On Capacity Scaling of Multi-Antenna Multi-Hop Networks: The Significance of the Relaying Strategy in the "Long Network Limit"

Jörg Wagner, *Student Member, IEEE*, and Armin Wittneben, *Member, IEEE*

***Abstract*—The sum-capacity C of a static uplink channel with n single-antenna sources and an n-antenna destination is known to scale linearly in n, if the random channel matrix fulfills the conditions for the Marcenko-Pastur law: if each source transmits at power P/n, there exists a positive c_0, such that $\lim_{n \longrightarrow \infty} C/n = c_0$ almost surely. This paper addresses the question to which extent this result carries over to multi-hop networks. Specifically, an $L+1$-hop network with n non-cooperative source antennas, n fully cooperative destination antennas, and L relay stages of $n_{\mathcal{R}}$ (cooperative or non-cooperative) relay antennas each is considered. Four relaying strategies are assessed based on the interrelationship between two sequences. For each considered strategy XF, there exists a sequence $(c_L^{\mathrm{XF}})_{L=0}^{\infty}$, such that $c_L^{\mathrm{XF}} = \lim_{n \longrightarrow \infty} R_L^{\mathrm{XF}}/n$ almost surely, where R_L^{XF} denotes the supremum of the set of sum-rates that are achievable by the strategy over $L+1$ hops. This sequence depends on the sequence $(P_L)_{L=0}^{\infty}$, where P_L corresponds to the power of the source stage and each of the relay stages in an $L+1$-hop network. Results are summarized as follows:**

- ***Decode & forward (DF)*: For $n_{\mathcal{R}} = n$, c_L^{DF} is constant with respect to L, if also P_L is constant with respect to L.**
- ***Quantize & forward with (CF) and without (QF) Slepian & Wolf compression*: For $n_{\mathcal{R}} = n$, there exists a sequence $(P_L)_{L=0}^{\infty}$, such that $c_L^{\mathrm{QF/CF}}$ is positive and constant with respect to L. The corresponding sequence $(P_L)_{L=0}^{\infty}$ grows linearly with L for CF and exponentially with L for QF.**
- ***Amplify & forward (AF)*: Fix $n_{\mathcal{R}}/n = \beta$ and $P_L \propto L$. Then, (i) there exists $c > 0$, such that $\lim_{L \longrightarrow \infty} c_L^{\mathrm{AF}} = c$, if $\beta \in \Omega\left(L^{1+\varepsilon}\right)$ and (ii) $\lim_{L \longrightarrow \infty} c_L^{\mathrm{AF}} = 0$, if $\beta \in \mathcal{O}\left(L^{1-\varepsilon}\right)$.**

***Index Terms*—Amplify & forward, capacity scaling, compress & forward, decode & forward, MIMO uplink, multi-hop, multiple-input multiple-output, quantize & forward, random matrix theory, Slepian and Wolf compression.**

I. Introduction

CONSIDER wireless transmission from n transmit antennas to an n-antenna receive terminal over a random and static frequency-flat channel. Assume that each transmit antenna transmits with power P/n, where P corresponds to the sum-transmit power. If the channel coefficients between all pairs of transmit and receive antennas are identically and independently distributed (i.i.d.) random variables with zero mean and nonzero variance, (sum-)capacity scales linearly in n almost surely, irrespective of whether transmit antennas can cooperate or not [1]. More precisely, the capacity of an $n \times n$ point-to-point multiple-input multiple-output channel with white transmit covariance matrix, coincides with the sum-capacity of an uplink channel with n single-antenna transmit terminals and an n-antenna receive terminal [2].

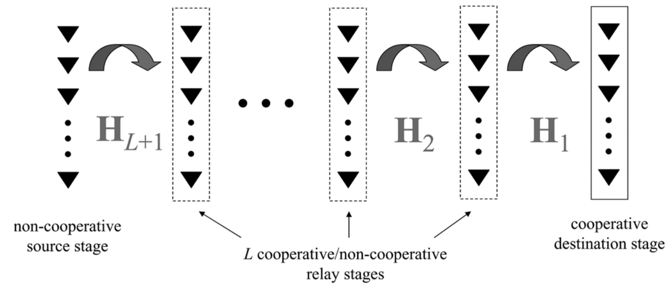

Fig. 1. $n_{\mathcal{S}}$ non-cooperating source antennas transmit to a destination terminal with $n_{\mathcal{D}}$ antennas via L stages of $n_{\mathcal{R}}$ relay antennas. Relay antennas are cooperative or non-cooperative depending on the relaying strategy.

Envision the scenario that the transmit antennas are shadowed from the receive terminal. Wireless connectivity can then be sustained through the installation of properly positioned intermediate nodes that relay the source signals to the destination via multiple hops (see Fig. 1). If the number of antennas in each relay stage, $n_{\mathcal{R}}$, grows linearly with n, then also the corresponding sum-capacity scales linearly in n for any fixed number of relay stages, L. This result holds true even for non-cooperative relay stages[1] [3]. Thus, as a generalization of the result that transmit antenna cooperation is not crucial in multi-antenna single-hop networks, *neither source nor relay antenna cooperation is crucial for linear sum-capacity scaling in multi-antenna multi-hop networks*.

The above statement says nothing about the asymptotic constant of proportionality c_L^{XF} that fulfills almost surely

$$c_L^{\mathrm{XF}} = \lim_{n \longrightarrow \infty} \frac{R_L^{\mathrm{XF}}}{n} \tag{1}$$

where R_L^{XF} denotes the supremum of the set of sum-rates that are achievable through a certain relaying scheme "XF" in an $L+1$-hop network, except for the fact that it is strictly positive

[1]The term non-cooperative relay stage refers in this paper to the scenario that joint processing of receive signals of antennas within a stage is disabled, i.e., each relay antenna is associated with a single-antenna node. Likewise, a cooperative relay stage can be viewed as a single multi-antenna node.

Manuscript received May 27, 2010; revised October 09, 2010; accepted June 15, 2011.

J. Wagner was with the Communication Technology Laboratory, ETH Zurich, Sternwartstrasse 7, 8092 Zurich, Switzerland. He is now with Swisscom AG, Alte Tiefenaustrasse 6, 3050 Bern, Switzerland (e-mail: joerg.wagner@swisscom.com).

A. Wittneben is with the Communication Technology Laboratory, ETH Zurich, Sternwartstrasse 7, 8092 Zurich, Switzerland (e-mail: wittneben@nari.ee.ethz.ch).

Communicated by L. Zheng, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2011.2177752

for every L. In particular, it does not preclude the scenario that $c_L^{\mathrm{XF}} \to 0$ as $L \to \infty$. Accordingly, the central question of this paper is, how c_L^{XF} evolves[2] for increasing L under various relaying strategies. The provided answers reveal fundamental differences in this asymptotic behavior with respect to L, not only between cooperative and non-cooperative relaying strategies, but also between different non-cooperative strategies. Specifically, the following four relaying strategies are investigated.

- *Decode & forward*: For this strategy, full cooperation among antennas within relay stages is assumed. Relay stages can then decode messages from their preceding stage efficiently, and re-encode them prior to the forwarding. Thus, messages are regenerated in every relay stage and propagate from stage to stage until they reach the destination. This strategy is optimal in multi-hop networks according to the data processing inequality [4].
- *Pure quantize & forward (QF)*: This strategy can be executed in a completely decentralized fashion, i.e., without any relay antenna cooperation. The receive signal of each relay antenna is quantized. The index of the quantization codeword is then encoded and forwarded. For decoding, the destination recursively recovers the quantized relay receive signals of each stage, until it can decode the source messages based on the quantized receive signals of the first relay stage.
- *Quantize & forward with Slepian & Wolf compression (CF)*: The above quantize & forward strategy QF can be enhanced through Slepian and Wolf compression [5] in each relay stage. This additional step exploits the spatial correlation among relay receive signals, but requires dissemination of channel state information within relay stages.
- *Amplify & forward*: This strategy operates in a completely decentralized fashion, too. The receive signal of each relay antenna is amplified by a constant gain factor prior to transmission to the succeeding relay or destination stage. Thus, end-to-end transmission from source to destination stage occurs over an equivalent single-hop channel. The strategy is particularly appealing due to its simplicity and low complexity.

In order to ensure a fair comparison, the sequence $(c_L^{\mathrm{XF}})_{L=0}^{\infty}$ must be considered together with a second sequence $(P_L)_{L=0}^{\infty}$, whose $L\mathrm{th}$ element corresponds to the per-stage transmit power that is allocated to the source antenna stage and each of the relay stages in an $L+1$-hop network.

The results of this paper are outlined as follows:

- *Decode & forward*: For $n_{\mathcal{R}} = n$, c_L^{DF} is constant with respect to L, if also P_L is constant with respect to L.
- *Quantize & forward with Slepian & Wolf compression*: For $n_{\mathcal{R}} = n$, there exists a linearly increasing sequence $(P_L)_{L=0}^{\infty}$, such that c_L^{CF} is constant with respect to L. This is the best among the non-cooperative strategies.
- *Pure quantize & forward*: For $n_{\mathcal{R}} = n$, there exists an exponentially increasing sequence $(P_L)_{L=0}^{\infty}$, such that c_L^{QF} is constant with respect to L.
- *Amplify & forward*: The convergence $c_L^{\mathrm{AF}} \to 0$ as $L \to \infty$ is avoided for a linear growth of P_L with L, if and only if also the ratio $\beta = n_{\mathcal{R}}/n$ grows at least linearly with L.

Remarks: It is not surprising, that a constant P_L does not suffice for a constant c_L^{XF}, if a non-cooperative strategy is applied. This is an immediate consequence of the inherent noise accumulation, which needs to be compensated by an increased transmit power. Moreover, the obtained results do not allow for a fair comparison between amplify & forward and pure quantize & forward. No statement is made about c_L^{AF} for $n_{\mathcal{R}} = n$ and an exponential growth of P_L. Vice versa, no statement is made on how c_L^{QF} can benefit from more than n relay nodes per stage.

