $r_\ell = 1$, and $r_r = 0$.

to the independent Receivers $(r_\ell + 1), \ldots, (t_r + 1)$. where Receiver $(r_\ell + 1)$ decodes based on antennas $1, \ldots, r_\ell + 1$ (and ignores the other antennas), Receivers $(r_\ell + 2), \ldots, t_\ell$ decode based only on their own antennas, and Receiver $(t_r + 1)$ decodes based on antennas $t_r + 1, \ldots, t_r + r_r + 1$.[8] Using an optimal MIMO broadcast scheme for this transmission we can achieve a multiplexing gain of $\operatorname{rank}(\mathsf{H}_\kappa(\alpha))$ over the subnet.

For parameters $t_\ell = 2$, $t_r = 3$, $r_\ell = 1$, and $r_r = 0$ the scheme is illustrated in Figure 4.

We finally describe the MIMO multi-access scheme for case 3.). Here, (13) and (68) imply

$$t_r < r_\ell. \tag{71}$$

By (13) and (65), each transmitter knows at least one of the Messages $M_{t_r+1}, \ldots, M_{r_\ell+1}$, and Receivers $(t_r + 1), \ldots, (r_\ell + 1)$ all have access to all κ receive antennas in the subnet. In our scheme the first $t_r + 1$ transmitters $1, \ldots, t_r + 1$ act as a big common transmitter that transmits Message M_{t_r+1}. Similarly, the last $t_\ell + 1$ transmitters $r_\ell + 1, \ldots, r_\ell + t_\ell + 1$ act as a big common transmitter that transmits Message $M_{r_\ell+1}$. Transmitters $t_r + 2, \ldots, r_\ell$ act as single transmitters that transmit their own messages. Receivers $(t_r + 1), \ldots, (r_\ell + 1)$ act as a single big common receiver that decodes Messages $M_{t_r+1}, \ldots, M_{r_\ell+1}$ based on all the antennas in the network. Applying an optimal MIMO MAC scheme for this transmission achieves multiplexing gain $\operatorname{rank}(\mathsf{H}_\kappa(\alpha))$ over the subnet.

For parameters $t_\ell = 2$, $t_r = 0$, $r_\ell = 1$, and $r_r = 3$ the scheme is illustrated in Figure 5.

We conclude that with the above described schemes we can achieve a multiplexing gain of $\operatorname{rank}(\mathsf{H}_\kappa(\alpha))$ when $\kappa = t_\ell + r_\ell + 1$, irrespective of the specific values of t_ℓ and r_ℓ.

We now consider the case where

$$\kappa < t_\ell + r_\ell + 1. \tag{72}$$

[8]The described assignment of antennas to receivers is only one possible assignment that leads to the desired multiplexing gain. Other assignments are possible.

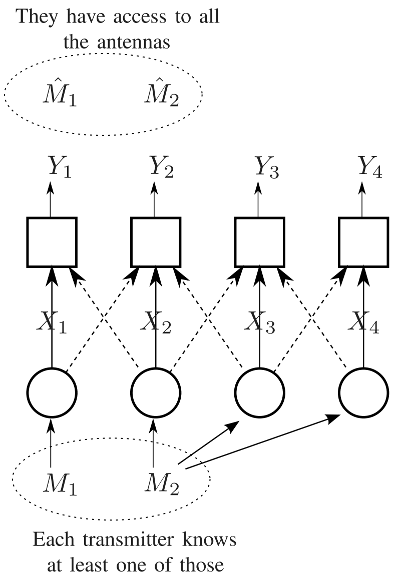

Fig. 5. Multi-access scheme employed in a subnets for parameters $\kappa = 4$, $t_\ell = 2$, $t_r = 0$, $r_\ell = 1$, and $r_r = 3$.

In this case we choose parameters $t'_\ell \leq t_\ell$, $t'_r \leq t_r$, $r'_\ell \leq r_\ell$, and $r'_r \leq r_r$ such that

$$\kappa = t'_\ell + r'_\ell + 1 = t'_r + r'_r + 1, \tag{73}$$

and depending on the choice of $t'_\ell, t'_r, r'_\ell, r'_r$ we apply one of the three schemes above. This way, we achieve multiplexing gain $\text{rank}(\mathsf{H}_\kappa(\alpha))$ over the subnet also when (72) holds.

We obtain the following proposition.

Proposition 10: For every subnet with $\kappa \leq t_\ell + r_\ell + 1$ transmitter/receiver pairs one of the three schemes described above achieves a multiplexing gain of $\text{rank}(\mathsf{H}_\kappa(\alpha))$.

This result relies on the cross-gains only through the rank of the subnet's channel matrix. Therefore, also in the setup with general cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$, over any subnet with $\kappa \leq t_\ell + r_\ell + 1$ transmitter/receiver pairs the described schemes achieve a multiplexing gain equal to the rank of the subnet's channel matrix.

B. Auxiliary Results

The following auxiliary results will be used in the proofs ahead.

Lemma 11: Let a real number α and a positive integer p be given such that $\det\left(\mathsf{H}_p(\alpha)\right) = 0$. *Then, the following statements hold.*

1) *The integer $p \geq 2$.*
2) *The determinants* $\det\left(\mathsf{H}_{p-1}(\alpha)\right)$, $\det\left(\mathsf{H}_{p+1}(\alpha)\right)$, *and* $\det\left(\mathsf{H}_{p+2}(\alpha)\right)$ *are all non-zero. Moreover, if $p > 2$ (and thus $\mathsf{H}_{p-2}(\alpha)$ is defined) also* $\det\left(\mathsf{H}_{p-2}(\alpha)\right)$ *is non-zero.*

Proof: See Appendix A. ■

Corollary 12: For every real number α and positive integer p, the rank of the matrix $\mathsf{H}_p(\alpha)$ is either p or $p-1$.

Proof: Follows by noting that $H_{p-1}(\alpha)$ is a sub-matrix of $H_p(\alpha)$ and by Lemma 11. ■

C. Achieving the Lower Bound in (14)

Recall that (14) holds under the assumption that $K \leq t_\ell + r_\ell + 1$. In this case, we do not silence any transmitter/receiver pairs but we directly apply one of the threes schemes in the previous Subsection VI-A. By Proposition 10 this way we can achieve a multiplexing gain of $\text{rank}(\mathsf{H}_K(\alpha))$, which trivially equals K if $\det(\mathsf{H}_K(\alpha)) \neq 0$ and by Corollary 12 equals $K-1$ otherwise.

D. Achieving the Lower Bound in (15)

Recall that (15) holds under the assumption that $K > (t_\ell + r_\ell + 2)$ and $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$. We define

$$\tilde{\kappa} \triangleq K \mod (t_\ell + r_\ell + 2) \tag{74}$$

$$\tilde{\gamma} \triangleq \left\lfloor \frac{K}{t_\ell + r_\ell + 2} \right\rfloor \tag{75}$$

and notice that by assumption $\tilde{\gamma} \geq 1$.

We switch off the transmitter/receiver pairs $\{g(t_\ell + r_\ell + 2)\}_{g=1}^{\tilde{\gamma}}$, i.e., in total $\tilde{\gamma}$ transmitter/receiver pairs. This decomposes the network into $\tilde{\gamma}$ subnets with $(t_\ell + r_\ell + 1)$ transmitter/receiver pairs and possibly a smaller last network with $\tilde{\kappa} \leq (t_\ell + r_\ell + 1)$ transmitter/receiver pairs. Thus, in each subnet we can apply one of the schemes described in Subsection VI-A. By Proposition 10, this achieves multiplexing gain $\text{rank}\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right)$ over the first $\tilde{\gamma}$ subnets and multiplexing gain $\text{rank}\left(\mathsf{H}_{\tilde{\kappa}}(\alpha)\right)$ over the last smaller network (if it exists). By assumption $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$ and thus $\text{rank}\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) = (t_\ell + r_\ell + 1)$; moreover, by Corollary 12, $\text{rank}\left(\mathsf{H}_{\tilde{\kappa}}(\alpha)\right)$ is either equal to $\tilde{\kappa}$ or to $\tilde{\kappa}-1$. Thus, we achieve at least the desired multiplexing gain of $K - \left\lfloor \frac{K}{t_\ell+r_\ell+2} \right\rfloor - 1$. In fact, whenever $\tilde{\kappa} = 0$ or $\det\left(\mathsf{H}_{\tilde{\kappa}}(\alpha)\right) \neq 0$, then we can even achieve a multiplexing gain of $K - \left\lfloor \frac{K}{t_\ell+r_\ell+2} \right\rfloor$.

E. Achieving the Lower Bound in (16)

Recall that (16) holds under the assumption that $K > (t_\ell + r_\ell + 2)$; that $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$; and that $\det\left(\mathsf{H}_{t_\ell+r_\ell}(\alpha)\right) \neq 0$.

We distinguish two cases depending on $\tilde{\kappa}$ as defined in (74):

1) $\text{rank}\left(\mathsf{H}_{\tilde{\kappa}}(\alpha)\right) = \tilde{\kappa}$;
2) $\text{rank}\left(\mathsf{H}_{\tilde{\kappa}}(\alpha)\right) < \tilde{\kappa}$.

In case 1) we use the same scheme as in the previous Subsection VI-D. As described above, this scheme achieves a multiplexing gain of $\text{rank}\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right)$ over each of the first $\left\lfloor \frac{K}{t_\ell+r_\ell+2} \right\rfloor$ subnets and a multiplexing gain of $\text{rank}\left(\mathsf{H}_{\tilde{\kappa}}(\alpha)\right)$ over the last smaller network. Since we assumed that $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$ and that $\text{rank}\left(\mathsf{H}_{\tilde{\kappa}}(\alpha)\right) = \tilde{\kappa}$, we conclude we achieve the desired multiplexing gain of $K - \left\lfloor \frac{K}{t_\ell+r_\ell+2} \right\rfloor$ over the entire network.

We now treat case 2). Notice that in this case $\tilde{\kappa} < t_\ell + r_\ell$ because we assumed that $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$ and that $\det\left(\mathsf{H}_{t_\ell+r_\ell}(\alpha)\right) \neq 0$.

We switch off transmitter/receiver pairs $\{g(t_\ell + r_\ell + 2)\}_{g=1}^{\tilde{\gamma}-1}$ and transmitter/receiver pair $\tilde{\gamma}(t_\ell + r_\ell + 2) - 1$, where $\tilde{\gamma}$ is defined in (75). This way, the first $\tilde{\gamma} - 1$ subnets are of size $t_\ell + r_\ell + 1$, the next subnet is of size $(t_\ell + r_\ell)$, and the last is of size $\tilde{\kappa}+1$ (where $\tilde{\kappa}$ is defined in (74)). Thus, all the subnets consist of at most $t_\ell + r_\ell + 1$ transmitter/receiver pairs, and we can apply one of the three schemes described in Subsection VI-A.

Since $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) \neq 0$, by Proposition 10, we achieve a multiplexing gain of $t_\ell + r_\ell + 1$ over each of the first $\tilde{\gamma}-1$ subnets. Moreover, since we assumed that $\det\left(\mathsf{H}_{t_\ell+r_\ell}(\alpha)\right) \neq 0$, we further achieve a multiplexing gain of $(t_\ell + r_\ell)$ over the $\tilde{\gamma}$-th subnet. Finally, since we assumed that $\det(\mathsf{H}_{\tilde{\kappa}}(\alpha)) = 0$, by Lemma 11, $\det\left(\mathsf{H}_{\tilde{\kappa}+1}(\alpha)\right) \neq 0$, and thus we achieve a multiplexing gain of $\tilde{\kappa}+1$ over the last subnet. We conclude that our scheme achieves full multiplexing gain (i.e., multiplexing gain equal to the number of transmitter/receiver pairs) in each subnet and hence a multiplexing gain of $K - \left\lfloor \frac{K}{t_\ell+r_\ell+2} \right\rfloor$ over the entire network.

F. Achieving the Lower Bound in (17)

Recall that (17) holds under the assumptions that $K > t_\ell + r_\ell + 2$ and $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}\right) = 0$.

We switch off every $(t_\ell + r_\ell + 1)$-th transmitter/receiver pair, i.e., in total $\left\lfloor \frac{K}{t_\ell+r_\ell+1} \right\rfloor$ transmitter/receiver pairs, and, depending on the values of t_ℓ, t_r, r_ℓ, r_r, we apply one of the three schemes in Subsection VI-A over the resulting subnets. Following similar lines as in the previous proof, it can be shown that all the resulting subnets have full-rank channel matrices and thus by Proposition 10 a multiplexing gain of $K - \left\lfloor \frac{K}{t_\ell+r_\ell+1} \right\rfloor$ is achieved over the entire network. The details of the proof are omitted.

G. General Cross-Gains $\{\alpha_{k,\ell}\}$ *and* $\{\alpha_{k,r}\}$

The performance analysis of the schemes presented in the previous subsections rely on the cross-gains only through the ranks of various principal submatrices of the network's channel matrix H_K and on the fact that the cross-gains are nonzero. Thus, our proofs and results generalize to non-equal cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$.