Related Work: Multiple-input multiple-output communication over multiple hops has received a lot of attention recently. [3] establishes linear sum-capacity scaling in n (the number of antennas per stage) of amplify & forward multi-hop networks for finite numbers of hops. This work is a generalization of a result on two-hop networks in [6]. [7] considers a system that corresponds to an amplify & forward multi-hop network with noiseless relays. For this network, it is shown that the asymptotic constant of proportionality between sum-capacity and the number of nodes per stage (assumed not to grow with the number of hops) tends to zero as the number of hops grows large. [8] studies the same setting for a finite number of hops with correlated fading and cooperative relay antennas that not only amplify, but also linearly combine the receive signals within a stage. In contrast to the above works, [9] studies multi-hop networks for finite n in terms of the achievable degrees of freedom in the limit of infinitely many hops. Moreover, the works [10] (decode & forward) and [11] (amplify & forward) study the diversity-multiplexing tradeoff of multi-antenna multi-hop networks. Multi-antenna multi-hop systems with non-cooperative destination antennas have been studied with respect to achievable degrees of freedom in [12] and [13].

Notation: For the matrix $\mathbf{A}$, $\det(\mathbf{A})$, $\mathbf{A}^H$, $\mathrm{Tr}(\mathbf{A})$ and $\|\mathbf{A}\|_*$ stand for its determinant, Hermitian transpose, trace and nuclear norm. Moreover, $\lambda_i\{\mathbf{A}\}$ and $\sigma_i\{\mathbf{A}\}$ denote its $i\mathrm{th}$ eigenvalue and the $i\mathrm{th}$ singular value (in descending order). Throughout the paper, all logarithms, unless specified otherwise, are to the base 2. For the vector $\mathbf{a}$, $\|\mathbf{a}\|$ denotes its Euclidean norm and $a(i)$ its $i\mathrm{th}$ element. The probability of the event A is denoted by $\mathsf{Pr}[A]$ and the expectation with respect to the random variable X by E_X. $\mathcal{A}^C$ denotes the complement of set $\mathcal{A}$. Furthermore, the standard $\mathcal{O}(\cdot), \Omega(\cdot), \Theta(\cdot)$ notation is used for the characterization the asymptotic behavior of some function $f(\cdot)$ according to

$$f(n) \in \mathcal{O}(g(n)), \text{ if } \exists M,\ n_0 > 0{:}M|g(n)| > |f(n)|\ \forall n \geq n_0$$
$$f(n) \in \Omega(g(n)), \text{ if } \exists M,\ n_0 > 0{:}M|g(n)| < |f(n)|\ \forall n > n_0$$
$$f(n) \in \Theta(g(n)), \text{ if } f(n) \in \mathcal{O}(g(n)) \text{ and } f(n) \in \Omega(g(n)).$$

Finally, the function $1\{x\}$ is defined to be 1, if x is true, and zero otherwise. The functions $\delta(x)$ and $\sigma(x)$ denote Dirac delta and Heaviside step, respectively.

Outline: Section II introduces the signal model and the applied multi-hop communication protocol. Section III presents in detail the relaying strategies that are considered. Section IV constitutes the core of this work and provides four Theorems that

[2]In this paper, the focus is not on the pre-log factor that is obviously incurred, if the source stage does not inject new signals into the network in every time slot. Accordingly, it is not taken into account in R_L^{XF} and set to one.

characterize the scaling of the supremum of the set of achievable rates under each relaying strategy.

II. Signal Model and Communication Protocol

A stage of $n_{\mathcal{S}}$ single-antenna source nodes, $\mathcal{S}$, aims to transmit data to a destination node that has access to the receive signals of a cluster $\mathcal{D}$ of $n_{\mathcal{D}}$ antennas. Communication is assisted by L relay stages of $n_{\mathcal{R}}$ antennas each (see Fig. 1). Two cases are distinguished:

- *Fully cooperative relay stages*: All relay antennas in a stage are connected to a central node. This central node has access to the receive signals of all antennas in the stage, and, based on this knowledge, determines the transmit signals of the relay antennas.
- *Non-cooperative relay stages*: Each relay antenna corresponds to a single node that has to determine its transmit signal solely based on the knowledge of its own receive signal.

The relay stages are labeled by $\mathcal{R}_l,\ l \in \{1,\ldots,L\}$. Moreover, the $k\mathrm{th}$ antenna in a source, relay or destination stage is labeled by $\mathrm{S}^{(k)}$, $\mathrm{R}_l^{(k)}$ and $\mathrm{D}^{(k)}$, respectively. The network employs a multi-hop protocol. More precisely, source signals traverse all relay stages in descending order of indexes, i.e., $\mathcal{R}_L, \mathcal{R}_{L-1}, \ldots, \mathcal{R}_1$, before they are received by the destination stage. Transmission is divided into $L+1$ time slots, one dedicated to each hop, of N symbol durations each. That is, transmissions of different stages are orthogonal in time. Specifically,

- stage $\mathcal{S}$ transmits to stage $\mathcal{R}_L$ in time slot $j=1$,
- stage $\mathcal{R}_{L-j+2}$ transmits to stage $\mathcal{R}_{L-j+1}$ in time slot j, where $j \in \{2,\ldots,L\}$,
- stage $\mathcal{R}_1$ transmits to stage $\mathcal{D}$ in time slot $j = L+1$.

Channels between any two antennas are quasi-static and frequency-flat over the bandwidth of interest. The channel coefficient that corresponds to the link between transmit node A and receive node B is denoted by h_{BA}. With this notation, the channel matrices are written as

$$\mathbf{H}_l = \begin{cases} \left(h_{\mathrm{R}_L^{(k)}\mathrm{S}^{(k')}}\right)_{k=1,\ldots,n_{\mathcal{R}},k'=1,\ldots,n_{\mathcal{S}}}, & \text{if } l = L+1 \\ \left(h_{\mathrm{R}_l^{(k)}\mathrm{R}_{l-1}^{(k')}}\right)_{k=1,\ldots,n_{\mathcal{R}},k'=1,\ldots,n_{\mathrm{R}}}, & \text{if } l \in \{2,\ldots,L\} \\ \left(h_{\mathrm{D}^{(k)}\mathrm{R}_1^{(k')}}\right)_{k=1,\ldots,n_{\mathcal{D}},k'=1,\ldots,n_{\mathcal{R}}}, & \text{if } l = 1. \end{cases} \tag{2}$$

Furthermore, the sequence of signals that is transmitted by antenna A in its dedicated transmit time slot is denoted by $(x_{\mathrm{A}}^{(i)})_{i=1}^N$. The vector of transmit signals of antennas within a stage $\mathcal{A} = \{\mathrm{A}^{(1)},\ldots,\mathrm{A}^{(|\mathcal{A}|)}\}$ is denoted by

$$\mathbf{x}_{\mathcal{A}}^{(i)} = \left[x_{\mathrm{A}^{(1)}}^{(i)},\ldots,x_{\mathrm{A}^{(|\mathcal{A}|)}}^{(i)}\right]^T. \tag{3}$$

Likewise, the sequences of receive and additive noise signals, as they are observed by antenna B in its dedicated receive time slot, are denoted by $(y_{\mathrm{B}}^{(i)})_{i=1}^N$ and $(w_{\mathrm{B}}^{(i)})_{i=1}^N$, respectively. The receive signal and noise vectors of a stage $\mathcal{B} = \{\mathrm{B}^{(1)},\ldots,\mathrm{B}^{|\mathcal{B}|}\}$ are denoted by

$$\mathbf{y}_{\mathcal{B}}^{(i)} = \left[y_{\mathrm{B}^{(1)}}^{(i)},\ldots,y_{\mathrm{B}^{(|\mathcal{B}|)}}^{(i)}\right]^T \tag{4}$$

$$\mathbf{w}_{\mathcal{B}}^{(i)} = \left[w_{\mathrm{B}^{(1)}}^{(i)},\ldots,w_{\mathrm{B}^{(|\mathcal{B}|)}}^{(i)}\right]^T. \tag{5}$$

The transmission in time slot j is thus described by the input-output (I-O) relation

$$\mathbf{y}_{\mathcal{R}_L}^{(i)} = \mathbf{H}_{L+1}\mathbf{x}_{\mathcal{S}}^{(i)} + \mathbf{w}_{\mathcal{R}_L}^{(i)}, \quad \text{if } j=1 \tag{6}$$

$$\mathbf{y}_{\mathcal{R}_{L-j+1}}^{(i)} = \mathbf{H}_{L-j+2}\mathbf{x}_{\mathcal{R}_{L-j+2}}^{(i)} + \mathbf{w}_{\mathcal{R}_{L-j+1}}^{(i)}, \text{ if } j \in \{2,\ldots,L\} \tag{7}$$

$$\mathbf{y}_{\mathcal{D}}^{(i)} = \mathbf{H}_1\mathbf{x}_{\mathcal{R}_1}^{(i)} + \mathbf{w}_{\mathcal{D}}^{(i)}, \quad \text{if } j = L+1. \tag{8}$$

Note that transmit antennas within the same stage are assumed to be symbol-synchronized.