More specifically, the three MIMO coding strategies discussed in Subsection VI-A achieve multiplexing gains equal to the rank of the subnet's channel matrix if the subnet consists of no more than $t_\ell + r_\ell + 1$ active transmitters and receivers, irrespective of the actual values of the cross-gains and of whether they are all equal or different. Thus, for general cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$, if we silence pairs of consecutive transmitters, ignore the corresponding receivers' antennas, and apply the appropriate MIMO strategies over the resulting subnets, then we achieve a multiplexing gain over the entire network that is equal to the sum of the ranks of the subnets' channel matrices.

The best choice of the pairs of transmitters to silence depends on the values of the cross-gains. In the case of equal cross-gains α, Lemma 11 and Corollary 12 helped us determining the best choices. Finding the optimal choices for general cross-gains seems very involved. Lemma 11 and Corollary 12 however generalize, and can provide some help. Lemma 11, for example, generalizes to arbitrary nonzero cross-gains in the following way. For each positive integer $p \leq K$, let $\mathsf{H}_{\text{Net},p}$ denote the p-th principal submatrix of H_{Net}. Then, Lemma 11 remains valid if the matrices $\mathsf{H}_q(\alpha)$ are replaced by $\mathsf{H}_{\text{Net},q}$ for $q \in \{p-2, p-1, p, p+1, p\}$. This can be verified by inspecting the proof. (The main change concerns (135), where α^2 needs to be replaced by the product $\alpha_{k,\ell} \cdot \alpha_{k-1,r}$, for some $k \in \mathcal{K}$, which by assumption is again nonzero. All other steps remain unchanged.)

In a randomized setup where the cross-gains are chosen independently according to a continuous distribution all cross-gains are nonzero with probability 1 and all the principal submatrices of the network's channel matrix H_{Net} are full rank with probability 1. This implies in particular that the coding schemes in Subsections VI-C (for $K \leq t_\ell + r_\ell + 1$) or VI-E (for $K \geq t_\ell + r_\ell + 2$) achieve the optimal multiplexing gain $K - \left\lfloor \frac{K}{t_\ell+r_\ell+2} \right\rfloor$ with probability 1.

VII. Proof of Proposition 6

We first prove the lower bound in 2), followed by the lower bounds in 3), 1), and 4).

A. Proof of Lower Bound 2), i.e., (21)

If $t_\ell = 0$, then (21) follows from lower bound (22). Moreover, if $t_\ell + r_\ell \leq 1$, then there is nothing to prove, as the multiplexing gain cannot be negative.

Thus, in the following we assume that $t_\ell + r_\ell \geq 2$ and $t_\ell \geq 1$, and present a scheme that achieves the lower bound in (21) under this assumption. Our scheme is similar to the scheme for the asymmetric network described in Section V-A when this latter is specialized to $t_r = r_r = 0$. (In particular our scheme here disregards the right side-information available to the transmitters and the receivers.)

The idea is again to silence some of the transmitters, which decomposes our asymmetric network into several subnets, and to apply a scheme based on Costa's dirty-paper coding and successive interference cancellation to communicate over the subnets. However, here, due to the two-sided interference, *pairs* of consecutive transmitters are silenced and the dirty-paper coding and the successive interference cancellation strategies are used to "cancel" *two* interference signals.

Define

$$\beta_2 \triangleq (t_\ell + r_\ell + 1) \tag{76}$$

$$\gamma_2 \triangleq \left\lfloor \frac{K}{\beta_2} \right\rfloor \tag{77}$$

and recall that in Proposition 6 we defined $\kappa_2 \triangleq K \mod \beta_2$ and

$$\theta_2 \triangleq \begin{cases} 2, & \text{if } \kappa_2 \geq 2 \\ 1, & \text{if } \kappa_2 = 1 \\ 0, & \text{if } \kappa_2 = 0. \end{cases} \tag{78}$$

1) Splitting the Network Into Subnets: We silence transmitters $\{m\beta_2 + 1\}_{m=0}^{\gamma_2-1}$ and transmitters $\{m\beta_2\}_{m=1}^{\gamma_2}$. Moreover, if $\theta_2 = 1$ we also silence transmitter $(\gamma_2\beta_2 + 1)$ and if $\theta_2 = 2$ then also transmitters $(\gamma_2\beta_2+1)$ and K. Notice that in total we silence $2\gamma_2 + \theta_2$ transmitters. Silencing the chosen subset of transmitters splits the network into γ_2 non-interfering subnets if $\theta_2 = 0$ and into $\gamma_2 + 1$ non-interfering subnets if $\theta_2 \geq 1$. In both cases, the first γ_2 subnets all have the same topology and consist of $\beta_2 - 2$ active transmit antennas and of β_2 receive antennas. In fact, the m-th subnet, for $m \in \{1, \ldots, \gamma_2\}$, consists of transmit antennas $((m-1)\beta_2 + 2), \ldots, (m\beta_2 - 1)$ and receive antennas $((m-1)\beta_2 + 1), \ldots, m\beta_2$. We call these subnets generic. If $\theta_2 \geq 1$, then there is an additional last smaller subnet which consists of $\max\{\kappa_2 - 2, 0\}$ active transmit antennas and κ_2 receive antennas. More precisely, it consists of transmit antennas $(K-\kappa_2+2), \ldots, (K-1)$ (i.e., of no transmit antennas if $\kappa \leq 2$) and of receive antennas $(K-\kappa_2+1), \ldots, K$.

The scheme employed over a subnet depends on whether the scheme is generic or reduced and on the parameter $r_\ell \geq 0$. We describe the schemes in the following subsections.

2) Scheme Over a Generic Subnet When $r_\ell \geq 1$: We assume that the first subnet is generic and describe the scheme for this first subnet.

We transmit Messages $M_2, \ldots, M_{r_\ell + t_\ell}$ over the subnet. Define the sets

$$\mathcal{F}_1 \triangleq \{2, \ldots, r_\ell + 1\} \tag{79}$$
$$\mathcal{F}_2 \triangleq \{r_\ell + 2, \ldots, r_\ell + t_\ell\}. \tag{80}$$

Messages $\{M_k\}_{k \in \mathcal{F}_1}$ are transmitted as follows.

- For each $k \in \mathcal{F}_1$ we construct a single-user Gaussian codebook $\mathcal{C}_k$ of power $\alpha^2 P$, blocklength n, and rate $R_k = \frac{1}{2}\log(1 + \alpha^2 P)$. The code $\mathcal{C}_k$ is revealed to Transmitter k and to Receivers $2, \ldots, k$.
- Each Transmitter $k \in \mathcal{F}_1$ ignores the side-information about other transmitters' messages and codes for a Gaussian single-user channel. That is, it picks the codeword from codebook $\mathcal{C}_k$ that corresponds to its message M_k. Denoting the resulting sequence by $\tilde{X}_k^n$, the transmitter sends the scaled version
$$X_k^n = \frac{1}{\alpha}\tilde{X}_k^n \tag{81}$$
over the channel.
- Receiver $k \in \mathcal{F}_1$, uses successive interference cancellation to decode its desired Message M_k. Let $\hat{X}_0^n$ and $\hat{X}_1^n$ be two all-zero sequences of length n. Receiver k initializes j to 2, and while $j \leq k$:
  - It decodes Message M_j based on the difference
  $$Y_{j-1}^n - \alpha\hat{X}_{j-2}^n - \hat{X}_{j-1}^n \tag{82}$$
  using an optimal ML-decoder. Let $\hat{M}_j$ denote the resulting guess.
  - It picks the codeword from codebook $\mathcal{C}_j$ that corresponds to the guess $\hat{M}_j$ and scales it by $\frac{1}{\alpha}$ to form the guess $\hat{X}_j^n$.
  - It increases the index j by 1.
- Notice that Receiver $k \in \mathcal{F}_1$ has access to the output signals $Y_1^n, \ldots, Y_k^n$ because $k \leq r_\ell + 1$.
- For each $k \in \mathcal{F}_1$, if the previous two messages were decoded correctly, $\hat{M}_{k-2} = M_{k-2}$ and $\hat{M}_{k-1} = M_{k-1}$,
$$Y_{k-1}^n - \alpha\hat{X}_{k-2}^n - \hat{X}_{k-1}^n = \tilde{X}_k^n + N_{k-1}^n. \tag{83}$$
Thus, in this case, Message M_k is decoded based on the interference-free outputs $\alpha X_k^n + N_{k-1}^n$, and, by construction of the code $\mathcal{C}_k$, the average probability of error
$$\Pr[\hat{M}_k = M_k] \to 0 \quad \text{as } n \to \infty. \tag{84}$$

If $t_\ell \geq 2$, Messages $\{M_k\}_{k \in \mathcal{F}_2}$ are transmitted as follows:

- For each $k \in \mathcal{F}_2$, construct a dirty-paper code $\mathcal{C}_k$ of power $\alpha^2 P$ and rate $R_k = \frac{1}{2}\log(1 + \alpha^2 P)$ for noise variance 1 and interference variance $(\alpha^2 P + P)$ (which is the variance of $\alpha X_{k-2} + X_{k-1}$). The code $\mathcal{C}_k$ is revealed to Transmitters $k, \ldots, r_\ell + t_\ell$ and to Receiver k.
- Each Transmitter $k \in \mathcal{F}_2$ computes the "interference term" $\alpha X_{k-2}^n + X_{k-1}^n$ and applies the dirty-paper code $\mathcal{C}_k$ to encode its message M_k and mitigate the "interference" $\alpha X_{k-2}^n + X_{k-1}^n$. Denoting the resulting sequence by $\tilde{X}_k^n$, the transmitter sends the scaled version
$$X_k^n = \frac{1}{\alpha}\tilde{X}_k^n. \tag{85}$$
- Each Receiver $k \in \mathcal{F}_2$ considers only the outputs at the antenna of its left neighbor, Y_{k-1}^n. It uses code $\mathcal{C}_k$ to apply dirty-paper decoding based on the outputs
$$Y_{k-1}^n = \alpha X_{k-2}^n + X_{k-1}^n + \alpha X_k^n + N_k^n \tag{86}$$
$$= \tilde{X}_k^n + \underbrace{\alpha X_{k-2}^n + X_{k-1}^n}_{\text{"interference"}} + N_k^n. \tag{87}$$
- Notice that Transmitter $k \in \mathcal{F}_2$ can compute the sequences X_{k-2}^n and X_{k-1}^n, because in our scheme they only depend on Messages $M_{r_\ell}, \ldots, M_{k-2}$ and $M_{r_\ell}, \ldots, M_{k-1}$, respectively.
- By construction, the sequence $\tilde{X}_k^n$, which encodes Message M_k, can completely mitigate the "interference" $\alpha X_{k-2}^n + X_{k-1}^n$, and the average probability of error
$$\Pr[\hat{M}_k = M_k] \to 0 \quad \text{as } n \to \infty. \tag{88}$$

To summarize, with the described scheme, we sent Messages $M_2, \ldots, M_{r_\ell + t_\ell}$ with vanishingly small probability of error, see (84) and (88), and at rates

$$R_2 = \cdots = R_{r_\ell + t_\ell} = \frac{1}{2}\log(1 + \alpha^2 P). \tag{89}$$

3) Scheme Over a Generic Subnet When $r_\ell = 0$: In this case the set $\mathcal{F}_1$ is empty whereas by the assumption $t_\ell + r_\ell \geq 2$, $t_\ell \geq 2$ and the set $\mathcal{F}_2$ is non-empty. We transmit Messages $\{M_{k-1}\}_{k \in \mathcal{F}_2}$ over the subnet.

Specifically, each Transmitter $k \in \mathcal{F}_2$ employs the dirty-paper scheme as described in th previous subsection VII-A.2, except that now, instead of sending its own message M_k, it sends its left-neighbor's message M_{k-1} (to which it has access because $t_\ell \geq 1$). Accordingly, the outputs Y_{k-1}^n, for $k \in \mathcal{F}_2$, are now used by Receiver $k-1$ to decode its desired message M_{k-1}.

Here, for each $k \in \mathcal{F}_2$, the probability of error of Message M_{k-1} equals the probability of error of Message M_k in the previous subsection VII-A.2. Thus, by (88), for all $k \in \mathcal{F}_2$:

$$\Pr[\hat{M}_{k-1} = M_{k-1}] \to 0 \quad \text{as } n \to \infty. \tag{90}$$

We conclude that with the described scheme, the messages $M_1, \ldots, M_{t_\ell - 1}$ are communicated with vanishingly small probability of error and at rates

$$R_1 = \cdots = R_{t_\ell - 1} = \frac{1}{2}\log(1 + \alpha^2 P). \tag{91}$$

Conclusion 3: Our schemes for generic subnets described here and in the previous subsection VII-A.2 achieve a multiplexing gain of $r_\ell + t_\ell - 1$ over a generic subnet when $r_\ell \geq 1$ and when $r_\ell = 0$, respectively. Both schemes use all the $(t_\ell + r_\ell - 1)$ active transmit antennas of the subnet; but they use only the first $(t_\ell + r_\ell - 1)$ receive antennas and ignore the last two receive antennas of the subnet.

4) Scheme Over a Reduced Subnet: Over the reduced subnet we use one of the two schemes for generic subnets of Subsections VII-A2 and VII-A3, but with reduced side-information parameters

$$r'_\ell \triangleq \min\left[(\kappa_2 - 1), r_\ell\right] \tag{92a}$$

$$t'_\ell \triangleq \min\left[(\kappa_2 - r_\ell - 1)_+, t_\ell\right]. \tag{92b}$$

By Conclusion 3, this achieves a multiplexing gain of $\max\{\kappa_2 - 2, 0\}$ over a reduced subnet.