The elements of the vectors $\mathbf{w}_{\mathcal{R}_l}^{(i)}$, $l \in \{1,\ldots,L\}$, and $\mathbf{w}_{\mathcal{D}}^{(i)}$ represent the thermal noise that is introduced at the receiver front-ends. They are i.i.d. (both in space and time) circularly symmetric complex Gaussian random variables of variance σ_{w}^2. The elements of the channel matrices $\mathbf{H}_l,\ l \in \{1,\ldots,L+1\}$, are i.i.d. random variables that follow an arbitrary distribution with zero-mean, unit variance and bounded fourth moment. The transmit signals of all antennas in source and relay stages are subject to an average power constraint[3]. Specifically, source and relay signals must fulfill almost surely

$$P_{\mathrm{S}^{(k)}} \triangleq \lim_{N\longrightarrow\infty}\frac{1}{N}\sum_{i=1}^N\left|x_{\mathrm{S}^{(k)}}^{(i)}\right|^2 \le \frac{P_L}{n_{\mathcal{S}}}\ \forall k \in \{1,\ldots,n_{\mathcal{S}}\} \tag{9}$$

$$P_{\mathrm{R}_l^{(k)}} \triangleq \lim_{N\longrightarrow\infty}\frac{1}{N}\sum_{i=1}^N\left|x_{\mathrm{R}_l^{(k)}}^{(i)}\right|^2 \le \frac{P_L}{n_{\mathcal{R}}}\ \forall k \in \{1,\ldots,n_{\mathcal{R}}\} \tag{10}$$

where P_L corresponds to the sum-power of each stage. The parameter P_L will be treated as a sequence in L.

It remains to specify, how each relay stage determines its transmit signals from its receive signals, i.e., a map

$$g_{\mathcal{R}_l} : \mathbb{C}^{N\times n_{\mathcal{R}}} \longrightarrow \mathcal{C}_{\mathcal{R}_l} : \left(\mathbf{y}_{\mathcal{R}_l}^{(i)}\right)_{i=1}^N \longrightarrow \left(\mathbf{x}_{\mathcal{R}_l}^{(i)}\right)_{i=1}^N \tag{11}$$

where $\mathcal{C}_{\mathcal{R}_l}$ denotes the set of transmit signal vector sequences of relay stage $\mathcal{R}_l$. In the case of non-cooperative relay stages, the transmit signal sequence of each antenna must solely depend on the corresponding receive signal sequence of the antenna. That is, the map $g_{\mathcal{R}_l}$ decouples into the following maps:

$$g_{\mathrm{R}_l^{(k)}} : \mathbb{C}^N \longrightarrow \mathcal{C}_{\mathrm{R}_l^{(k)}} : \left(y_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{i=1}^N \longrightarrow \left(x_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{i=1}^N, \quad k \in \{1,\ldots,n_{\mathcal{R}}\} \tag{12}$$

where $\mathcal{C}_{\mathrm{R}_l^{(k)}}$ denotes the set of transmit signals of relay $\mathrm{R}_l^{(k)}$. Three fundamentally different relaying techniques are investigated: (i) decode & forward, (ii) quantize & forward (with an optional Slepian and Wolf compression), and (iii) amplify & forward. While the decode & forward strategy is optimal in terms of sum-capacity as an immediate consequence of the data processing inequality [4], it requires fully cooperative relay stages. The other forwarding strategies are known to be suboptimal, but do not require cooperative relay stages. For all relaying strategies, transmitting nodes are assumed not to possess any transmit channel state information. The amount of receive channel state

[3]This power constraint will be slightly relaxed in the case of amplify & forward relaying (see Section III-C).

TABLE I
RECEIVE CHANNEL STATE INFORMATION

	Relay Stages	Destination Node
DF	preceding hop channel matrix	preceding hop channel matrix
CF	receive signal powers	all channel matrices
QF	receive signal powers	all channel matrices
AF	no requirements	all channel matrices

TABLE II
RATE FEEDBACK

	Source Nodes	Relay Stages
DF	channel code-rates	channel code-rates
CF	channel code-rates	channel code- and compression-rates
QF	channel code-rates	channel code-rates
AF	channel code-rates	no requirements

information in the relay stages and at the destination node is specified in Table I for each scheme individually. Moreover, transmitting nodes require certain rate feedback, which is provided through perfect feedback links from the destination. This rate feedback is summarized in Table II for each scheme.

The three relaying techniques and corresponding achievable rates are revisited in the context of the considered multi-hop network in the next section.

III. Multi-Hop Relaying Techniques

In the following, each source node $\mathrm{S}^{(k)}$ chooses randomly according to a uniform distribution a message $m_{\mathrm{S}^{(k)}}$ out of the message set $\mathcal{M}_{\mathrm{S}^{(k)}}$ with $2^{NR_{\mathrm{S}^{(k)}}}$ messages for transmission. The channel codebook of node $\mathrm{S}^{(k)}$ has rate $R_{\mathrm{S}^{(k)}}$ and is denoted by $\mathcal{C}_{\mathrm{S}^{(k)}}$. Furthermore, for each node $\mathrm{S}^{(k)}$ an encoding function is defined as

$$f_{\mathrm{S}^{(k)}}: \mathcal{M}_{\mathrm{S}^{(k)}} \longrightarrow \mathcal{C}_{\mathrm{S}^{(k)}} : m_{\mathrm{S}^{(k)}} \longrightarrow \left(x^{(i)}_{\mathrm{S}^{(k)}}\right)_{i=1}^{N}. \quad (13)$$

A. Decode & Forward Relaying

Although decode & forward relaying can be applied to networks with arbitrary $n_{\mathcal{S}}, n_{\mathcal{R}}$ and $n_{\mathcal{D}}$, attention is restricted to the case $n_{\mathcal{S}} = n_{\mathcal{R}} = n_{\mathcal{D}} \triangleq n$ in this work. Decode & forward relaying refers to the technique of decoding and re-encoding messages in each relay stage. In order to enable coherent decoding in all relay stages and at the destination node, each relay stage as well as the destination node is assumed to possess perfect channel state information of its preceding hop. The maps (11) can be split into two parts each, i.e., $g_{\mathcal{R}_l} = \hat{g}_{\mathcal{R}_l} \circ \tilde{g}_{\mathcal{R}_l}$. First, the $n_{\mathcal{S}}$ messages—one corresponding to each source node—as transmitted by the preceding stage are decoded based on the observed sequence $\left(\mathbf{y}^{(i)}_{\mathcal{R}_l}\right)_{i=1}^{N}$ and the knowledge of the channel matrix $\mathbf{H}_{l+1}$. The corresponding decoding function is

$$\tilde{g}_{\mathcal{R}_l}: \mathbb{C}^{N\times n_{\mathcal{R}}} \longrightarrow \mathcal{M}_{\mathrm{S}^{(1)}} \times \ldots \times \mathcal{M}_{\mathrm{S}^{(n_{\mathcal{S}})}} : \left(\mathbf{y}^{(i)}_{\mathcal{R}_l}\right)_{i=1}^{N} \longrightarrow \left(\hat{m}^{(l)}_{\mathrm{S}^{(1)}}, \ldots, \hat{m}^{(l)}_{\mathrm{S}^{(n_{\mathcal{S}})}}\right). \quad (14)$$

Then, the decoded messages are re-encoded as

$$\hat{g}_{\mathcal{R}_l}: \mathcal{M}_{\mathrm{S}^{(1)}} \times \ldots \times \mathcal{M}_{\mathrm{S}^{(n_{\mathcal{S}})}} \longrightarrow \mathcal{C}_{\mathcal{R}_l} : \left(\hat{m}^{(l)}_{\mathrm{S}^{(1)}}, \ldots, \hat{m}^{(l)}_{\mathrm{S}^{(n_{\mathcal{S}})}}\right) \longrightarrow \left(\mathbf{x}^{(i)}_{\mathcal{R}_l}\right)_{i=1}^{N} \quad (15)$$

where $\mathcal{C}_{\mathcal{R}_l}$ denotes the channel codebook of relay stage $\mathcal{R}_l$. The rate of this channel codebook is given by $R_l = \sum_{k=1}^{n} R_{\mathrm{S}^{(k)}} \triangleq R^{\Sigma}_{\mathcal{S}}$ for all $l \in \{1, \ldots, L\}$.

A sum-rate $R^{\Sigma}_{\mathcal{S}}$ is achievable with the decode & forward strategy, if it is supported by each individual hop. For the first hop between the source stage $\mathcal{S}$ and the fully cooperative relay stage $\mathcal{R}_L$, transmission occurs over an equivalent uplink channel with n single-antenna transmit terminals and an n-antenna receive terminal. A sum-rate $R^{\Sigma}_{\mathcal{S}}$ is achievable over the first hop channel [2], if

$$R^{\Sigma}_{\mathcal{S}} < \log\det\left(\mathbf{I}_n + \frac{P_L}{n\cdot\sigma^2_{\mathrm{w}}}\mathbf{H}_{L+1}\mathbf{H}^H_{L+1}\right). \quad (16)$$

Note that this work focuses on the achievable sum-rate $R^{\Sigma}_{\mathcal{S}}$ and does not care about the individual rates $R_{\mathrm{S}^{(k)}}$. For all following hops, transmission occurs over equivalent point-to-point channels, since both transmit and receive stage are fully cooperative. Relay stage $\mathcal{R}_l, l \in \{1, \ldots, L\}$, can transmit reliably to its succeeding stage at a rate R_l, if [1]

$$R_l < \log\det\left(\mathbf{I}_n + \frac{P_L}{n\cdot\sigma^2_{\mathrm{w}}}\mathbf{H}_l\mathbf{H}^H_l\right). \quad (17)$$

Here, a spatially white transmit covariance matrix is used due to the absence of transmit channel state information. In order to achieve $R^{\Sigma}_{\mathcal{S}}$ over the end-to-end channel, (16) and(17) need to be fulfilled for $R^{\Sigma}_{\mathcal{S}} = R_L = \cdots = R_1$. This corresponds to the condition

$$R^{\Sigma}_{\mathcal{S}} < \min_{l\in\{1,\ldots,L+1\}} \log\det\left(\mathbf{I}_n + \frac{P_L}{n\cdot\sigma^2_{\mathrm{w}}}\mathbf{H}_l\mathbf{H}^H_l\right) \triangleq R^{\mathrm{DF}}_L. \quad (18)$$