5) Analysis of the Performance Over the Entire Network: Over the first $\lfloor K/\beta_2 \rfloor$ generic subnets we achieve a multiplexing gain of $\beta_2 - 2$ and, if it exists, then over the last reduced subnet we achieve a multiplexing gain of $\max\{\kappa_2 - 2, 0\}$. Thus, over the entire network we achieve a multiplexing gain of

$$K - 2\gamma_2 - \theta_2 = \begin{cases} K - 2\lfloor K/\beta_2 \rfloor - 2, & \text{if } \kappa_2 \geq 2 \\ K - 2\lfloor K/\beta_2 \rfloor - \kappa_2 & \text{if } \kappa_2 < 2. \end{cases} \tag{93}$$

This establishes the desired lower bound.

B. Proof of Lower Bound 3)

By symmetry, this lower bound follows directly from (21). In particular, if $t_r \geq 1$ and $t_r + r_r \geq 2$, a scheme that is symmetric to the scheme described in the previous subsection VII-A achieves the desired multiplexing gain in 3). We briefly sketch this scheme because we will use it to prove the lower bound in 1), (20), in Subsection VII-C ahead.

Define

$$\beta'_2 \triangleq (t_r + r_r + 1), \tag{94}$$

$$\gamma'_2 \triangleq \left\lfloor \frac{K}{\beta'_2} \right\rfloor, \tag{95}$$

$$\kappa'_2 \triangleq K \mod \beta'_2, \tag{96}$$

and

$$\theta'_2 \triangleq \begin{cases} 2, & \text{if } \kappa'_2 \geq 2 \\ 1, & \text{if } \kappa'_2 = 1 \\ 0, & \text{if } \kappa'_2 = 0. \end{cases} \tag{97}$$

1) Splitting the Network Into Subnets: We silence transmitters $\{m\beta'_2 + 1\}_{m=0}^{\gamma'_2 - 1}$ and transmitters $\{m\beta'_2\}_{m=1}^{\gamma'_2}$. Moreover, if $\theta'_2 = 1$ then we also silence transmitter $(\gamma'_2\beta'_2 + 1)$ and if $\theta'_2 = 2$ then also transmitters $(\gamma'_2\beta'_2 + 1)$ and K. This splits the network into γ'_2 generic subnets with $\beta'_2 - 2$ active transmit antennas and β'_2 receive antennas, and if $\theta'_2 \in \{1, 2\}$ then there is an additional last reduced subnet with $\max\{\kappa'_2 - 2, 0\}$ active transmit antennas and κ'_2 receive antennas.

The scheme that we employ in the subnets depends on whether the subnet is generic or reduced and on the parameter $r_r \geq 0$.

2) Scheme Over a Generic Subnet When $r_r \geq 1$: Define the sets $\mathcal{F}_3$ and $\mathcal{F}_4$ as:

$$\mathcal{F}_3 \triangleq \{2, \ldots, t_r\}$$

$$\mathcal{F}_4 \triangleq \{t_r + 1, \ldots, t_r + r_r\}.$$

Assume that the first subnet is generic. Then, over this first subnet we transmit messages $M_2, \ldots, M_{t_r + r_r}$.

Messages $\{M_k\}_{k \in \mathcal{F}_3}$ are transmitted in a similar way as Messages $\{M_k\}_{k \in \mathcal{G}_3}$ in the scheme in Subsection V-A, and Message $\{M_k\}_{k \in \mathcal{F}_4}$ are transmitted in a similar way as Messages $\{M_k\}_{k \in \mathcal{G}_4}$ in that scheme. The only difference is that here, each dirty-paper code $\mathcal{C}_k$, for $k \in \mathcal{F}_3$, has to be designed for an interference variance $(\alpha^2 P + P)$ so that it can mitigate the "interference" $X^n_{k+1} + \alpha X^n_{k+2}$; likewise, during the successive interference cancellation steps, each Receiver $k \in \mathcal{F}_4$ has to cancel the two "interference" terms X^n_{k+1} and αX^n_{k+2}.

For brevity, we omit the details of the scheme and of the analysis. It can be shown that the scheme achieves a multiplexing gain of $t_r + r_r - 1$ over the generic subnet.

3) Scheme Over a Generic Subnet When $r_r = 0$: In this case, the set $\mathcal{F}_4$ is empty whereas, by the assumption $t_r + r_r \geq 2$, the set $\mathcal{F}_3$ is nonempty. We transmit messages $M_3, \ldots, M_{t_r + r_r + 1}$ over the subnet.

Messages $\{M_{k+1}\}_{k \in \mathcal{F}_3}$ are transmitted in the same way as messages $\{M_{k+1}\}_{k \in \mathcal{G}_3}$ in Subsection V-A. For brevity, we omit details and analysis. It can be shown that such a scheme achieves a multiplexing gain of $t_r + r_r - 1$ over the generic subnet.

Conclusion 4: Our schemes in the previous subsection VII-B2 and here achieve a multiplexing gain of $r_r + t_r - 1$ over a generic subnet when $r_r \geq 1$ and when $r_r = 0$, respectively. Both schemes use all $(t_r + r_r - 1)$ active transmit antennas of the subnet; but they use only the last $(t_r + r_r - 1)$ receive antennas and ignore the first two receive antennas of the subnet.

4) Scheme Over a Reduced Subnet: Over a reduced subnet we employ the schemes for a generic subnet described above, but with reduced side-information parameters

$$t'_r \triangleq \min\left[\left(\kappa'_2 - 2\right)_+, t_r\right] \tag{98a}$$

$$r'_r \triangleq \min\left[\left(\kappa'_2 - t_r - 2\right)_+, r_r\right] \tag{98b}$$

By Conclusion 4, such a scheme achieves a multiplexing gain of $\max\{\kappa'_2 - 2, 0\}$ over the reduced subnet.

C. Proof of Lower Bound 1), i.e., (20)

If $t_\ell + r_\ell = 0$ or $t_r + r_r = 0$, then the proof follows directly from the lower bounds in 2) or 3). If $t_\ell + t_r + r_\ell + r_r \leq 2$, there is nothing to prove.

Thus in the following we assume that $t_\ell + t_r + r_\ell + r_r \geq 3$ and $(t_\ell + r_\ell), (t_r + r_r) \geq 1$. Define

$$\beta_1 \triangleq (t_\ell + t_r + r_\ell + r_r) \tag{99}$$

$$\gamma_1 \triangleq \left\lfloor \frac{K}{\beta_1} \right\rfloor, \tag{100}$$

and recall that in Proposition 6 we defined $\kappa_1 \triangleq K \mod \beta_1$ and

$$\theta_1 \triangleq \begin{cases} 2, & \text{if } \kappa_1 \geq 2 \\ 1, & \text{if } \kappa_1 = 1 \\ 0, & \text{if } \kappa_1 = 0. \end{cases} \tag{101}$$

1) Splitting the Network Into Subnets: We silence transmitters $\{m\beta_1 + 1\}_{m=0}^{\gamma_1-1}$ and transmitters $\{m\beta_1\}_{m=1}^{\gamma_1}$. Moreover, if $\theta_1 = 1$, then we also silence transmitter $(\gamma_1\beta_1 + 1)$, and if $\theta_1 = 2$, then also transmitters $(\gamma_1\beta_1 + 1)$ and K. Thus, in total we silence $2\gamma_1 + \theta_1$ transmitters. This splits the network into γ_1 or $\gamma_1 + 1$ non-interfering subnets: the first γ_1 generic subnets consist of $(\beta_1 - 2)$ transmit antennas and β_1 receive antennas, and if there is an additional last subnet then it is smaller and consists of $\max\{\kappa_1 - 2, 0\}$ transmit antennas and of κ_1 receive antennas.

The scheme employed in each subnet depends on whether the subnet is generic or reduced.

2) Scheme Over a Generic Subnet: We assume that the first subnet is generic and describe the scheme for this first subnet. To this end, define the groups

$$\mathcal{F}_{1/2} \triangleq \{2, \ldots, r_\ell + t_\ell\}$$
$$\mathcal{F}_{3/4} \triangleq \{(r_\ell + t_\ell + 1), \ldots, (r_\ell + t_\ell + t_r + r_r - 1)\}.$$

Our scheme is a combination of the two schemes for generic subnets described in Sections VII-A and VII-B. Over the left part of the subnet that consists of transmit antennas $k \in \mathcal{F}_{1/2}$ and receive antennas $1, \ldots, (r_\ell + t_\ell - 1)$ we use the scheme in Section VII-A. Over the right part of the subnet that consists of transmit antennas $k \in \mathcal{F}_{3/4}$ and receive antennas $(r_\ell + t_\ell + 2), \ldots, (r_\ell + t_\ell + t_r + r_r)$ we use the scheme in Section VII-B where the set $\mathcal{F}_3$ needs to be replaced by $\{(r_\ell + t_\ell + 1), \ldots, (r_\ell + t_\ell + t_r - 1)\}$ and the set $\mathcal{F}_4$ by $\{(r_\ell + t_\ell + t_r), \ldots, (r_\ell + t_\ell + t_r + r_r - 1)\}$. Thus, the combined scheme utilizes all the transmit antennas in the subnet but only receive antennas $1, \ldots, r_\ell + t_\ell - 1$ and $r_\ell + t_\ell + 2, \ldots, r_\ell + t_\ell + t_r + r_r + 2$, i.e., it ignores the two receive antennas $(r_\ell + t_\ell)$ and $(r_\ell + t_\ell + 1)$, see also Conclusions 3 and 4.

Since the transmit antennas $k \in \mathcal{F}_{1/2}$ in the "left-hand" scheme do not influence the signals observed at receive antennas $(r_\ell + t_\ell + 2), \ldots, (r_\ell + t_\ell + t_r + r_r)$ employed in the "left-hand" scheme, and the signals sent at transmit antennas $k \in \mathcal{F}_{3/4}$ in the "right-hand" scheme do not influence the signals observed at receive antennas $1, \ldots, (r_\ell + t_\ell - 1)$ employed in the "left-hand" scheme, the performance of the two schemes can be analyzed separately. By Conclusions 3 and 4 we achieve a multiplexing gain of $r_\ell + t_\ell - 1$ over the left part of the subnet and a multiplexing gain of $t_r + r_r - 1$ over the right part of the subnet. Thus, we achieve a multiplexing gain $r_\ell + t_\ell + t_r + r_r - 2$ over the entire subnet.

3) Scheme Over a Reduced Subnet: We employ the same scheme as over a generic subnet but with reduced side-information parameters. Details and analysis are omitted for brevity. Such a scheme can achieve a multiplexing gain of $\max\{\kappa_1 - 2, 0\}$ over a reduced subnet.

4) Analysis of Performance Over the Entire Network: Over the first $\lfloor K/\beta_1 \rfloor$ generic subnets we achieve a multiplexing gain of $\beta_1 - 2$ and, if it exists, then over the last reduced subnet we achieve a multiplexing gain of $\max\{\kappa_1 - 2, 0\}$. Thus, over the entire network we achieve a multiplexing gain of

$$K - 2\gamma_1 - \theta_1 = \begin{cases} K - 2\lfloor K/\beta_1 \rfloor - 2, & \text{if } \kappa_1 \geq 2 \\ K - 2\lfloor K/\beta_1 \rfloor - \kappa_1 & \text{if } \kappa_1 < 2. \end{cases} \tag{102}$$

This establishes the desired lower bound.

D. Proof of Lower Bound 4), i.e., (22)

In our scheme the transmitters ignore their side-information. Define

$$\beta_3 \triangleq (r_\ell + r_r + 3) \tag{103}$$

$$\gamma_3 \triangleq \left\lfloor \frac{K}{\beta_3} \right\rfloor, \tag{104}$$

and recall that in Proposition 6 we defined $\kappa_3 \triangleq K \mod \beta_3$ and

$$\theta_3 \triangleq \begin{cases} 2, & \text{if } \kappa_3 \geq 2 \\ 1, & \text{if } \kappa_3 = 1 \\ 0, & \text{if } \kappa_3 = 0. \end{cases} \tag{105}$$

1) Splitting the Network Into Subnets: We silence transmitters $\{m\beta_3 + 1\}_{m=0}^{\gamma_3-1}$ and transmitters $\{m\beta_3\}_{m=1}^{\gamma_3}$. Moreover, if $\theta_3 = 1$, we also silence transmitter $\beta_3\gamma_3 + 1$, and if $\theta_3 = 2$, we also silence transmitters $\beta_3\gamma_3 + 1$ and K. Notice that in total we have silenced $2\gamma_3 + \theta_3$ transmitters.

This splits the network into γ_3 or $\gamma_3 + 1$ non-interfering subnets: the first γ_3 subnets consist of $\beta_3 - 2$ active transmit antennas and β_3 receive antennas (we call these subnets generic), and if an additional last subnet exists it is smaller and consists of $\max\{\kappa_3 - 2, 0\}$ transmit and κ_3 receive antennas.

The scheme employed over the subnets depends on whether the subnet is generic or reduced.

2) Scheme Over a Generic Subnet: We assume that the first subnet is generic and describe our scheme for this first subnet. Define

$$\mathcal{H}_1 \triangleq \{2, \ldots, r_\ell + 1\}$$
$$\mathcal{H}_2 \triangleq \{r_\ell + 2\}$$
$$\mathcal{H}_3 \triangleq \{r_\ell + 3, \ldots, r_\ell + r_r + 2\}.$$

We only sketch the scheme.

- Messages $M_2, \ldots, M_{r_\ell + r_r + 2}$ are transmitted over the subnet.
- For each $k \in (\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3)$, Transmitter k encodes its Message M_k using a Gaussian point-to-point code.

- For each $k \in \mathcal{H}_1$, Receiver k decodes its message using successive interference cancellation from the *left*, starting with the *first* antenna in the subnet. These messages can be decoded with arbitrary small probability of error (for sufficiently large blocklengths), whenever

$$R_k \leq \frac{1}{2}\log\left(1+\alpha^2 P\right), \quad \forall k \in \mathcal{H}_1. \tag{106}$$

- Similarly, for each $k \in \mathcal{H}_3$, Receiver k decodes its message using successive interference cancellation but now from the *right* and starting with the *last* antenna in the subnet. These messages can be decoded with arbitrary small probability of error (for sufficiently large blocklengths), whenever

$$R_k \leq \frac{1}{2}\log\left(1+\alpha^2 P\right), \quad \forall k \in \mathcal{H}_3. \tag{107}$$

- Receiver $r_\ell+2$, which has access to antennas $2,\ldots,(r_\ell+r_r+2)$, decodes its desired Message $M_{r_\ell+2}$ by decoding *all* the transmitted messages $M_2,\ldots,M_{r_\ell+r_r+2}$ using an optimal MIMO decoder [41]. In this step, we have arbitrary small probability of error, whenever

$$\sum_{i=2}^{r_\ell+r_r+2} R_i \leq \frac{1}{2}\log\left(\det\left(\mathsf{I}+P\mathsf{H}^{\mathsf{T}}_{r_\ell+r_r+1}\mathsf{H}_{r_\ell+r_r+1}\right)\right); \tag{108}$$

where here for ease of notation we wrote $\mathsf{H}_{r_\ell+r_r+1}$ instead of $\mathsf{H}_{r_\ell+r_r+1}(\alpha)$. Notice that since the channel matrix $H_{r_\ell+r_r+1}(\alpha)$ is non-singular and does not depend on the power P, by [41]:

$$\overline{\lim_{P\to\infty}} \frac{\frac{1}{2}\log\left(\det\left(\mathsf{I}+P\mathsf{H}^{\mathsf{T}}_{r_\ell+r_r+1}\mathsf{H}_{r_\ell+r_r+1}\right)\right)}{\frac{1}{2}\log(P)} = r_\ell+r_r+1. \tag{109}$$

Combining (106)–(109), we conclude that the described scheme can achieve a multiplexing gain of $r_\ell+r_r+1$ over the entire subnet.

3) Scheme Over a Reduced Subnet: We employ the same scheme as over a generic subnet but with reduced side-information parameters. Such a scheme can achieve a multiplexing gain of $\max\{\kappa_3-2,0\}$ over a reduced subnet. Details and analysis omitted.

4) Analysis of Performance Over the Entire Network: Over the first $\lfloor K/\beta_3\rfloor$ generic subnets we achieve a multiplexing gain of β_3-2 and, if it exists, then over the last reduced subnet we achieve a multiplexing gain of $\max\{\kappa_3-2,0\}$. Thus, over the entire network we achieve a multiplexing gain of

$$K-2\gamma_3-\theta_3 = \begin{cases} K-2\lfloor K/\beta_3\rfloor-2, & \text{if } \kappa_3\geq 2\\ K-2\lfloor K/\beta_3\rfloor-\kappa_3 & \text{if } \kappa_3<2. \end{cases} \tag{110}$$

This establishes the desired lower bound.

E. General Cross-Gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$

Our proofs and results presented in the previous subsections generalize to non-equal cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$.

For example, the three coding schemes in Subsections VII-A–VII-C are solely based on silencing a subset of the transmitters (which splits the network into subnets), on dirty-paper coding interference sequences that are known at the transmitter, and on successive interference cancellation at the receivers. All these techniques do not rely on the fact that the values of the cross-gains are all equal and apply in the same way also to setups with general cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$. Thus, the results under items 1)–3) remain valid also for general cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$.

We present now in detail how to adapt the description of the scheme and the analysis in Subsection VII-A to the setup with general cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$. In (81), (85), (89), and (91) α needs to be replaced by $\alpha_{k-1,r}$; in (82) α needs to be replaced by $\alpha_{j-1,\ell}$; in (83) and (87) it needs to be replaced by $\alpha_{k-1,\ell}$; and in (86) the first α needs to be replaced by $\alpha_{k-1,\ell}$ and the second by $\alpha_{k-1,r}$. Also, the codebooks $\mathcal{C}_k$, for $k\in\mathcal{F}_1$ or $k\in\mathcal{F}_2$, should be of power $\alpha^2_{k-1,r}P$ and rate $R_k=\frac{1}{2}\log(1+\alpha^2_{k-1,r}P)$, and the codebooks $\mathcal{C}_k$, for $k\in\mathcal{F}_2$, should be designed for the interference $\alpha_{k-1,\ell}X^n_{k-2}+X^n_{k-1}$ which is of power $\alpha^2_{k-1,\ell}P+P$. Finally, Receivers $k\in\mathcal{F}_1$ produce their estimates $\hat{X}^n_j$, $j\leq k$, by picking the codeword in $\mathcal{C}_j$ that corresponds to their estimate $\hat{M}_j$ and scaling it by $\alpha_{j-1,r}$. This way, if $\hat{M}_{k-2}=M_{k-2}$ and $\hat{M}_{k-1}=M_{k-1}$, Receiver k can decode its desired message M_k based on the interference-free output $\alpha_{k-1,r}X^n_k+N^n_{k-1}$.

The scheme in Subsection VII-D is based on silencing a subset of the transmitters (which splits the network into subnets), on successive interference cancellation, and on MIMO decoding. The multiplexing gain achieved by the scheme relies on the cross-gains only through the rank of the subnets' channel matrices which shows up in the performance analysis of the MIMO decoding, (108) and (109). To achieve the multiplexing gain in (110) all the subnets' channel matrices need to have full rank $t_\ell+r_\ell+1$. Thus, the multiplexing-gain in (110) is achievable also in a setup with general cross-gains $\{\alpha_{k,\ell}\}$ and $\{\alpha_{k,r}\}$ if all the subnets have full-rank channel matrices. This is in particular the case (with probability 1) in a randomized setup where all the cross-gains are drawn according to continuous distribution, where all the subnets have full-rank channel matrices.

Proof of Proposition 7

A. Proof of Upper Bound 1), i.e., (23)

Define

$$\beta_4 \triangleq t_\ell+t_r+r_\ell+r_r+4,$$
$$\gamma_4 \triangleq \left\lfloor\frac{K}{\beta_4}\right\rfloor,$$

and recall that $\kappa_4\triangleq K-\gamma_4\beta_4$ and that θ_4 equals 1 if $\kappa_4\geq\min\{t_\ell+r_\ell+1,t_r+r_r+1\}$, and it equals 0 otherwise.

The proof is based on the Dynamic-MAC Lemma 9. To describe the choice of parameters for which we wish to apply this lemma, we need the following definitions. Define for every positive integer $p\geq 2$ and every non-zero number α the matrix $\mathsf{M}_p(\alpha)$ as the $p\times p$ matrix with diagonal elements α, first upper off-diagonal elements 1, second upper off-diagonal elements α, and all other elements 0. That means, the row-j_r column-j_c entry of the matrix $\mathsf{M}_p(\alpha)$ equals α if $j_r = j_c$ or $j_r = j_c - 2$, it equals 1 if $j_r = j_c - 1$, and it equals 0 otherwise. Let $\mathsf{M}_p^{\text{inv}}(\alpha)$ denote the inverse matrix of $\mathsf{M}_p(\alpha)$. This inverse always exists because $\det(\mathsf{M})_p(\alpha) = \alpha^p$, which by our assumption $\alpha \neq 0$ is nonzero. As we will see shortly, our main interest lies in the inverses $\mathsf{M}^{\text{inv}}_{t_\ell+r_\ell+1}(\alpha)$ and $\mathsf{M}^{\text{inv}}_{t_r+r_r+1}(\alpha)$. To simplify notation, we therefore denote the row-j_r column-j_c entry of $\mathsf{M}^{\text{inv}}_{t_\ell+r_\ell+1}(\alpha)$ by a_{j_r,j_c} and the row-j_r column-j_c entry of $\mathsf{M}^{\text{inv}}_{t_r+r_r+1}(\alpha)$ by b_{j_r,j_c}.

We treat the cases $\theta_4 = 0$ and $\theta_4 = 1$ separately. If $\theta_4 = 1$, then we apply Lemma 9 to the following choices:

- $q = 1$;
- $g = 2\gamma_4$;
- $\mathcal{A} = \bigcup_{m=0}^{\gamma_4-1} \mathcal{A}'(m)$, where for $m \in \{0, \ldots, \gamma_4 - 2\}$:

$$\mathcal{A}'(m) \triangleq \{(m\beta_4 + r_\ell + 2), \ldots, (m\beta_4 + r_\ell + t_\ell + t_r + 3)\}, \tag{111}$$

and

$$\mathcal{A}'(\gamma_4 - 1) \triangleq \{((\gamma_4 - 1)\beta_4 + r_\ell + 2), \ldots, (\gamma_4\beta_4 - r_r + 3)\} \cup\{(\gamma_4\beta_4 + r_\ell + 2), \ldots, K\}; \tag{112}$$

- $\mathcal{B}_1 = \mathcal{K}\backslash\mathcal{A}$;
- for i even and $0 \le i \le g$:

$$\boldsymbol{V}_i = \sum_{j=1}^{t_r+r_r+1} \alpha b_{1,j} \boldsymbol{N}_{\frac{i}{2}\beta_4 - j} + \sum_{j=1}^{t_\ell+r_\ell+1} (a_{1,j} + \alpha a_{2,j}) \boldsymbol{N}_{\frac{i}{2}\beta_4+1+j} - \boldsymbol{N}_{\frac{i}{2}\beta_4+1}, \tag{113}$$

and for i odd and $1 \le i \le g - 1$:

$$\boldsymbol{V}_i = \sum_{j=1}^{t_r+r_r+1} (b_{1,j} + \alpha b_{2,j}) \boldsymbol{N}_{\frac{i-1}{2}\beta_4 - j} + \sum_{j=1}^{t_\ell+r_\ell+1} \alpha a_{1,j} \boldsymbol{N}_{\frac{i-1}{2}\beta_4+1+j} - \boldsymbol{N}_{\frac{i-1}{2}\beta_4}. \tag{114}$$

Thus, if $\theta_4 = 1$,

$$\mathcal{K}\backslash\mathcal{R}_{\mathcal{A}} = \big\{m\beta_4 + 1, (m+1)\beta_4\big\}_{m=0}^{\gamma_4-1} \cup \{\gamma_4\beta_4 + 1\}. \tag{115}$$

If $\theta_4 = 0$, we apply Lemma 9 to the choices

- $q = 1$;
- $g = 2\gamma_4 - 1$;
- $\mathcal{A} = \bigcup_{m=0}^{\gamma_4-1} \mathcal{A}'(m)$, where $\{\mathcal{A}'(m)\}_{m=0}^{\gamma_4-2}$ are defined in (111) and where

$$\mathcal{A}'(\gamma_4-1) \triangleq \{((\gamma_4-1)\beta_4+r_\ell+2), \ldots, (K-r_r-1)\}; \tag{116}$$

- $\mathcal{B}_1 = \mathcal{K}\backslash\mathcal{A}$;
- $\{\boldsymbol{V}_m\}_{m=0}^{2(\gamma_4-1)}$ are given by (113) and (114) and

$$\boldsymbol{V}_{2\gamma_4-1} = \sum_{j=1}^{t_r+r_r+1} (b_{1,j} + \alpha b_{2,j}) \boldsymbol{N}_{K-j} - \boldsymbol{N}_K. \tag{117}$$

Thus, if $\theta_4 = 0$,

$$\mathcal{K}\backslash\mathcal{R}_{\mathcal{A}} = \big\{m\beta_4 + 1, (m+1)\beta_4\big\}_{m=0}^{\gamma_4-2} \cup \big\{(\gamma_4 - 1)\beta_4 + 1, K\big\}. \tag{118}$$

One readily verifies that both for $\theta_4 = 0$ and $\theta_4 = 1$ the differential entropy $h\big(\{\boldsymbol{N}_k\}_{k\in\mathcal{R}_{\mathcal{A}}} | \boldsymbol{V}_0, \ldots, \boldsymbol{V}_q\big)$ is finite and does not depend on the power constraint P, since neither does the genie-information. In Appendix B we show that also Assumption (26) in the Dynamic MAC Lemma is satisfied, and hence the lemma applies. It gives the desired upper bound, because by (115) and (118),

$$|\mathcal{R}_{\mathcal{A}}| = 2\gamma_4 + \theta_4. \tag{119}$$

B. Proof of Upper Bound 3), i.e., (24)

The proof is again based on the Dynamic-MAC Lemma 9. We first give some definitions.