B. Quantization-Based Relaying

Although quantization-based relaying can be applied to networks with arbitrary $n_{\mathcal{S}}$, $n_{\mathcal{R}}$ and $n_{\mathcal{D}}$, attention is restricted to the case $n_{\mathcal{S}} = n_{\mathcal{R}} = n_{\mathcal{D}} \triangleq n$ in this work. In this non-cooperative relaying scheme the maps $g_{\mathrm{R}^{(k)}_l}$, $k \in \{1, \ldots, n\}$, as defined in (12), are implemented in two steps, i.e., $g_{\mathrm{R}^{(k)}_l} = \hat{g}_{\mathrm{R}^{(k)}_l} \circ \tilde{g}_{\mathrm{R}^{(k)}_l}$. In a first step, each relay $\mathrm{R}^{(k)}_l$ quantizes its receive sequence $\left(y^{(i)}_{\mathrm{R}^{(k)}_l}\right)_{i=1}^{N}$. The quantized sequence is denoted by $\left(\hat{y}^{(i)}_{\mathrm{R}^{(k)}_l}\right)_{i=1}^{N} \in \tilde{\mathcal{C}}_{\mathrm{R}^{(k)}_l}$, where $\tilde{\mathcal{C}}_{\mathrm{R}^{(k)}_l}$ denotes the quantization codebook of the relay. The sequence of additive quantization noise is defined as $\left(q_{\mathrm{R}^{(k)}_l}\right)_{i=1}^{N} \triangleq \left(\hat{y}^{(i)}_{\mathrm{R}^{(k)}_l} - y^{(i)}_{\mathrm{R}^{(k)}_l}\right)_{i=1}^{N}$, the quantization noise vector of stage $\mathcal{R}_l$ is denoted by

$$\mathbf{q}^{(i)}_{\mathcal{R}_l} = \left[q^{(i)}_{\mathrm{R}^{(1)}_l}, \ldots, q^{(i)}_{\mathrm{R}^{(n)}_l}\right]^T. \quad (19)$$

The corresponding map is

$$\tilde{g}^{(k)}_l: \mathbb{C}^N \longrightarrow \tilde{\mathcal{C}}_{\mathrm{R}^{(k)}_l} : \left(y^{(i)}_{\mathrm{R}^{(k)}_l}\right)_{i=1}^{N} \longrightarrow \left(\hat{y}^{(i)}_{\mathrm{R}^{(k)}_l}\right)_{i=1}^{N} \quad (20)$$

where the quantization codebook $\tilde{\mathcal{C}}_{\mathrm{R}^{(k)}_l}$ has rate $\tilde{R}^{(k)}_l$.

In a second step, the index of the quantized sequence is encoded by the channel coder, i.e., mapped onto a codeword of the channel codebook $\mathcal{C}_{\mathrm{R}_l^{(k)}}$ whose rate is larger than or equal to $\tilde{R}_l^{(k)}$

$$\hat{g}_l^{(k)} : \tilde{\mathcal{C}}_{\mathrm{R}_l^{(k)}} \longrightarrow \mathcal{C}_{\mathrm{R}_l^{(k)}} : \left(\hat{y}^{(i)}_{\mathrm{R}_l^{(k)}}\right)_{i=1}^{N} \longrightarrow \left(x^{(i)}_{\mathrm{R}_l^{(k)}}\right)_{i=1}^{N}. \quad (21)$$

The decoding in the destination stage is then performed in a successive fashion. In a first step, the messages sent by the relay stage $\mathcal{R}_1$ are decoded based on the sequence of receive vectors $(\mathbf{y}_{\mathcal{D}}^{(i)})_{i=1}^N$ and the knowledge of $\mathbf{H}_1$. Note that the transmission from $\mathcal{R}_1$ to $\mathcal{D}$ occurs over an uplink channel with n single-antenna transmit terminals and a receive terminal with n antennas. Since the functions $g_1^{(k)}$ are not injective, and thus not invertible, it is impossible to obtain a perfect reconstruction of the sequence of receive vectors $(\mathbf{y}_{\mathcal{R}_1}^{(i)})_{i=1}^N$. This ambiguity is accounted for by the sequence of additive quantization noise vectors $(\mathbf{q}_{\mathcal{R}_1}^{(i)})_{i=1}^N$. The $i\mathrm{th}$ element in the sequence of quantized receive vectors $\hat{\mathbf{y}}_{\mathcal{R}_1}^{(i)}$ is then written in terms of the $i\mathrm{th}$ element in the corresponding sequence of transmit vectors $\mathbf{x}_{\mathcal{R}_2}^{(i)}$ as

$$\hat{\mathbf{y}}_{\mathcal{R}_1}^{(i)} = \mathbf{H}_2\mathbf{x}_{\mathcal{R}_2}^{(i)} + \mathbf{w}_{\mathcal{R}_1}^{(i)} + \mathbf{q}_{\mathcal{R}_1}^{(i)}. \quad (22)$$

With the knowledge of the sequence $(\hat{\mathbf{y}}_{\mathcal{R}_1}^{(i)})_{i=1}^N$ and $\mathbf{H}_2$, the destination proceeds with decoding the messages of the nodes in $\mathcal{R}_2$. These messages can be considered as being transmitted over a virtual uplink with n single-antenna transmit terminals and a receive terminal with n antennas.

Proceeding this way iteratively allows for tracing back through the relay chain stage by stage based on the sequences of quantized receive vectors and the knowledge of the respective channel matrix. In the $l\mathrm{th}$ iteration, the decoder obtains the sequence of quantized receive vectors $(\hat{\mathbf{y}}_{\mathcal{R}_l}^{(i)})_{i=1}^N$ by decoding the messages of $\mathcal{R}_l$. These messages are transmitted over a virtual uplink channel with n single-antenna transmit terminals and a receive terminal with n antennas. The effective I-O relation between $\mathcal{R}_{l+1}$ and $\mathcal{D}$ can thus be written as

$$\hat{\mathbf{y}}_{\mathcal{R}_l}^{(i)} = \mathbf{H}_{l+1}\mathbf{x}_{\mathcal{R}_{l+1}}^{(i)} + \mathbf{w}_{\mathcal{R}_l}^{(i)} + \mathbf{q}_{\mathcal{R}_l}^{(i)} \quad (23)$$

where $\mathbf{q}_{\mathcal{R}_l}^{(i)}$ is the $i\mathrm{th}$ element in the respective sequence of quantization noise vectors. In the $(L+1)$-st iteration the decoder finally arrives at the source stage, whose messages are transmitted over a virtual uplink channel with n single-antenna transmit terminals and a receive terminal with n antennas. They are decoded based on the effective I-O relation

$$\hat{\mathbf{y}}_{\mathcal{R}_L}^{(i)} = \mathbf{H}_{L+1}\mathbf{x}_{\mathcal{S}}^{(i)} + \mathbf{w}_{\mathcal{R}_L}^{(i)} + \mathbf{q}_{\mathcal{R}_L}^{(i)} \quad (24)$$

and the knowledge of $\mathbf{H}_{L+1}$ and $\hat{\mathbf{y}}_{\mathcal{R}_L}^{(i)}$. The channel and quantization codebooks that are used later on render the quantization noise vectors $\mathbf{q}_{\mathcal{R}_l}^{(i)}, l \in \{1,\ldots,L\}$, i.i.d. (both in space and time) circularly symmetric complex Gaussian distributed with variance σ_l^2. Thus, for the end-to-end channel from $\mathcal{S}$ to $\mathcal{D}$, a sum-rate rate $R_{\mathcal{S}}^{\Sigma} = \sum_{k=1}^{n} R_{\mathrm{S}^{(k)}}$, where $R_{\mathrm{S}^{(k)}}$ denotes the rate of the channel codebook of $\mathrm{S}^{(k)}$, is achievable, if it fulfills

$$R_{\mathcal{S}}^{\Sigma} < \log\det\left(\mathbf{I}_n + \frac{P_L}{n\left(\sigma_{\mathrm{w}}^2 + \sigma_L^2\right)}\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\right) \triangleq R_L^{\mathrm{QF/CF}}. \quad (25)$$

The rate of the channel codebook of relay node $\mathrm{R}_l^{(k)}$ is denoted by $R_l^{(k)}$ in the following. In this work, rates of channel codebooks are required to coincide within each stage $\mathcal{R}_l$, i.e., $R_l^{(1)} = \cdots = R_l^{(n)} \triangleq R_l$. The rate R_l will in each case determine the quantization noise variance σ_l^2. For given the quantization noise variances $\sigma_{l-1}^2, l \in \{1,\ldots,L\}$, where $\sigma_0^2 = 0$, the rate R_l is achievable, if [1], [2]

$$|\tilde{\mathcal{R}}_l|R_l < \log\det\left(\mathbf{I}_n + \frac{P_L}{n\left(\sigma_{\mathrm{w}}^2 + \sigma_{l-1}^2\right)}\left(\mathbf{H}_l\right)_{\tilde{\mathcal{R}}_l}\left(\mathbf{H}_l\right)_{\tilde{\mathcal{R}}_l}^H\right) \quad \text{for all } \tilde{\mathcal{R}}_l \subseteq \mathcal{R}_l. \quad (26)$$

Here, $(\mathbf{H}_l)_{\tilde{\mathcal{R}}_l}$ denotes the $n \times |\tilde{\mathcal{R}}_l|$ matrix, that collects all columns of $\mathbf{H}_l$ that correspond to the nodes contained in $\tilde{\mathcal{R}}_l$. Moreover, the supremum of the set of rates R_l that fulfill (26) is denoted by $\bar{R}_l$ in the sequel.

In order to achieve these rates, the source and relay nodes need to generate their channel codebooks by choosing the entries $x_{\mathrm{S}^{(k)}}^{(i)}$ and $x_{\mathrm{R}_l^{(k)}}^{(i)}$ as independent realizations of circularly symmetric complex Gaussian random variable $X_{\mathcal{S}}$ and $X_{\mathcal{R}}$, respectively. The variances of these random variable are chosen to fulfill the average power constraints (9) and(10) with equality, i.e., $\mathsf{E}_{X_{\mathcal{S}}}[|X_{\mathcal{S}}|^2] = P_L/n$ and $\mathsf{E}_{X_{\mathcal{R}}}[|X_{\mathcal{R}}|^2] = P_L/n$.