Define

$$\beta_5 \triangleq t_\ell + t_r + r_\ell + r_r + 3 \tag{120}$$

$$\gamma_5 \triangleq \left\lfloor \frac{K}{\beta_5} \right\rfloor \tag{121}$$

and recall that $\kappa_5 \triangleq K - \beta_5\gamma_5$ and θ_5 equals 1 if $\kappa_5 \ge t_r + r_r + 2$ and 0 otherwise.

For $j_r, j_c \in \{1, \ldots, t_\ell + r_\ell + 1\}$, denote the row-$j_r$ column-j_c entry of the matrix $\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)$ by h_{j_r,j_c}. Also, choose a set of real numbers $\{d_2, \ldots, d_{t_\ell+r_\ell+1}\}$ so that

$$h_{1,j_c} = \sum_{j_r=2}^{t_\ell+r_\ell+1} d_{j_r} h_{j_r,j_c}, \quad j_c \in \{1, \ldots, t_\ell + r_\ell + 1\}. \tag{122}$$

Such a choice always exists because of the assumption $\det(\mathsf{H}_{t_\ell+r_\ell+1}) = 0$.

We treat the cases $\theta_5 = 1$ and $\theta_5 = 0$ separately. If $\theta_5 = 1$, we apply the Dynamic-MAC Lemma to the choices:

- $q = 2\gamma_5 + 1$;
- $g = 2\gamma_5$;
- $\mathcal{A} = \bigcup_{m=0}^{\gamma_5} \mathcal{A}''(m)$, where

$$\mathcal{A}''(0) \triangleq \{1, \ldots, t_r + 1\}, \tag{123}$$

for $m \in \{1, \ldots, \gamma_5 - 1\}$:

$$\mathcal{A}''(m) \triangleq \{m\beta_5 - t_\ell + 1, \ldots, m\beta_5 + t_r + 1\}, \tag{124}$$

and

$$\mathcal{A}''(\gamma_5) \triangleq \{(\gamma_5\beta_5 - t_\ell + 1), \ldots, (K - r_r - 1)\}; \tag{125}$$

- for i odd and $1 \le i \le 2\gamma_5 - 1$,

$$\mathcal{B}_i = \left\{\frac{(i-1)}{2}\beta_5 + t_r + r_r + r_\ell + 3\right\}; \tag{126}$$

for i even and $2 \le i \le 2\gamma_5$:

$$\mathcal{B}_i = \left\{\left(\left(\frac{i}{2} - 1\right)\beta_5 + t_r + 2\right), \ldots, \left(\left(\frac{i}{2} - 1\right)\beta_5 + t_r + r_r + r_\ell + 2\right)\right\}, \tag{127}$$

and

$$\mathcal{B}_{2\gamma_5+1} = \{(K - r_r), \ldots, K\}; \tag{128}$$

- for i even and $0 \leq i \leq 2(\gamma_5 - 1)$[9]:

$$\begin{aligned} \boldsymbol{V}_i = & \sum_{j=2}^{t_r+r_r+1} d_j \boldsymbol{N}_{\frac{i}{2}\beta_4+t_r+r_r+2+j} \\ & + \alpha \sum_{j_c=1}^{t_r+r_r+1} b_{1,j_c} \boldsymbol{N}_{\frac{i}{2}\beta_4+t_r+r_r+1-j_c} \\ & - \boldsymbol{N}_{\frac{i}{2}\beta_4+t_r+r_r+3}, \end{aligned} \tag{129}$$

for i odd and $1 \leq i \leq 2\gamma_5 - 1$:

$$\begin{aligned} \boldsymbol{V}_i = & \sum_{j_c=1}^{t_r+r_r+1} (\alpha b_{2,j_c} + b_{1,j_c}) \boldsymbol{N}_{\frac{i-3}{2}\beta_4+t_r+r_r+1-j_c} \\ & + \sum_{j_c=1}^{t_\ell+r_\ell+1} \alpha a_{1,j_c} \boldsymbol{N}_{\frac{i-3}{2}\beta_4+t_r+r_r+3+j_c} \\ & - \boldsymbol{N}_{\frac{i-3}{2}\beta_4+t_r+r_r+2}, \end{aligned} \tag{130}$$

and

$$\boldsymbol{V}_{2\gamma_5} \triangleq \sum_{j_c=1}^{t_r+r_r+1} (\alpha b_{2,j_c} + b_{1,j_c}) \boldsymbol{N}_{K-1-j_c} - \boldsymbol{N}_K. \tag{131}$$

Thus, if $\theta_5 = 1$,

$$\begin{aligned} \mathcal{K} \backslash \mathcal{R}_{\mathcal{A}} = & \left\{(m\beta_5 + t_r + r_r + 2), (m\beta_5 + t_r + r_r + 3)\right\}_{m=0}^{\gamma_5-1} \\ & \cup \{K\}. \end{aligned} \tag{132}$$

If $\theta_5 = 0$, we apply the Dynamic-MAC Lemma to the following choices:

- $q = 2\gamma_5$;
- $g = 2\gamma_5 - 1$;
- $\mathcal{A} = \bigcup_{m=0}^{\gamma_5} \mathcal{A}''(m)$, where $\left\{\mathcal{A}''(m)\right\}_{m=0}^{\gamma_5-1}$ are defined in (123) and (124), and where

$$\mathcal{A}''(\gamma_5) \triangleq \{(\gamma_5\beta_5 - t_\ell + 1), \ldots, K\};$$

- the sets $\left\{\mathcal{B}_i\right\}_{i=1}^{2\gamma_5}$ are defined in (126) and (127);
- $\{\boldsymbol{V}_m\}_{m=0}^{\gamma_5-1}$ is defined in (129) and (130).

Thus, if $\theta_5 = 0$,

$$\mathcal{K} \backslash \mathcal{R}_{\mathcal{A}} = \left\{(m\beta_5 + t_r + r_r + 2), (m\beta_5 + t_r + r_r + 3)\right\}_{m=0}^{\gamma_5-1}. \tag{133}$$

One readily verifies that both for $\theta_5 = 0$ and $\theta_5 = 1$ the differential entropy $h\left(\{\boldsymbol{N}_k\}_{k \in \mathcal{R}_{\mathcal{A}}} | \boldsymbol{V}_0, \ldots, \boldsymbol{V}_q\right)$ is finite and does not depend on the power constraint P, since neither does the genie-information. In Appendix B we show that also Assumption (26) of the Dynamic-MAC Lemma is satisfied, and hence the lemma applies. It gives the desired upper bound, because by (132) and (133),

$$|\mathcal{R}_{\mathcal{A}}| = 2\gamma_5 + \theta_5. \tag{134}$$

[9]Recall that a_{j_r,j_c} denotes the row-j_r column-j_c entry of the matrix $\mathsf{M}^{\text{inv}}_{t_\ell+r_\ell+1}(\alpha)$ defined in the previous Subsection A; and where similarly b_{j_r,j_c} denotes the row-j_r column j_c entry of the matrix $\mathsf{M}^{\text{inv}}_{t_r+r_r++1}(\alpha)$ also defined in Subsection A.

Appendix A
Proof of Lemma 11

By definition, $\det(\mathsf{H}_1(\alpha)) = 1$. Therefore, the integer p has to be at least 2 and Statement 1.) in the lemma follows.

Statement 2.) can be proved as follows. We define $\mathsf{H}_0(\alpha) \triangleq 1$ and note that also $\mathsf{H}_1(\alpha) = 1$, irrespective of α. We then have for each positive integer $q \geq 2$:

$$\det\left(\mathsf{H}_q(\alpha)\right) = \det\left(\mathsf{H}_{q-1}(\alpha)\right) - \alpha^2 \det\left(\mathsf{H}_{q-2}(\alpha)\right). \tag{135}$$

Thus, $\det\left(\mathsf{H}_p(\alpha)\right) = 0$ implies that the two determinants $\det\left(\mathsf{H}_{p-1}(\alpha)\right)$ and $\det\left(\mathsf{H}_{p-2}(\alpha)\right)$ are either both 0 or both non-zero, and similarly, that the two determinants $\det\left(\mathsf{H}_{p+1}(\alpha)\right)$ and $\det\left(\mathsf{H}_{p+2}(\alpha)\right)$ are either both 0 or both non-zero. Applying this argument iteratively, we see that the determinants $\det\left(\mathsf{H}_{p-2}(\alpha)\right)$ and $\det\left(\mathsf{H}_{p-1}(\alpha)\right)$ can only be 0 if all "previous" determinants $\det\left(\mathsf{H}_0(\alpha)\right), \ldots, \det\left(\mathsf{H}_{p-3}(\alpha)\right)$ are zero. Similarly, for the determinants $\det\left(\mathsf{H}_{p+1}(\alpha)\right)$ and $\det\left(\mathsf{H}_{p+2}(\alpha)\right)$. However, since $\det\left(\mathsf{H}_0(\alpha)\right) = \det\left(\mathsf{H}_1(\alpha)\right) = 1$, we conclude that $\det\left(\mathsf{H}_{p-2}(\alpha)\right), \det\left(\mathsf{H}_{p-1}(\alpha)\right), \det\left(\mathsf{H}_{p+1}(\alpha)\right)$, and $\det\left(\mathsf{H}_{p+2}(\alpha)\right)$ must be non-zero, which proves Statement 2.)

Appendix B
Proof That Assumption (26) Holds in Section VIII-A

By (115) and (118) it suffices to show that if $\theta_4 = 0$, then the output sequences $\{\boldsymbol{Y}_{m\beta_4+1}, \boldsymbol{Y}_{(m+1)\beta_4}\}_{m=0}^{\gamma_4-2}$, $\boldsymbol{Y}_{(\gamma_4-1)\beta_4+1}$, and $\boldsymbol{Y}_K$ can be reconstructed, and if $\theta_4 = 1$, then the output sequences $\{\mathbf{Y}_{\mathbf{m}\beta_4+\mathbf{1}}, \mathbf{Y}^{\mathbf{n}}_{(\mathbf{m}+\mathbf{1})\beta_4}\}_{\mathbf{m}=\mathbf{0}}^{\gamma_4-\mathbf{1}}$ and $\boldsymbol{Y}_{\gamma_4\beta_4+1}$ can be reconstructed.

Notice first that using the given encoding functions $f_1, \ldots, f_n$ the input sequences $\{\boldsymbol{X}_{m\beta_4+t_\ell+r_\ell+2}, \boldsymbol{X}_{m\beta_4+t_\ell+r_\ell+3}\}_{m=0}^{\gamma_4-1}$ can be computed from Messages $\{M_k\}_{k \in \mathcal{A}}$. Moreover, if $\theta_4 = 0$ then additionally also the input sequences $\boldsymbol{X}_{(\gamma_4-1)+r_\ell+t_\ell+4}, \ldots, \boldsymbol{X}_{K-r_r-t_r-1}$ can be computed from $\{M_k\}_{k \in \mathcal{A}}$, and if $\theta_4 = 1$ additionally also the input sequences $\boldsymbol{X}_{\gamma_4\beta_4+t_\ell+r_\ell+2}, \boldsymbol{X}_{K-r_r-t_r-1}$ can be computed from $\{M_k\}_{k \in \mathcal{A}}$. The result is then proved by showing that each of the desired output sequences can be expressed as a linear combination of the genie-information, these reconstructed inputs, and the outputs observed by the group-$\mathcal{A}$ receivers.

We start with $\boldsymbol{Y}_{\beta_4}$. Notice that by the channel law (9), the linear systems (136) and (137) on top of the next page hold for every time-instant $t \in \{1, \ldots, n\}$. Recalling that a_{j_r,j_c} denotes the row-j_r column-j_c entry of the inverse matrix $\mathsf{M}^{\text{inv}}_{t_\ell+r_\ell+1}(\alpha)$ and that b_{j_r,j_c} denotes the row-j_r column-j_c entry of the inverse matrix $\mathsf{M}^{\text{inv}}_{t_r+r_r+1}(\alpha)$, it is easily checked that (136) implies:

$$\begin{aligned} & \sum_{j=1}^{t_r+r_r+1} b_{2,j} \boldsymbol{Y}_{\beta_4-j} - (b_{2,t_r+r_r}\alpha + b_{2,t_r+r_r+1}) \boldsymbol{X}_{t_\ell+r_\ell+3} \\ & \qquad - b_{2,t_r+r_r+1}\alpha \boldsymbol{X}_{t_\ell+r_\ell+2} \\ & = \boldsymbol{X}_{\beta_4-1} + \sum_{j=1}^{t_r+r_r+1} b_{2,j} \boldsymbol{N}_{\beta_4-j}; \end{aligned} \tag{138}$$

$$\begin{pmatrix} Y_{\beta_4-1,t} \\ Y_{\beta_4-2,t} \\ \vdots \\ Y_{t_\ell+r_\ell+4,t} \\ Y_{t_\ell+r_\ell+3,t} \end{pmatrix} = \mathsf{M}_{t_r+r_r+1}(\alpha) \begin{pmatrix} X_{\beta_4,t} \\ X_{\beta_4-1,t} \\ \vdots \\ X_{t_\ell+r_\ell+5,t} \\ X_{t_\ell+r_\ell+4,t} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \alpha X_{t_\ell+r_\ell+3} \\ \alpha X_{t_\ell+r_\ell+2,t} + X_{t_\ell+r_\ell+3,t} \end{pmatrix} + \begin{pmatrix} N_{\beta_4-1,t} \\ N_{\beta_4-2,t} \\ \vdots \\ N_{t_\ell+r_\ell+4,t} \\ N_{t_\ell+r_\ell+3,t} \end{pmatrix} \tag{136}$$

$$\begin{pmatrix} Y_{\beta_4+2,t} \\ Y_{\beta_4+3,t} \\ \vdots \\ Y_{\beta_4+t_\ell+r_\ell+1,t} \\ Y_{\beta_4+t_\ell+r_\ell+2,t} \end{pmatrix} = \mathsf{M}_{t_\ell+r_\ell+1}(\alpha) \begin{pmatrix} X_{\beta_4+1,t} \\ X_{\beta_4+2,t} \\ \vdots \\ X_{\beta_4+t_\ell+r_\ell,t} \\ X_{\beta_4+t_\ell+r_\ell+1,t} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \alpha X_{\beta_4+t_\ell+r_\ell+2} \\ X_{\beta_4+t_\ell+r_\ell+2,t} + \alpha X_{\beta_4+t_\ell+r_\ell+3,t} \end{pmatrix} + \begin{pmatrix} N_{\beta_4+2,t} \\ N_{\beta_4+3,t} \\ \vdots \\ N_{\beta_4+t_\ell+r_\ell+1,t} \\ N_{\beta_4+t_\ell+r_\ell+2,t} \end{pmatrix} \tag{137}$$

and

$$\begin{aligned} \sum_{j=1}^{t_r+r_r+1} & b_{1,j} \boldsymbol{Y}_{\beta_4-j} - (b_{1,t_r+r_r}\alpha + b_{1,t_r+r_r+1})\boldsymbol{X}_{t_\ell+r_\ell+3} \\ & \qquad - b_{1,t_r+r_r+1}\alpha \boldsymbol{X}_{t_\ell+r_\ell+2} \\ &= \boldsymbol{X}_{\beta_4} + \sum_{j=1}^{t_r+r_r+1} b_{1,j}\boldsymbol{N}_{\beta_4-j}; \end{aligned} \tag{139}$$

and that (137) implies:

$$\begin{aligned} \sum_{j=1}^{t_\ell+r_\ell+1} & a_{1,j}\boldsymbol{Y}_{\beta_4+1+j} - a_{1,t_r+r_r+1}\alpha \boldsymbol{X}_{\beta_4+t_\ell+r_\ell+3} \\ & -(a_{1,t_r+r_r} + a_{1,t_r+r_r+1}\alpha)\boldsymbol{X}_{\beta_4+t_\ell+r_\ell+2} \\ &= \boldsymbol{X}_{\beta_4+1} + \sum_{j=1}^{t_\ell+r_\ell+1} a_{1,j}\boldsymbol{N}_{\beta_4+1+j}. \end{aligned} \tag{140}$$

Since the genie-information has been chosen so that

$$\begin{aligned} \boldsymbol{Y}_{\beta_4} &= \alpha \left(\boldsymbol{X}_{\beta_4-1} + \sum_{j=1}^{t_r+r_r+1} b_{2,j}\boldsymbol{N}_{\beta_4-j} \right) \\ &\quad + \left(\boldsymbol{X}_{\beta_4} + \sum_{j=1}^{t_r+r_r+1} b_{1,j}\boldsymbol{N}_{\beta_4-j} \right) \\ &\quad + \alpha \left(\boldsymbol{X}_{\beta_4+1} + \sum_{j=1}^{t_\ell+r_\ell+1} a_{1,j}\boldsymbol{N}_{\beta_4+1+j} \right) - \boldsymbol{V}_1 \end{aligned} \tag{141}$$

the desired linear combination representing $\boldsymbol{Y}_{\beta_4}$ is obtained by combining the linear combinations on the left-hand sides of Equations (138)–(140) with the genie-information $\boldsymbol{V}_1$.

We next consider $\boldsymbol{Y}_{\beta_4+1}$. By (137),

$$\begin{aligned} \sum_{j=1}^{t_\ell+r_\ell+1} & a_{1,j}\boldsymbol{Y}_{\beta_4+1+j} - a_{2,t_r+r_r+1}\alpha \boldsymbol{X}_{\beta_4+t_\ell+r_\ell+3} \\ & -(a_{2,t_r+r_r} + a_{2,t_r+r_r+1}\alpha)\boldsymbol{X}_{\beta_4+t_\ell+r_\ell+2} \\ &= \boldsymbol{X}_{\beta_4+2} + \sum_{j=1}^{t_\ell+r_\ell+1} a_{2,j}\boldsymbol{N}_{\beta_4+1+j}. \end{aligned} \tag{142}$$

Since the genie-information $\boldsymbol{V}_2$ has been chosen so that

$$\begin{aligned} \boldsymbol{Y}_{\beta_4+1} &= \alpha \left(\boldsymbol{X}_{\beta_4} + \sum_{j=1}^{t_r+r_r+1} b_{1,j}\boldsymbol{N}_{\beta_4-j} \right) \\ &\quad + \left(\boldsymbol{X}_{\beta_4+1} + \sum_{j=1}^{t_\ell+r_\ell+1} a_{1,j}\boldsymbol{N}_{\beta_4+1+j} \right) \\ &\quad + \alpha \left(\boldsymbol{X}_{\beta_4+2} + \sum_{j=1}^{t_\ell+r_\ell+1} a_{2,j}\boldsymbol{N}_{\beta_4+1+j} \right) - \boldsymbol{V}_2 \end{aligned} \tag{143}$$

the desired linear combination representing $\boldsymbol{Y}_{\beta_4+1}$ is obtained by combining the left-hand sides of (139), (140), and (142) with the genie-information $\boldsymbol{V}_2$.

The desired linear combinations representing the outputs $\{\boldsymbol{Y}_{m\beta_4}\}_{m=2}^{\gamma_4-1+\theta_4}$ can be obtained from the equations that result when in (138)–(141) each vector $\boldsymbol{X}_k$, for $k \in \{1, \ldots, K\}$, is replaced by $\boldsymbol{X}_{k+(m-1)\beta_4}$, each vector $\boldsymbol{Y}_k$ by $\boldsymbol{Y}_{k+(m-1)\beta_4}$, each vector $\boldsymbol{N}_k$ by $\boldsymbol{N}_{k+(m-1)\beta_4}$, and the genie-information $\boldsymbol{V}_1$ is replaced by $\boldsymbol{V}_{2m-1}$.

The linear combinations representing the outputs $\{\boldsymbol{Y}_{m\beta_4+1}\}_{m=2}^{\gamma_4-1+\theta_4}$ are obtained from the equations that result when in (139), (140), (142), and (143) the vectors $\boldsymbol{X}_k$, $\boldsymbol{Y}_k$, and $\boldsymbol{N}_k$, for $k \in \{1, \ldots, K\}$, are replaced by the vectors $\boldsymbol{X}_{k+(m-1)\beta_4}$,$\boldsymbol{Y}_{k+(m-1)\beta_4}$, and $\boldsymbol{N}_{k+(m-1)\beta_4}$ and the genie-information $\boldsymbol{V}_2$ is replaced by $\boldsymbol{V}_{2m}$. When $m = 0$ all the out-of-range indices should be ignored, that means, $\boldsymbol{X}_k$, $\boldsymbol{Y}_k$, $\boldsymbol{N}_k$ are assumed to be deterministically 0 for all $k \leq 0$.

Finally, if $\theta_4 = 0$, then the desired linear combination representing $\boldsymbol{Y}_K$ can be obtained by combining the equations that result when in Equations (138), (139), and (141) the vectors $\boldsymbol{X}_k$, $\boldsymbol{Y}_k$, and $\boldsymbol{N}_k$ are replaced by the vectors $\boldsymbol{X}_{K-\beta_4}$,$\boldsymbol{Y}_{K-\beta_4}$, and $\boldsymbol{N}_{K-\beta_4}$ and the genie-information $\boldsymbol{V}_1$ is replaced by $\boldsymbol{V}_{2\gamma_4-1}$. Again, all out-of-range indices should be ingored, i.e., $\boldsymbol{X}_k$, $\boldsymbol{Y}_k$, $\boldsymbol{N}_k$ are assumed to be deterministically 0 for all $k > K$.

Appendix C
Proof That Assumption (26) Holds in Section VIII-B

Notice that for $i \in \{1, \ldots, 2\gamma_5\}$ odd,

$$\mathcal{R}_{\mathcal{B}_i} \backslash (\mathcal{R}_{\mathcal{B}_i} \cap \mathcal{R}_{\mathcal{A}_i}) = \left\{ \frac{(i-1)}{2}\beta_5 + t_r + r_r + 3 \right\}, \tag{144}$$

$$\begin{pmatrix} Y_{t_r+r_r+3,t} \\ Y_{t_r+r_r+4,t} \\ \vdots \\ Y_{\beta_5-1,t} \\ Y_{\beta_5,t} \end{pmatrix} = H_{t_\ell+r_\ell+1}(\alpha) \begin{pmatrix} X_{t_r+r_r+3,t} \\ X_{t_r+r_r+4,t} \\ \vdots \\ X_{\beta_5-1,t} \\ X_{\beta_5,t} \end{pmatrix} + \begin{pmatrix} \alpha X_{t_r+r_r+2,t} \\ 0 \\ \vdots \\ 0 \\ \alpha X_{\beta_5+1,t} \end{pmatrix} + \begin{pmatrix} \boldsymbol{N}_{t_r+r_r+3,t} \\ \boldsymbol{N}_{t_r+r_r+4,t} \\ \vdots \\ \boldsymbol{N}_{\beta_5-1,t} \\ \boldsymbol{N}_{\beta_5} \end{pmatrix} \tag{147}$$

and for i even,

$$\mathcal{R}_{\mathcal{B}_i} \backslash (\mathcal{R}_{\mathcal{B}_i} \cap \mathcal{R}_{\mathcal{A}_i}) = \left\{ \frac{(i-1)}{2}\beta_5 + t_r + r_r + 2 \right\}, \tag{145}$$

and moreover, if $\theta_5 = 1$,

$$\mathcal{R}_{\mathcal{B}_{2\gamma_5+1}} \backslash (\mathcal{R}_{\mathcal{B}_{2\gamma_5+1}} \cap \mathcal{R}_{\mathcal{A}_{2\gamma_5+1}}) = \{K\}. \tag{146}$$

Thus, for $i \le 2\gamma_5 - 1$ odd we need to show that the output sequence $\boldsymbol{Y}_{\frac{(i-1)}{2}\beta_5+t_r+r_r+3}$ can be reconstructed from the messages $\{M_k\}_{k\in\mathcal{A}_i}$, the outputs $\{\boldsymbol{Y}_k\}_{k\in\mathcal{R}_{\mathcal{A}_i}}$, and the genie-information $\{\boldsymbol{V}_m\}_{m=0}^{g}$. Similarly, for i even we need to show that $\boldsymbol{Y}_{\frac{(i-1)}{2}\beta_5+t_r+r_r+2}$ can be reconstructed, and for $i = 2\gamma_5+1$, we need to show that $\boldsymbol{Y}_K$ can be reconstructed.

Using the encoding functions $f_1, \ldots, f_n$, for each i that is odd and satisfies $1 \le i \le 2\gamma_5 - 1$ the inputs $\boldsymbol{X}_{\frac{i-1}{2}\beta_5}$, $\boldsymbol{X}_{\frac{i-1}{2}\beta_5+1}$, and $\boldsymbol{X}_{\frac{i+1}{2}\beta_5}$ can be computed from the messages $\{M_k\}_{k\in\mathcal{A}_i}$. For each i that is even and that satisfies $2 \le i \le 2\gamma_5$ the inputs $\boldsymbol{X}_{\frac{i-1}{2}\beta_5}$, $\boldsymbol{X}_{\frac{i-1}{2}\beta_5+1}$, $\boldsymbol{X}_{\frac{i+1}{2}\beta_5}$, and $\boldsymbol{X}_{\frac{i+1}{2}\beta_5+1}$ can be computed from messages $\{M_k\}_{k\in\mathcal{A}_i}$. Finally, if $\theta_5 = 1$, then inputs $X_{K-t_\ell-r_\ell-2}$ and $X_{K-t_\ell-r_\ell-1}$ can be computed from the messages $\{M_k\}_{k\in\mathcal{A}_{2\gamma_5+1}}$.

We start with $i = 1$ and outputs $\boldsymbol{Y}_{t_r+r_r+3}$. By the channel law (9), the linear system (147) on top of this page holds for every time $t \in \{1, \ldots, n\}$.