In the following, the quantization codebooks of the relay nodes and the resulting quantization noise variances are specified. In this context, "pure quantization" and "quantization with Slepian and Wolf compression" is distinguished.

Pure Quantization: Pure quantization refers to a receive signal quantization method that does not exploit the statistical correlation among receive signals of relay nodes within the same stage. Each relay node $\mathrm{R}_l^{(k)}$ generates a quantization codebook $\tilde{\mathcal{C}}_l^{(k)}$ of rate $\tilde{R}_l^{(k)}$. The elements of the quantization noise vector $\mathbf{q}_{\mathcal{R}_l}^{(i)}$ are rendered i.i.d. circularly symmetric complex Gaussian with variance σ_l^2 by choosing the entries of the quantization codebook as statistically independent realizations of a circularly symmetric complex Gaussian random variable $\hat{Y}$ of variance $Q_l^{(k)} + \sigma_{\mathrm{w}}^2 + \sigma_l^2$. Here, $Q_l^{(k)}$ denotes the average power of the desired part of the receive signal at node $\mathrm{R}_l^{(k)}$. A codeword is declared to be the quantization of the observed sequence, if both are jointly typical. In order to ensure that a quantization codeword for the observed sequence is found based on a joint typicality check with probability one as $N \rightarrow \infty$, the mutual information between original and quantized observation of node $\mathrm{R}_l^{(k)}$ must fulfill

$$\tilde{R}_l^{(k)} > I\left(Y_{\mathrm{R}_l^{(k)}}^{(i)}; \hat{Y}_{\mathrm{R}_l^{(k)}}^{(i)}\right) = \log\left(1 + \frac{Q_l^{(k)} + \sigma_{\mathrm{w}}^2}{\sigma_l^2}\right) \quad (27)$$

or equivalently

$$\sigma_l^2 > \frac{Q_l^{(k)} + \sigma_{\mathrm{w}}^2}{2^{\tilde{R}_l^{(k)}} - 1}. \quad (28)$$

This inequality must be fulfilled for all nodes in $\mathcal{R}_l$. Moreover, the rates of the quantization codebooks $\tilde{R}_l^{(k)}$ must be smaller than or equal to the rate of the channel codebook, R_l. For a given $R_l < \bar{R}_l$, a quantization noise variance σ_l^2 is achievable, if it fulfills

$$\sigma_l^2 > \max_{k:\mathrm{R}_l^{(k)} \in \mathcal{R}_l} \frac{Q_l^{(k)} + \sigma_{\mathrm{w}}^2}{2^{R_l} - 1}. \tag{29}$$

Quantization With Slepian and Wolf Compression: A more efficient relaying method performs a Slepian and Wolf compression of the receive signal upon quantization. Thus, the correlation in the receive signals of relay nodes within the same stage can be efficiently exploited. As in the case of pure quantization, the elements of the quantization noise vector $\mathbf{q}_{\mathcal{R}_l}^{(i)}$ are rendered i.i.d. circularly symmetric complex Gaussian with variance σ_l^2. The rate of the compressed quantization codebook of relay node $\mathrm{R}_l^{(k)}$ is denoted by $\tilde{R}_l^{(k)}$ in the sequel. Rates of quantization codebooks within each stage $\mathcal{R}_l$ are required to coincide, i.e., $\tilde{R}_l^{(1)} = \cdots = \tilde{R}_l^{(n)} \triangleq \tilde{R}_l$. The compression problem at hand has been studied in [14] in the context of a two-hop setup with orthogonal second hop. The quantized observation vector can be reconstructed with probability one as $N \longrightarrow \infty$, if for all $\tilde{\mathcal{R}}_l \subseteq \mathcal{R}_l$

$$|\tilde{\mathcal{R}}_l|\tilde{R}_l > I\left(\left(Y_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{\mathrm{R}_l^{(k)} \in \tilde{\mathcal{R}}_l}; \left(\hat{Y}_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{\mathrm{R}_l^{(k)} \in \tilde{\mathcal{R}}_l} \middle| \left(\hat{Y}_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{\mathrm{R}_l^{(k)} \in \tilde{\mathcal{R}}_l^C}\right). \tag{30}$$

For a given quantization noise variance, the conditional mutual information expression can be written as (refer to the proof of Lemma 3)

$$I\left(\left(Y_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{\mathrm{R}_l^{(k)} \in \tilde{\mathcal{R}}_l}; \left(\hat{Y}_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{\mathrm{R}_l^{(k)} \in \tilde{\mathcal{R}}_l} \middle| \left(\hat{Y}_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{\mathrm{R}_l^{(k)} \in \tilde{\mathcal{R}}_l^C}\right) \tag{31}$$

$$= \log\det\left(\mathbf{I}_n + \frac{P_L}{(\sigma_{\mathrm{w}}^2 + \sigma_l^2)\cdot n}\mathbf{H}_{l+1}\mathbf{H}_{l+1}^H\right) - \log\det\left(\mathbf{I}_n + \frac{P_L}{(\sigma_{\mathrm{w}}^2 + \sigma_l^2)\cdot n}\left((\mathbf{H}_{l+1})_{\tilde{\mathcal{R}}_l^C}\right)^H(\mathbf{H}_{l+1})_{\tilde{\mathcal{R}}_l^C}\right) + \left|\tilde{\mathcal{R}}_l\right|\log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_l^2}\right). \tag{32}$$

Here, $(\mathbf{H}_l)_{\tilde{\mathcal{R}}_l^C}$ denotes the $|\tilde{\mathcal{R}}_l^C| \times n$ matrix that collects all rows of $\mathbf{H}_l$ which correspond to the nodes contained in $\tilde{\mathcal{R}}_l^C$. Again, the rate of the quantization codebook of each node is required to be smaller or equal to the rate of the channel codebook of the node, i.e., $\tilde{R}_l \leq R_l$. For a given $R_l < \bar{R}_l$, a quantization noise variance σ_l^2 is achievable, if it fulfills (30) for

$$\tilde{R}_l = R_l. \tag{33}$$

C. Amplify & Forward

Amplify & forward refers to the technique of deriving a transmit signal from the receive signal through simple amplification. This work focuses on the case that each relay within stage $\mathcal{R}_l$ applies the same gain factor $\sqrt{\alpha/n_{\mathcal{R}}}$, $\alpha > 0$ and sufficiently small, such that (10) is fulfilled. That is, $g_{\mathrm{R}_l^{(k)}}$ as defined in (12) is given by

$$g_{\mathrm{R}_l^{(k)}} : \mathbb{C}^N \longrightarrow \mathbb{C}^N : \left(y_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{i=1}^N \longrightarrow \left(\sqrt{\alpha/n_{\mathcal{R}}} \cdot y_{\mathrm{R}_l^{(k)}}^{(i)}\right)_{i=1}^N. \tag{34}$$

Equivalently, the relay receive and transmit vectors are related as follows:

$$\mathbf{x}_{\mathcal{R}_l}^{(i)} = \sqrt{\frac{\alpha}{n_{\mathcal{R}}}}\mathbf{y}_{\mathcal{R}_l}^{(i)} = \begin{cases} \sqrt{\frac{\alpha}{n_{\mathcal{R}}}}\left(\mathbf{H}_{l+1}\mathbf{x}_{\mathcal{R}_{l+1}}^{(i)} + \mathbf{w}_{\mathcal{R}_l}\right), & \text{if } l \in \{1, \ldots, L-1\}, \\ \sqrt{\frac{\alpha}{n_{\mathcal{R}}}}\left(\mathbf{H}_{l+1}\mathbf{x}_{\mathcal{S}}^{(i)} + \mathbf{w}_{\mathcal{R}_1}\right), & \text{if } l = L \end{cases} \tag{35}$$

where the second equality follows from (6) and(7), respectively. The effective I-O relation from source to destination stage is then obtained through recursive derivation of $\mathbf{x}_{\mathcal{R}_{l+1}}^{(i)}$ from $\mathbf{x}_{\mathcal{R}_l}^{(i)}$ according to (35), and given by

$$\mathbf{y}_{\mathcal{D}}^{(i)} = \frac{\alpha^{L/2}}{n_{\mathcal{R}}^{L/2}}\mathbf{H}_1 \cdots \mathbf{H}_{L+1}\mathbf{x}_{\mathcal{S}}^{(i)} + \mathbf{w}_{\mathcal{D}}^{(i)} + \sum_{l=1}^{L}\frac{\alpha^{l/2}}{n_{\mathcal{R}}^{l/2}}\mathbf{H}_1 \cdots \mathbf{H}_l\mathbf{w}_{\mathcal{R}_l}^{(i)}. \tag{36}$$

This I-O relation represents an $n_{\mathcal{S}}$-user uplink channel with $n_{\mathcal{D}}$ receive antennas. The sum-capacity for a fixed channel realization is achieved by a Gaussian codebook, i.e., the $x_{\mathrm{S}^{(k)}}^{(i)}$ are i.i.d. circular symmetric complex Gaussian with variance $P_L/n_{\mathcal{S}}$. Under this input-distribution a sum-rate $R_{\mathcal{S}}^{\Sigma}$ is achievable, if it fulfills [1]

$$R_{\mathcal{S}}^{\Sigma} < \log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \mathbf{R}_s\mathbf{R}_n^{-1}\right) \triangleq R_L^{\mathrm{AF}} \tag{37}$$

where $\mathbf{R}_s$ and $\mathbf{R}_n$ denote the covariance matrices of desired receive signal and accumulated noise at the destination, respectively

$$\mathbf{R}_s = \frac{P_L\alpha^L}{n_{\mathcal{S}}n_{\mathcal{R}}^L}\mathbf{H}_1 \cdots \mathbf{H}_{L+1}\mathbf{H}_{L+1}^H \cdots \mathbf{H}_1^H \tag{38}$$

$$\mathbf{R}_n = \sigma_{\mathrm{w}}^2 \cdot \left(\mathbf{I}_{n_{\mathcal{D}}} + \sum_{l=1}^{L}\frac{\alpha^l}{n_{\mathcal{R}}^l}\mathbf{H}_1 \cdots \mathbf{H}_l\mathbf{H}_l^H \cdots \mathbf{H}_1^H\right). \tag{39}$$

In the sequel, α is chosen as $\alpha = P_L/(P_L + \sigma_{\mathrm{w}}^2)$. This choice does not fulfill the power constraints (10) exactly, but in the following sense:

Proposition 1: Let $\mathbf{H}_2, \ldots, \mathbf{H}_L \in \mathbb{C}^{n_{\mathcal{R}} \times n_{\mathcal{R}}}$ and $\mathbf{H}_{L+1} \in \mathbb{C}^{n_{\mathcal{R}} \times n_{\mathcal{S}}}$ be statistically independent random matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment, and fix the ratios $n_\mathcal{R}/n_\mathcal{D}$ and $n_\mathcal{R}/n_\mathcal{S}$. If $\alpha = P_L/(P_L+\sigma_\mathrm{w}^2)$, then

- the sum-transmit power of each stage $\mathcal{R}_l$ fulfills

$$\lim_{n_\mathcal{R}\longrightarrow\infty}\sum_{k=1}^{n_\mathcal{R}} P_{\mathrm{R}_l^{(k)}} = P_L \text{ almost surely} \tag{40}$$

- for every fixed $\gamma < 1$ and every stage $\mathcal{R}_l$, $l \in \{1,\dots,L\}$

$$\Pr\left[\forall\varepsilon>0\exists n_0: \frac{1}{n_\mathcal{R}}\sum_{k=1}^{n_\mathcal{R}} 1\left\{\left|n_\mathcal{R}P_{\mathrm{R}_l^{(k)}}-P_L\right|<\varepsilon\forall n_\mathcal{R}\geq n_0\right\}>\gamma\right] = 1. \tag{41}$$

Remarks: Note that this proposition does not imply for a given l that

$$\lim_{n_\mathcal{R}\longrightarrow\infty}\max_{k\in\{1,\dots,n_\mathcal{R}\}}\left|n_\mathcal{R}\cdot P_{\mathrm{R}_l^{(k)}}-P_L\right| = 0 \text{ almost surely.} \tag{42}$$

The proof of the proposition is provided in Appendix C. Since it relies on notation and concepts introduced in Section IV-C, it is recommended to return to this proposition and its proof later.

IV. Capacity Scaling

This section establishes for each of the introduced relaying schemes the capacity scaling result as outlined above. Four theorems, one for each relaying scheme, are provided and proved. Decode & forward is examined in Section IV-A, quantize & forward without and with Slepian and Wolf compression in Section IV-B-I and Section IV-B-II, and amplify & forward in Section IV-C.

A. Decode & Forward Networks

The following theorem characterizes the scaling of the supremum of achievable sum-rates in decode & forward multi-antenna multi-hop networks:

Theorem 1: Let $\mathbf{H}_1,\dots,\mathbf{H}_{L+1}\in\mathbb{C}^{n\times n}$ be statistically independent random channel matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment. Then, the supremum of the set of sum-rates that are achievable by the decode & forward strategy, R_L^DF, fulfills for all $L\in\mathbb{N}_0$

$$\lim_{n\longrightarrow\infty}\frac{1}{n}R_L^\mathrm{DF} = \psi\left(\frac{P_L}{\sigma_\mathrm{w}^2}\right) \triangleq c_L^\mathrm{DF} \text{ almost surely} \tag{43}$$

where

$$\psi(x) \triangleq 2\log\left(1+x-\frac{1}{4}\left(\sqrt{4x+1}-1\right)^2\right) - \frac{\log e}{4x}\left(\sqrt{4x+1}-1\right)^2. \tag{44}$$

Note that the constant of proportionality that asymptotically relates R_L^DF and n is constant with respect to L, if P_L is constant with respect to L.

Proof of Theorem 1: Under the assumptions of the theorem the following holds for every $l\in\{1,\dots,L\}$ according to [15], [16] (see also [17, p. 10]):

$$\lim_{n\to\infty}\frac{1}{n}\cdot\log\det\left(\mathbf{I}_n+\frac{P_L}{n\cdot\sigma_\mathrm{w}^2}\mathbf{H}_l\mathbf{H}_l^H\right) = \psi\left(\frac{P_L}{\sigma_\mathrm{w}^2}\right) \text{ almost surely.} \tag{45}$$

Thus, there exists almost surely for every $\varepsilon>0$ and arbitrary L an $n_0(L)$, such that for all $n\geq n_0(L)$

$$\max_l\left|\frac{1}{n}\log\det\left(\mathbf{I}_n+\frac{P_L}{n\sigma_\mathrm{w}^2}\mathbf{H}_l\mathbf{H}_l^H\right)-\psi\left(\frac{P_L}{\sigma_\mathrm{w}^2}\right)\right|<\varepsilon. \tag{46}$$

For the supremum of the set of achievable sum-rates (18), this implies

$$\lim_{n\longrightarrow\infty}\frac{1}{n}R_L^\mathrm{DF} = \lim_{n\longrightarrow\infty}\min_l\frac{1}{n}\log\det\left(\mathbf{I}_n+\frac{P_L}{n\cdot\sigma_\mathrm{w}^2}\mathbf{H}_l\mathbf{H}_l^H\right) = \psi\left(\frac{P_L}{\sigma_\mathrm{w}^2}\right) \text{ almost surely.} \tag{47}$$

This establishes the theorem. □

B. Quantize & Forward Networks

Quantize & Forward Without Slepian and Wolf Compression: The following theorem characterizes the scaling of the supremum of achievable sum-rates in quantize & forward multi-antenna multi-hop networks that do not apply Slepian and Wolf compression:

Theorem 2: Let $\mathbf{H}_1,\dots,\mathbf{H}_{L+1}\in\mathbb{C}^{n\times n}$ be statistically independent random channel matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment. Then, the supremum of the set of sum-rates that are achievable by the quantize & forward strategy, R_L^QF, fulfills for all $L\in\mathbb{N}_0$

$$\lim_{n\longrightarrow\infty}\frac{1}{n}R_L^\mathrm{QF} = \psi(\mathrm{snr}_\mathcal{D}) \triangleq c_L^\mathrm{QF} \text{ almost surely} \tag{48}$$

where $\mathrm{snr}_\mathcal{D} = P_L/(\sigma_L^2+\sigma_\mathrm{w}^2)$, and $\psi(\cdot)$ is defined in (44). Moreover, the per-stage transmit power P_L that is required for rendering $\mathrm{snr}_\mathcal{D}$ constant with respect to L grows exponentially with L.

Note that the constant of proportionality that asymptotically relates R_L^QF and n is constant with respect to L, if the SNR at the destination stage, $\mathrm{snr}_\mathcal{D}$, is constant with respect to L.

Proof of Theorem 2: It has to be shown that there exists a sequence $(P_L)_{L=0}^\infty$, such that the SNR at the destination antennas, given by $\mathrm{snr}_\mathcal{D} = P_L/(\sigma_\mathrm{w}^2+\sigma_L^2)$, is kept constant and bounded away from zero as $L\longrightarrow\infty$. Then, the supremum of the set of achievable sum-rates as given by (25) fulfills [15], [16]

$$\lim_{n\longrightarrow\infty}\frac{1}{n}\cdot R_L^\mathrm{QF} = \psi(\mathrm{snr}_\mathcal{D}) \text{ almost surely} \tag{49}$$

independently of L. Since $\mathrm{snr}_\mathcal{D} > 0$, this limit is strictly positive.

First, a result on the multiple-input multiple-ouput (MIMO) uplink channel with n single-antenna sources and an n antenna destination is stated. Under the assumptions on the fading distributions of the theorem, all sources can, in the limit of large n, simultaneously achieve a fraction $1/n$ of the supremum of achievable sum-rates. More precisely, the following lemma, whose proof is given in Appendix A, holds:

Lemma 1: Consider an uplink channel with n single-antenna transmit terminals with transmit power P/n each, and an n-antenna receive terminal with spatial noise covariance matrix $\sigma^2 \mathbf{I}_n$. The elements of the channel matrix $\mathbf{H} \in \mathbb{C}^{n\times n}$ are distributed according to the assumptions of Theorem 2. Let

$$\xi = \psi\left(\frac{P}{\sigma^2}\right). \tag{50}$$

Then, there exists almost surely for every rate $R < \xi$ an n_0, such that for all $n \geq n_0$ the rate tuple $(R^{(1)}, R^{(2)}, \ldots, R^{(n)}) = (R, R, \ldots, R)$, where $R^{(i)}$ denotes the rate of the ith transmit terminal, is achievable. This lemma implies that in the limit of large n the set of achievable R_l is fully determined by the constraint in (26) that corresponds to the set $\tilde{\mathcal{R}}_l = \mathcal{R}_l$, since [15], [16]

$$\lim_{n\to\infty} \frac{1}{n} \log\det\left(\mathbf{I}_n + \frac{P_L}{n(\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2)} \mathbf{H}_l \mathbf{H}_l^H\right) = \xi_l \quad \text{almost surely} \tag{51}$$

where

$$\xi_l \triangleq \psi\left(\frac{P_L}{\sigma_{\mathrm{w}}^2 + \sigma_{l-1}^2}\right). \tag{52}$$

Thus, there is for every fixed L and given sequence $(\sigma_l^2)_{l=0}^{L-1}$ almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ all rates R_l, $l \in \{1, \ldots, L\}$, are achievable simultaneously, if $R_l < \xi_l$ for all l.