Recalling the definition of the parameters $\{d_2, \ldots, d_{t_\ell+r_\ell+1}\}$ in Section VII-B and because $\det\left(\mathsf{H}_{t_\ell+r_\ell+1}(\alpha)\right) = 0$, (147) implies:

$$\begin{aligned} \boldsymbol{Y}_{t_r+r_r+3} = & \sum_{j=2}^{t_\ell+r_\ell+1} d_j \left(\boldsymbol{Y}_{t_r+r_r+2+j} - \boldsymbol{N}_{t_r+r_r+2+j}\right) \\ & -\alpha d_{t_\ell+r_\ell+1}\boldsymbol{X}_{\beta_5+1} + \alpha \boldsymbol{X}_{t_r+r_r+2} \\ & +\boldsymbol{N}_{t_r+r_r+3}. \end{aligned} \tag{148}$$

We next notice that by the channel law (9), for every time $t \in \{1, \ldots, n\}$, the linear system in (149) as shown at the next top of the page holds, where the matrix $\mathsf{M}_{t_r+r_r+1}(\alpha)$ is defined in Section VII-A. Recalling that b_{j_r,j_c} denotes the row-j_r column-j_c entry of the inverse $\mathsf{M}^{\text{inv}}_{t_r+r_r+1}(\alpha)$, Equation (149) implies:

$$\begin{aligned} & \sum_{j_c=1}^{t_r+r_r+1} b_{1,j_c}\boldsymbol{Y}_{t_r+r_r+2-j_c} - (b_{1,t_r+r_r+1} + \alpha b_{1,t_r+r_r})\boldsymbol{X}_1 \\ & = \boldsymbol{X}_{t_r+r_r+2} + \sum_{j_c=1}^{t_r+r_r+1} b_{1,j_c}\boldsymbol{N}_{t_r+r_r+2-j_c}. \end{aligned} \tag{150}$$

Finally, by the definition of the genie-information $\boldsymbol{V}_0$, combining (148) with (150) yields the desired linear combination

$$\begin{aligned} \boldsymbol{Y}_{t_r+r_r+3} = & \sum_{j=2}^{t_\ell+r_\ell+1} d_j\boldsymbol{Y}_{t_r+r_r+2+j} \\ & +\alpha \sum_{j_c=1}^{t_r+r_r+1} b_{1,j_c}\boldsymbol{Y}_{t_r+r_r+2-j_c} \\ & -(b_{1,t_r+r_r+1} + \alpha b_{1,t_r+r_r})\boldsymbol{X}_1 \\ & -\alpha d_{t_\ell+r_\ell+1}\boldsymbol{X}_{\beta_5+1} - \boldsymbol{V}_0. \end{aligned} \tag{151}$$

For each i odd and $3 \le i \le 2\gamma_5 - 1$ the desired linear combination representing $\boldsymbol{Y}_{\frac{i-1}{2}\beta_5+t_r+r_r+3}$ can be found in a similar way. Specifically, using Equations similar to (147)–(151) one can show that

$$\begin{aligned} \boldsymbol{Y}_{\frac{i-1}{2}\beta_5+t_r+r_r+3} = & \sum_{j=2}^{t_\ell+r_\ell+1} d_j\boldsymbol{Y}_{\frac{i-1}{2}\beta_5+t_r+r_r+2+j} \\ & +\alpha \sum_{j_c=1}^{t_r+r_r+1} b_{1,j_c}\boldsymbol{Y}_{\frac{i-1}{2}\beta_5+t_r+r_r+1-j_c} \\ & -(b_{1,t_r+r_r+1} + \alpha b_{1,t_r+r_r})\boldsymbol{X}_{\frac{i-1}{2}\beta_5+1} \\ & -\alpha b_{1,t_r+r_r+1}\boldsymbol{X}_{\frac{i-1}{2}\beta_5} \\ & -\alpha\beta_{t_\ell+r_\ell+1}\boldsymbol{X}_{\frac{i+1}{2}\beta_5+1} - \boldsymbol{V}_{i-1}. \end{aligned} \tag{152}$$

We next consider the case where i is even and $2 \le i \le 2\gamma_5$, where we wish to reconstruct $\boldsymbol{Y}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2}$. The construction of the desired linear combination is similar to Appendix B, that means it is based on equations that are similar to equations (138)–(141). Obviously, (136) remains valid if for each $k \in \{1, \ldots, K\}$ the symbols $X_{k,t}$, $Y_{k,t}$, and $N_{k,t}$ are replaced by $X_{k+(\frac{i}{2}-1)\beta_5-(t_\ell+r_\ell+2),t}$, $Y_{k+(\frac{i}{2}-1)\beta_5-(t_\ell+r_\ell+2),t}$, and $N_{k+(\frac{i}{2}-1)\beta_5-(t_\ell+r_\ell+2),t}$, and therefore similar to (138) and (139) we obtain:

$$\begin{aligned} & \sum_{j=1}^{t_r+r_r+1} b_{2,j}\boldsymbol{Y}_{(\frac{i}{2}-1)\beta_5+t_\ell+r_\ell+2-j} \\ & \qquad -b_{2,t_r+r_r+1}\alpha\boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+1} \\ & \qquad -(b_{2,t_r+r_r}\alpha + b_{2,t_r+r_r+1})\boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+1} \\ & \qquad = \boldsymbol{X}_{t_r+r_r+1} + \sum_{j=1}^{t_r+r_r+1} b_{2,j}\boldsymbol{N}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2-j} \end{aligned} \tag{153}$$

and

$$\begin{aligned} & \sum_{j=1}^{t_r+r_r+1} b_{1,j}\boldsymbol{Y}_{(\frac{i}{2}-1)\beta_5+t_\ell+r_\ell+2-j} \\ & \qquad -b_{1,t_r+r_r+1}\alpha\boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+1} \\ & \qquad -(b_{1,t_r+r_r}\alpha + b_{1,t_r+r_r+1})\boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+1} \end{aligned}$$

$$\begin{pmatrix} Y_{t_r+r_r+1,t} \\ Y_{t_r+r_r,t} \\ \vdots \\ Y_{2,t} \\ Y_{1,t} \end{pmatrix} = \mathsf{M}_{t_r+r_r+1}(\alpha) \begin{pmatrix} X_{t_r+r_r+2,t} \\ X_{t_r+r_r+1,t} \\ \vdots \\ X_{3,t} \\ X_{2,t} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \alpha X_{1,t} \\ X_{1,t} \end{pmatrix} + \begin{pmatrix} N_{t_r+r_r+1,t} \\ N_{t_r+r_r,t} \\ \vdots \\ N_{2,t} \\ N_{1,t} \end{pmatrix} \tag{149}$$

$$= \boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2} + \sum_{j=1}^{t_r+r_r+1} b_{1,j} \boldsymbol{N}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2-j}. \tag{154}$$

Since also (137) remains valid if for each $k \in \mathcal{K}$ the symbols $X_{k,t}$, $Y_{k,t}$, and $N_{k,t}$ are replaced by $X_{k+(\frac{i}{2}-1)\beta_5-(t_\ell+r_\ell+3),t}$, $Y_{k+(\frac{i}{2}-1)\beta_5-(t_\ell+r_\ell+3),t}$, and $N_{k+(\frac{i}{2}-1)\beta_5-(t_\ell+r_\ell+3),t}$, we obtain similarly to (140):

$$\begin{aligned}\sum_{j=1}^{t_\ell+r_\ell+1} & a_{1,j} \boldsymbol{Y}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2+j} \\ & -a_{1,t_r+r_r+1}\alpha \boldsymbol{X}_{(\frac{i}{2}+1)\beta_5+1} \\ & -(a_{1,t_r+r_r} + a_{1,t_r+r_r+1}\alpha)\boldsymbol{X}_{(\frac{i}{2}-1)\beta_5} \\ = & \boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+3} \\ & + \sum_{j=1}^{t_\ell+r_\ell+1} a_{1,j} \boldsymbol{N}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2+j}. \end{aligned} \tag{155}$$

Now, since the genie-information $\boldsymbol{V}_{i-1}$ has been chosen so that Equality (156) on top of the next page holds, the desired linear combination representing $\boldsymbol{Y}_{(\frac{i}{2}-1)\beta_5+t_\ell+r_\ell+2}$ can be obtained by combining (153)–(156).

If $\theta_5 = 1$, then the desired linear combination representing $\boldsymbol{Y}_K$ can be found in a similar manner as in the previous Appendix B. The details are omitted.

APPENDIX D
PROOF OF PROPOSITION 5

Lemma 13: For an integer p and a real number α, denote $u_p(\alpha) = \det\big(\mathsf{H}_p(\alpha)\big)$. Then the following holds.

1) *$u_p(\alpha)$ is a polynomial in α, $u_p(0) = 1$, and it satisfies the following second order recursion:*

$$u_{p+2}(\alpha) = u_{p+1}(\alpha) - \alpha^2 u_p(\alpha), \tag{157}$$

with the initial conditions $u_0(\alpha) = u_1(\alpha) = 1$. We denote by E_p the set of roots of $u_p(\alpha)$.

2) *For $\alpha \neq 0$, define*

$$\upsilon_p(\alpha) \triangleq \frac{u_p(\alpha)}{(-\alpha)^p}.$$

Then $\upsilon_p(\alpha)$ satisfies the second order recursion:

$$\upsilon_{p+2}(\alpha) = -\frac{1}{\alpha}\upsilon_{p+1}(\alpha) - \upsilon_p(\alpha), \tag{158}$$

with the initial conditions $\upsilon_{-1}(\alpha) = 0$ and $\upsilon_0(\alpha) = 1$.

Moreover, for all $p \geq 1$ and $l \geq 0$,

$$\begin{aligned}&\big(\upsilon_l \ \cdots \ \upsilon_{l+p-1}\big)\,\mathsf{H}_p \\ &\quad = \big(-\alpha\upsilon_{l-1} \quad 0 \quad \cdots \quad 0 \quad -\alpha\upsilon_{l+p}\big)\end{aligned} \tag{159}$$

where for simplicity we wrote υ_l for $\upsilon_l(\alpha)$.

Proof: Omitted. ■

We give a proof of Proposition 5 for the case where q is odd. The case q even goes along the same lines. We define $\gamma''' \triangleq (q-1)/2$ and

$$\begin{aligned} L &\triangleq r_\ell + t_\ell \\ \beta''' &\triangleq 2L + 4. \end{aligned}$$

The first part of the proof follows the first part of the proof of the Dynamic-MAC Lemma, see Section IV. We construct a Cognitive MAC as in Section IV using parameters

- $q = 2$;
- $g = 2\gamma'''$;
- $\mathcal{A} = \bigcup_{m=1}^{\gamma'''} \mathcal{A}'''(m)$ where

$$\mathcal{A}'''(0) \triangleq \{r_\ell + 2, \ldots, L + t_r + 2\},$$

 for $1 \leq m \leq \gamma''' - 1$,

$$\mathcal{A}'''(m) \triangleq \{m\beta''' + r_\ell + 1, \ldots, m\beta''' + L + t_r + 2\},$$

 and

$$\mathcal{A}'''(\gamma''') \triangleq \{\gamma'''\beta''' + r_\ell + 1, \ldots, K\};$$

- $\mathcal{B}_1 = \{r_\ell + 1\}$ and $\mathcal{B}_2 = \mathcal{K}\backslash(\mathcal{A} \cup \mathcal{B}_1)$;
- the genie-information

$$\boldsymbol{V}_0 \triangleq -\alpha\upsilon_{L+1}\boldsymbol{X}_{L+1} + \sum_{j=0}^{L} \upsilon_j \boldsymbol{N}_{j+1},$$

 and where the rest of the genie-informations $\{\boldsymbol{V}_i\}_{i=1}^{2\gamma'''}$ is similar to the genie-information described in (113) and (114).

By the choice above,

$$\mathcal{K}\backslash\mathcal{R}_{\mathcal{A}} = \{1\} \cup \big\{m\beta''' - 1, m\beta'''\big\}_{m=1}^{\gamma'''}. \tag{160}$$

Notice that unlike in the proof in Section VII-A, here, part of the genie-information depends on the transmitted signal $\boldsymbol{X}_{L+1}$. (But notice that the signal to noise ratio of $\boldsymbol{X}_{L+1}$ with respect to $\sum_{j=0}^{L} \upsilon_j \boldsymbol{N}_{j+1}$ goes to 0 like $(\alpha - \alpha^*)^\nu$ as α goes to α^*.)

Our choice of parameters satisfies Assumption (26) in the Dynamic-MAC Lemma, and thus we can follow the steps in the proof of (29) to deduce that the capacity region of the original network is included in the capacity region of the Cognitive MAC. That Assumption (26) is satisfied for

$$\begin{aligned} \boldsymbol{Y}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2} = \alpha & \left(\boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+1} + \sum_{j=1}^{t_r+r_r+1} b_{2,j} \boldsymbol{N}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2-j} \right) \\ & + \left(\boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2} + \sum_{j=1}^{t_r+r_r+1} b_{1,j} \boldsymbol{N}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+2-j} \right) \\ & + \alpha \left(\boldsymbol{X}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+3} + \sum_{j=1}^{t_\ell+r_\ell+1} a_{1,j} \boldsymbol{N}_{(\frac{i}{2}-1)\beta_5+t_r+r_r+1+j} \right) - \boldsymbol{V}_{i-1} \end{aligned} \tag{156}$$

$i = 1$ follows because from the messages $\{M_k\}_{k \in \mathcal{A}}$ one can reconstruct $\boldsymbol{X}_{L+2}$, and because by

$$\begin{pmatrix} \boldsymbol{Y}_1 \\ \vdots \\ \boldsymbol{Y}_L \\ \boldsymbol{Y}_{L+1} \end{pmatrix} = H_{L+1} \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \vdots \\ \boldsymbol{X}_{L+1} \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \boldsymbol{X}_{L+2} \end{pmatrix} + \begin{pmatrix} \boldsymbol{N}_1 \\ \vdots \\ \vdots \\ \boldsymbol{N}_{L+1} \end{pmatrix}$$

and by (159), applied to $p = L + 1$ and $l = 0$,

$$\begin{aligned} \sum_{j=0}^{L} \upsilon_j \boldsymbol{Y}_{j+1} &= \alpha \upsilon_L \boldsymbol{X}_{L+2} - \alpha \upsilon_{L+1} \boldsymbol{X}_{L+1} + \sum_{j=0}^{L} \upsilon_j \boldsymbol{N}_{j+1} \\ &= \alpha \upsilon_L \boldsymbol{X}_{L+2} + \boldsymbol{V}_0, \end{aligned}$$

and thus it is possible to reconstruct $\boldsymbol{Y}_1$.