The next step, relies on the fact that as $n \longrightarrow \infty$, the receive power of all receive antennas in the various relay stages converges almost surely to $P_L + \sigma_{\mathrm{w}}^2$. To make this precise, the following lemma, whose proof is given in Appendix A, is stated:

Lemma 2: Let $\mathbf{H} \in \mathbb{C}^{n\times n}$ be a random matrix whose elements are distributed according to the assumptions of Theorem 2 and $P > 0$. Denote by $\mathbf{h}_k^T$ the kth row of $\mathbf{H}$. Then

$$\lim_{n\longrightarrow\infty} \max_{k\in\{1,\ldots,n\}} \left|\frac{P}{n} \cdot \|\mathbf{h}_k^T\|^2 - P\right| = 0 \text{ almost surely}. \tag{53}$$

Thus, the $Q_l^{(k)}$ in (29) fulfill for every fixed L

$$\lim_{n\longrightarrow\infty} \max_{l\in\{1,\ldots,L\}} \max_{k\in\{1,\ldots,n\}} \left|Q_l^{(k)} - P_L\right| = 0 \text{ almost surely}. \tag{54}$$

From the conclusions of Lemmata 1 and 2, we infer that there is for every fixed L almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ the sequence of quantization noise variances $(\sigma_l^2)_{l=1}^{L}$ is achievable according to (29), if for all $l \in \{1, \ldots, L\}$

$$\sigma_l^2 > \frac{P_L + \sigma_{\mathrm{w}}^2}{2^{\xi_l} - 1} \triangleq \left(\sigma_l^{(\inf)}\right)^2 \tag{55}$$

where the $\left(\sigma_l^{(\inf)}\right)^2$, $l \in \{1, \ldots, L\}$, denote the infima of achievable noise variances. Substitution of the asymptotic suprema of achievable rates (52) into(55) yields a first-order difference equation in σ_l^2 with $\sigma_0^2 = 0$. This difference equation is the starting point for the proof that any SNR value, $\mathrm{snr}_{\mathcal{D}}$, can be sustained at the destination antennas for increasing L by increasing P_L appropriately.

We apply the inequality[4]

$$2^{\xi_l} - 1 > \frac{P_L}{e \cdot (\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2)} \tag{56}$$

which holds for $\xi_l > 0$, to upper-bound $\left(\sigma_l^{(\inf)}\right)^2$ in (55) as follows:

$$\left(\sigma_l^{(\inf)}\right)^2 = \frac{P_L + \sigma_{\mathrm{w}}^2}{2^{\xi_l} - 1} < \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e \cdot (\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2). \tag{57}$$

Hence, we conclude that the sequence of quantization noise variances $(\sigma_l)_{l=0}^{L}$ that is characterized by the following difference equation is achievable almost surely in the limit $n \longrightarrow \infty$:

$$\sigma_l^2 = \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e \cdot (\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2) \text{ with } \sigma_0^2 = 0. \tag{58}$$

The solution to this first-order difference equation is given by

$$\sigma_l^2 = \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e \cdot \frac{1 - \left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e\right)^l}{1 - \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) \cdot e}. \tag{59}$$

Thus, for sustaining a certain SNR, $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_L^2 + \sigma_{\mathrm{w}}^2)$, P_L and L need to be coupled as follows:

$$P_L = \mathrm{snr}_{\mathcal{D}} \sigma_{\mathrm{w}}^2 \left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) e \frac{1 - \left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) e\right)^L}{1 - \left(1 + \frac{\sigma_{\mathrm{w}}^2}{P_L}\right) e} + 1\right). \tag{60}$$

This implicit equation corresponds to an exponential growth of P_L with L, since

$$\lim_{L\longrightarrow\infty} \frac{P_L}{e^L} = \mathrm{snr}_{\mathcal{D}} \cdot \sigma \mathrm{w}^2 \cdot \frac{e}{e-1}. \tag{61}$$

In order to show that an exponential growth of P_L with L is necessary for sustaining a constant destination SNR, a lower-bound on $\left(\sigma_l^{(\inf)}\right)^2$, as defined in (55), is developed. There exists for every destination SNR value $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_L^2 + \sigma_{\mathrm{w}}^2)$ a $c > 1$, such that for arbitrarily large L

$$2^{\xi_l} - 1 \leq \frac{1}{c} \cdot \rho_l \text{ for all } l \in \{1, \ldots, L\} \tag{62}$$

[4]This inequality follows, since both sides evaluate to zero for $x \triangleq \frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2} = 0$, and the slope of the left-hand side is strictly larger than the slope of the right-hand side for all $x > 0$

$$\frac{\partial}{\partial x}(2^{\xi_l} - 1) = \exp\left(-\frac{(-1+\sqrt{1+4x})^2}{4x}\right) > e^{-1} = \frac{\partial}{\partial x}(e^{-1}x).$$

where $\rho_l \triangleq P_L/(\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2)$. In order to prove this, c is chosen as[5]

$$c = \min_{l\in\{1,\ldots,L\}} \frac{\rho_l}{2^{\xi_l}-1} = \frac{\rho_L}{2^{\xi_L}-1} = \frac{\mathrm{snr}_{\mathcal{D}}}{2^{\xi_L}-1}. \tag{65}$$

This c is larger than one for positive $\mathrm{snr}_{\mathcal{D}}$, since c takes on values on the interval $(1, e)$

$$\lim_{\mathrm{snr}_{\mathcal{D}} \longrightarrow 0} c = 1 \tag{66}$$

$$\lim_{\mathrm{snr}_{\mathcal{D}} \longrightarrow \infty} c = e, \tag{67}$$

$$\frac{\partial c}{\partial \mathrm{snr}_{\mathcal{D}}} > 0 \text{ for all } \mathrm{snr}_{\mathcal{D}} > 0. \tag{68}$$

Hence, we conclude that

$$\left(\sigma_l^{(\mathrm{inf})}\right)^2 = \frac{P_L + \sigma_{\mathrm{w}}^2}{2^{\xi_l}-1} \geq \frac{P_L}{\frac{1}{c}\cdot\rho_l} = c\cdot(\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2) \quad \text{for all } l \in \{1,\ldots,L\}. \tag{69}$$

With $\left(\sigma_0^{(\mathrm{inf})}\right)^2 = 0$, this yields the following lower-bound on $\left(\sigma_L^{(\mathrm{inf})}\right)^2$:

$$\left(\sigma_L^{(\mathrm{inf})}\right)^2 > \sigma_{\mathrm{w}}^2 \cdot c \cdot \frac{1-c^L}{1-c} \tag{70}$$

That is, the power P_L that is required for sustaining a constant SNR, $\mathrm{snr}_{\mathcal{D}}$, is lower-bounded according to

$$P_L = \mathrm{snr}_{\mathcal{D}}\cdot\left(\sigma_L^2 + \sigma_{\mathrm{w}}^2\right) > \mathrm{snr}_{\mathcal{D}}\cdot\sigma_{\mathrm{w}}^2\cdot\left(c\cdot\frac{1-c^L}{1-c}+1\right). \tag{71}$$

This lower-bound implies an exponential growth of the required P_L with L. □

Quantization With Slepian and Wolf Compression: The following theorem characterizes the scaling of the supremum of achievable sum-rates in quantize & forward multi-antenna multi-hop networks that apply Slepian and Wolf compression:

Theorem 3: Let $\mathbf{H}_1, \ldots, \mathbf{H}_{L+1} \in \mathbb{C}^{n\times n}$ be statistically independent random channel matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment. Then, the supremum of the set of sum-rates that are achievable by the quantize & forward strategy with additional Slepian and Wolf compression, R_L^{CF}, fulfills for all $L \in \mathbb{N}_0$

$$\lim_{n\longrightarrow\infty} \frac{1}{n} R_L^{\mathrm{CF}} = \psi\left(\mathrm{snr}_{\mathcal{D}}\right) \triangleq c_L^{\mathrm{CF}} \text{ almost surely} \tag{72}$$

where $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_L^2 + \sigma_{\mathrm{w}}^2)$, and $\psi(\cdot)$ is defined in (44). Moreover, the per-stage transmit power P_L that is required for rendering $\mathrm{snr}_{\mathcal{D}}$ constant with respect to L grows linearly with L.

That is, the constant of proportionality that asymptotically relates R_L^{CF} and n is constant with respect to L, if the SNR at the destination stage, $\mathrm{snr}_{\mathcal{D}}$, is constant with respect to L.

Proof of Theorem 3: The proof is along the lines of the proof of Theorem 2. It has to be shown that there is for every L a P_L, such that the SNR at the destination antennas, given by $\mathrm{snr}_{\mathcal{D}} = P_L/(\sigma_{\mathrm{w}}^2 + \sigma_L^2)$, is kept constant and nonzero. Then, the supremum of the set of achievable sum-rates, as defined in (25), fulfills [15], [16]

$$\lim_{n\longrightarrow\infty} \frac{1}{n} \cdot R_L^{\mathrm{CF}} = \psi\left(\mathrm{snr}_{\mathcal{D}}\right) \text{ almost surely} \tag{73}$$

independently of L. This limit is strictly positive for positive $\mathrm{snr}_{\mathcal{D}} > 0$.

Again, Lemma 1 implies that for every fixed L and given sequence $(\sigma_l^2)_{l=0}^{L-1}$ there exists almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ all rates R_l, $l \in \{1,\ldots,L\}$, are achievable simultaneously, if $R_l < \xi_l$ for all l.