For $i = 2$, Assumption (26) follows by similar considerations as in Appendix B. Appendix B also shows how to choose the genie-signals $\{\boldsymbol{V}_m\}_{m=1}^{2\gamma'''}$.

Let us now bound the sum-capacity of the Cognitive MAC:

$$\begin{aligned} n\mathcal{C}_{\text{MAC},\Sigma} &\leq I\left(\{\boldsymbol{Y}_i\}_{i \in \mathcal{A}'''}, \{\boldsymbol{V}_i\}_{0 \leq i \leq 2\gamma'''}; M_1 \ldots, M_K\right) \\ &= I\left(\{\boldsymbol{Y}_i\}_{i \in \mathcal{A}'''}; M_1, \ldots, M_K \,|\, \{\boldsymbol{V}_i\}_{0 \leq i \leq 2\gamma'''}\right) \\ &\quad + I\left(\{\boldsymbol{V}_i\}_{0 \leq i \leq 2\gamma'''}; M_1, \ldots, M_K\right). \end{aligned}$$

We deal with each term separately.

$$\begin{aligned} & I(\{\boldsymbol{Y}_i\}_{i \in \mathcal{A}}; M_1, \ldots, M_K \,|\, \{\boldsymbol{V}_i\}_{0 \leq i \leq 2\gamma'''}) \\ &\quad \leq \sum_{i \in \mathcal{A}} h(\boldsymbol{Y}_i) - h\left(\{\boldsymbol{N}_i\}_{i \in \mathcal{A}} \middle| \{\boldsymbol{V}_i\}_{1 \leq i \leq 2\gamma'''}, \sum_{j=0}^{L} \upsilon_j \boldsymbol{N}_{j+1}\right) \\ &\quad \leq n(K - 2\gamma''' - 1)\frac{1}{2}\log(P) + nf_1(P, \alpha), \end{aligned}$$

where f_1 is such that $\lim_{\alpha \to \alpha_0} \lim_{P \to \infty} f_1(P, \alpha)$ exists and is finite.

Moreover, as can be verified, the genie-information $\{\boldsymbol{V}_i\}_{1 \leq i \leq 2\gamma'''}$ is independent of $(\boldsymbol{V}_0, M_1, \ldots, M_K)$, and

$$\begin{aligned} & I(\{\boldsymbol{V}_i\}_{0 \leq i \leq 2\gamma'''}; M_1, \ldots, M_K) \\ &\quad = I(\boldsymbol{V}_0; M_1, \ldots, M_K) \\ &\quad \leq n\frac{1}{2}\log\left(1 + \frac{P\alpha^2\upsilon_{L+1}^2(\alpha)}{\left\|\left(\upsilon_0 \cdots \upsilon_L\right)\right\|_2^2}\right) \\ &\quad = n\frac{1}{2}\log\left(P\left|\alpha - \alpha^*\right|^{2\nu}\right) + nf_2(P, \alpha), \end{aligned} \tag{161}$$

where f_2 is such that $\lim_{\alpha \to \alpha_0} \lim_{P \to \infty} f_2(P, \alpha)$ exists and is finite. The last equality follows because for every non-zero α_0, the limit $\lim_{\alpha \to \alpha_0} \left\|\left(\upsilon_0 \cdots \upsilon_L\right)\right\|_2^2$ exists, is finite, and larger than 0, and because by definition α^* is a root of the polynomial $\upsilon_{L+1}^2(\alpha)$ with multiplicity 2ν.

Taking $c_0(\alpha) = \lim_{P \to \infty} (f_1(P, \alpha) + f_2(P, \alpha))$ concludes the proof.

References

[1] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.

[2] S. V. Hanly and P. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommun. Syst.*, vol. 1, no. 1, pp. 1–42, 1993.

[3] A. Lapidoth, S. Shamai (Shitz), and M. A. Wigger, "A linear interference network with local side-information," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Nice, France, Jun. 2007, pp. 2201–2205.

[4] N. Levy and S. Shamai (Shitz), "Clustered local decoding for Wyner-type cellular models," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4976–4985, Nov. 2009.

[5] S. Shamai (Shitz) and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels. I," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1877–1911, Nov. 1997.

[6] S.-J. Kim, S. Jain, and G. B. Giannakis, "Backhaul-constrained multi-cell cooperation using compressive sensing and spectral clustering," in *Proc. SPAWC*, Cesme, Turkey, Jun. 2012, pp. 65–69.

[7] S. Shamai (Shitz) and M. Wigger, "Rate-limited transmitter-cooperation in Wyner's asymmetric interference network," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, St. Petersburg, Russia, Jul./Aug. 2011, pp. 425–429.

[8] F. M. J. Willems, "The discrete memoryless multiple access channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 441–445, May 1983.

[9] I. Maric, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3536–3548, Oct. 2007.

[10] O. Simeone, O. Somekh, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Three-user Gaussian multiple access channel with partially cooperating encoders," in *Proc. Asilomar Conf.*, Oct. 2008, pp. 85–89.

[11] M. A. Wigger and G. Kramer, "Three-user MIMO MACs with cooperation," in *Proc. IEEE Int. Workshop Inf. Theory (ITW)*, Volos, Greece, Jun. 2009, pp. 221–225.

[12] S. I. Bross, A. Lapidoth, and M. Wigger, "Dirty-paper coding for the Gaussian multiaccess channel with conferencing," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5640–5668, Sep. 2012.

[13] A. El Gamal, V. S. Annapureddy, and V. V. Veervalli, "On optimal message assignments for interference channels with CoMP transmission," in *Proc. 46th Annu. Conf. Inf. Sci. Syst.*, Princeton, NJ, USA, Mar. 2012, pp. 1–6.

[14] A. El Gamal, V. S. Annapureddy, and V. V. Veervalli, "Degrees of freedom (DoF) of locally connected interference channels with coordinated multi-point (CoMP) transmission," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Ottawa, ON, Canada, Jun. 2012.

[15] A. El Gamal, V. S. Annapureddy, and V. V. Veervalli, "Degrees of freedom of interference channels with CoMP transmission and reception," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5740–5760, Sep. 2012.

[16] A. El Gamal, V. S. Annapureddy, and V. V. Veervalli, "Interference channels with coordinated multipoint transmission: Degrees of freedom, message assignment, and fractional reuse," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3483–3498, Jun. 2014.

[17] A. Lapidoth, S. Shamai (Shitz), and M. A. Wigger, "On cognitive interference networks," in *Proc. Inf. Theory Workshop (ITW)*, Sep. 2007, pp. 325–330.

[18] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *J. Select. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[19] A. Lozano, R. W. Heath, Jr., and J. G. Andrews. (2012, Mar.). *Fundamental Limits of Cooperation*. [Online]. Available: http://arxiv.org/abs/1204.0011

[20] C. Wang, S. A. Jafar, S. Shamai (Shitz), and M. Wigger, "Interference, cooperation and connectivity—A degrees of freedom perspective," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, St. Petersburg, Russia, Jul./Aug. 2011, pp. 430–434.

[21] Y. Liu and E. Erkip, "On the sum capacity of K-user cascade Gaussian Z-interference channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, St. Petersburg, Russia, Jul./Aug. 2011, pp. 1382–1386.

[22] L. Zhou and W. Yu, "On the capacity of the K-user cyclic Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 154–165, Jan. 2013.

[23] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5543–5562, Dec. 2008.

[24] A. Lapidoth, "On the high-SNR capacity of noncoherent networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3025–3036, Sep. 2005.

[25] I. Bergel, D. Yellin, and S. Shamai, "Linear precoding bounds for Wyner-type cellular networks with limited base-station cooperation and dynamic clustering," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3714–3725, Jul. 2012.

[26] S. Shamai (Shitz), O. Somekh, O. Simeone, A. Sanderovich, B. M. Zaidel, and H. V. Poor, "Cooperative multi-cell networks: Impact of limited-capacity backhaul and inter-users links," in *Proc. Joint Workshop Commun. Coding (JWCC)*, Dürnstein, Austria, Oct. 2007.

[27] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai (Shitz), "Local base station cooperation via finite-capacity links for the uplink of linear cellular networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 190–204, Jan. 2009.

[28] O. Simeone, O. Somekh, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Throughput of cellular systems with conferencing mobiles and cooperative base stations," *EURASIP J. Wireless Commun. Netw.*, vol. 27, Feb. 2008.

[29] R. Zakhaour and S. V. Hanly, "Min-max power allocation in cellular networks with coordinated beamforming," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 287–302, Feb. 2013.

[30] I. Bergel, D. Yellin, and S. Shamai, "Uplink downlink balancing using variable feedback rates," in *Proc. IEEE 27th Conv. Elect. Electron. Eng. Israel (IEEEI)*, Nov. 2012, pp. 1–4.

[31] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Found. Trends Commun. Inf. Theory (FnT)*, vol. 9, nos. 2–3, pp. 117–381, 2012.

[32] O. Simeone *et al.*, "Cooperative wireless cellular systems: An information-theoretic view," *Found. Trends Commun. Inf. Theory (FnT)*, vol. 8, nos. 1–2, pp. 1–177, 2012.

[33] H. Maleki, V. Cadambe, and S. Jafar, "Index coding: An interference alignment perspective," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2012, pp. 2236–2240.

[34] S. A. Jafar, "Blind interference alignment," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 3, pp. 216–227, Jun. 2012.

[35] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[36] S. A. Jafar and V. R. Cadambe, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[37] S. A. Jafar, "Interference alignment—A new look at signal dimensions in a communication network," *Found. Trends Commun. Inf. Theory*, vol. 7, no. 1, pp. 1–134, 2011.

[38] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. 27, no. 6, pp. 786–788, Nov. 1981.

[39] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 581–586, Mar. 2004.

[40] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[41] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–598, Nov./Dec. 1999.

Amos Lapidoth (S'89–M'95–SM'00–F'04) received the B.A. degree in mathematics (*summa cum laude*, 1986), the B.Sc. degree in electrical engineering (*summa cum laude*, 1986), and the M.Sc. degree in electrical engineering (1990) all from the Technion—Israel Institute of Technology. He received the Ph.D. degree in electrical engineering from Stanford University in 1995.

In the years 1995–1999 he was an Assistant and Associate Professor at the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, and was the KDD Career Development Associate Professor in Communications and Technology. He is now Professor of Information Theory at ETH Zurich in Switzerland. He is the author of the book *A Foundation in Digital Communication* (Cambridge University Press, 2009). His research interests are in digital communications and information theory.

Dr. Lapidoth served in the years 2003–2004 and 2009 as Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Nathan Levy received the B.A. degree in mathematics from the Ecole Normale Supérieure, Paris, France, in 2005, the M.A. degree in mathematics from the Université Paris-Sud, Paris, France, in 2006, and the Ph.D. degree in electrical engineering from the Technion—Israel Institute of Technology, Haifa, Israel.

His research interests are random matrices theory and information theory.

Shlomo Shamai (Shitz) (F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975–1985 he was with the Communications Research Labs, in the capacity of a Senior Research Engineer. Since 1986, he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now a Technion Distinguished Professor, and holds the William Fondiller Chair of Telecommunications. His research interests encompass a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is a member of the Israeli Academy of Sciences and Humanities and a foreign member of the U.S. National Academy of Engineering. He is the recipient of the 2011 Claude E. Shannon Award and the 2014 Rothschild Prize in Mathematics/Computer Sciences and Engineering.

He has been awarded the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and is a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++) Best Paper Award, the 2010 Thomson Reuters Award for International Excellence in Scientific Research, and the 2014 EURASIP Best Paper Award (for the *EURASIP Journal on Wireless Communications and Networking*). He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served twice on the Board of Governors of the Information Theory Society. He has served on the Executive Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY.

Michèle Wigger (S'05–M'09) received the M.Sc. degree in electrical engineering (with distinction) and the Ph.D. degree in electrical engineering both from ETH Zurich in 2003 and 2008, respectively. In 2009 she was a postdoctoral researcher at the ITA center at the University of California, San Diego. Since December 2009 she is an Assistant Professor at Telecom ParisTech, in Paris, France. Her research interests are in information and communications theory.