Next, the quantization noise variances are evaluated. The following lemma serves as a starting point:

Lemma 3: Let Y be a Gaussian random vector with zero-mean and covariance matrix $\mathbf{K}_{\mathsf{Y}} = \sigma_{\mathrm{w}}^2\mathbf{I}_n + \frac{P}{n}\mathbf{H}\mathbf{H}^H$, where $\mathbf{H} \in \mathbb{C}^{n\times n}$ is distributed according to the assumptions of Theorem 3 and $P > 0$. Let $\hat{\mathsf{Y}}$ be the quantization of Y, which is obtained as $\hat{\mathsf{Y}} = \mathsf{Y} + \mathsf{Z}$, where the quantization noise vector Z is a Gaussian random vector with zero-mean and covariance matrix $\mathbf{K}_{\mathsf{Z}} = \sigma_{\mathrm{q}}^2\mathbf{I}_n$. Let

$$\zeta = \psi\left(\frac{P}{\sigma_{\mathrm{w}}^2 + \sigma_{\mathrm{q}}^2}\right) + \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{\mathrm{q}}^2}\right) \tag{74}$$

where $\psi(\cdot)$ is defined in (44). Then, there exists for every tuple of rates of compressed quantization codebooks $(R^{(1)}, R^{(2)}, \ldots, R^{(n)}) = (R, R, \ldots, R)$ with $R > \zeta$ almost surely an n_0, such that for all $n \geq n_0$ the quantization noise variance σ_{q}^2 is achievable in the sense of (30).

This lemma implies that, in the limit of large n, the set of achievable σ_l^2 is fully determined by the constraint in (30) that corresponds to the set $\tilde{\mathcal{R}}_l = \mathcal{R}_l$, since [15], [16]

$$\lim_{n\to\infty} \frac{1}{n} \log\det\left(\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_l^2}\right)\mathbf{I}_n + \frac{P_L}{n(\sigma_l^2 + \sigma_{\mathrm{w}}^2)}\mathbf{H}_{l+1}\mathbf{H}_{l+1}^H\right) = \zeta_l \text{ almost surely} \tag{75}$$

where

$$\zeta_l = \psi\left(\frac{P_L}{\sigma_{\mathrm{w}}^2 + \sigma_l^2}\right) + \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_l^2}\right). \tag{76}$$

Thus, there is for every fixed L and given sequence $(\tilde{R}_l)_{l=1}^{L}$, where $\tilde{R}_l > \zeta_l$ for all l, almost surely an $n_0(L)$, such that for all

[5]Here, $l = L$ is the minimizer, since $\rho_l > \rho_L$ for every $l < L$, and $\rho_l/(2^{\xi_l} - 1)$ is monotonically increasing in ρ_l

$$\frac{\partial}{\partial\rho_l}\frac{\rho_l}{2^{\xi_l}-1} = \frac{2e^{\frac{\left(-1+\sqrt{1+4\rho_l}\right)^2}{4\rho_l}}\left(1+4\rho_l-\sqrt{1+4\rho_l}\left(2e^{\frac{\left(\sqrt{1+4\rho_l}-1\right)^2}{4\rho_l}}-1\right)\right)}{\sqrt{1+4\rho_l}\left(1-2e^{\frac{\left(-1+\sqrt{1+4\rho_l}\right)^2}{4\rho_l}}+2\rho_l+\sqrt{1+4\rho_l}\right)^2} > 0. \tag{63}$$

The positiveness of the derivative follows, since $e^x < (1-x)^{-1}$ for $x < 1$, and thus

$$2e^{\frac{\left(\sqrt{1+4\rho_l}-1\right)^2}{4\rho_l}} - 1 < \frac{2}{1-\frac{\left(\sqrt{1+4\rho_l}-1\right)^2}{4\rho_l}} - 1 = \sqrt{1+4\rho_l}. \tag{64}$$

$n \geq n_0(L)$ all quantization noise variances $\sigma_l^2, l \in \{1, \ldots, L\}$, that fulfill (76) are achievable simultaneously.

The conclusions of Lemmatas 1 and 3 allow to infer that there exists for every fixed L almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$ the sequence of quantization noise variances $(\sigma_l^2)_{l=1}^L$ is achievable according to (33), if for all $l \in \{1, \ldots, L\}$

$$\zeta_l < \tilde{R}_l = R_l < \xi_l. \tag{77}$$

For a given σ_{l-1}^2 the infimum of the achievable quantization noise variances in stage $\mathcal{R}_l$ is denoted by $\left(\sigma_l^{(\mathrm{inf})}\right)^2$. It is convenient to define

$$\Delta_l^{(\mathrm{inf})} \triangleq \begin{cases} \left(\sigma_l^{(\mathrm{inf})}\right)^2 - \sigma_{l-1}^2, \text{ if } l > 1, \\ \left(\sigma_l^{(\mathrm{inf})}\right)^2, \text{ if } l = 1. \end{cases} \tag{78}$$

The quantities $\left(\sigma_l^{(\mathrm{inf})}\right)^2$, or, equivalently $\Delta_l^{(\mathrm{inf})}$, $l \in \{1, \ldots, L\}$, are determined by the equations $\zeta_l = \xi_l$, $l \in \{1, \ldots, L\}$

$$\psi\left(\frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2}\right) = \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2}\right) + \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})}}\right). \tag{79}$$

Theorem 3 is established by showing that there exist constants $c > 0$ and $d > 1$, such that $\sigma_l^2 = (l+1)c$ and $P_L = Ld$ fulfill (77) for all $l \in \{1, \ldots, L\}$, where L can grow arbitrarily large. To this end, $\Delta_l^{(\mathrm{inf})}$ is upper-bounded by reducing the transmit power of relay stage $\mathcal{R}_l$ from $P_L = Ld$ to $P_L = ld$. This yields the upper-bound $\Delta_l^{(\mathrm{inf})} < \bar{\Delta}_l^{(\mathrm{inf})}$, since

(i) $\Delta_l^{(\mathrm{inf})}$ can only increase for a fixed σ_{l-1}^2, when the transmit power *both* of $\mathcal{R}_l$ and $\mathcal{R}_{l+1}$ are reduced from Ld to $(l+1)d$ in a first step[6].

(ii) the upper-bound on $\Delta_l^{(\mathrm{inf})}$ from (i) can again only increase for a fixed a transmit power $(l+1)d$ of stage $\mathcal{R}_{l+1}$ and fixed σ_{l-1}^2, when the transmit power of $\mathcal{R}_l$ is reduced from $(l+1)d$ to ld in a second step.

Equations (79) are rewritten in a first step as

$$\frac{\log e}{\frac{4ld}{lc+\overline{\Delta}_l^{(\mathrm{inf})}+\sigma_{\mathrm{w}}^2}}\left(\sqrt{\frac{4ld}{lc+\overline{\Delta}_l^{(\mathrm{inf})}+\sigma_{\mathrm{w}}^2}+1}-1\right)^2 + \phi(\overline{\Delta}_l^{(\mathrm{inf})}) = 0, \quad l \in \{1, \ldots, L\} \tag{80}$$

where

$$\begin{aligned}\phi(\bar{\Delta}_l^{(\mathrm{inf})}) = & -2\log\left(1 + \frac{ld}{lc + \bar{\Delta}_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2} - \frac{1}{4}\left(\sqrt{\frac{4ld}{lc + \bar{\Delta}_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2} + 1} - 1\right)^2\right) \\ & - \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{lc + \bar{\Delta}_l^{(\mathrm{inf})}}\right) + \xi_l \Big|_{P_L = ld, \sigma_{l-1}^2 = lc}\end{aligned} \tag{81}$$

and in a second step as

$$\overline{\Delta}_l^{(\mathrm{inf})} = -\frac{ld\log e}{\phi(\overline{\Delta}_l^{(\mathrm{inf})})} - \sigma_{\mathrm{w}}^2 - cl - 2ld - \frac{ld\phi(\overline{\Delta}_l^{(\mathrm{inf})})}{\log e}, \quad l \in \{1, \ldots, L\}. \tag{82}$$

We take limits on both sides (Bernoulli l'Hospital) and obtain

$$\begin{aligned}\lim_{l \longrightarrow \infty} \bar{\Delta}_l^{(\mathrm{inf})} = & \left(2d\left(c + 4d + c\sqrt{1 + \frac{4d}{c}}\right)\lim_{l \longrightarrow \infty} \bar{\Delta}_l^{(\mathrm{inf})} - \sigma_{\mathrm{w}}^2(c + 4d) \right.\\ & \left. \times\left(c + 2d + c\sqrt{1 + \frac{4d}{c}}\right)\right) \Big/ \left(2cd\sqrt{1 + \frac{4d}{c}}\right).\end{aligned} \tag{83}$$

This equation is solved for $\lim_{l \longrightarrow \infty} \bar{\Delta}_l^{(\mathrm{inf})}$ as follows:

$$\lim_{l \longrightarrow \infty} \bar{\Delta}_l^{(\mathrm{inf})} = \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}}\right). \tag{84}$$

Thus, there is for every $\varepsilon > 0$ an l_0, such that for all $l \geq l_0$

$$\left|\bar{\Delta}_l^{(\mathrm{inf})} - \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}}\right)\right| < \varepsilon. \tag{85}$$

In order to ensure that

$$\begin{aligned}\sigma_l^2 - \sigma_{l_0-1}^2 &= (l - l_0 + 1)c & (86)\\ &\geq (l - l_0 + 1) \max_{l \in \{l_0, \ldots, l\}} \left\{\bar{\Delta}_l^{(\mathrm{inf})}\right\} & (87)\\ &\geq \sum_{k=l_0}^{l} \bar{\Delta}_k^{(\mathrm{inf})} > \sum_{k=l_0}^{l} \Delta_k^{(\mathrm{inf})}, & (88)\end{aligned}$$

[6]Proof: The derivative $\partial\Delta_l^{(\mathrm{inf})}/\partial P_L$ is obtained according to the implicit function theorem

$$\begin{aligned}\frac{\partial\Delta_l^{(\mathrm{inf})}}{\partial P_L} = & \Big[-2\left(4P_L + \sigma_{\mathrm{w}}^2 + \Delta_l^{(\mathrm{inf})} + \sigma{\mathrm{w}}^2\beta - \Delta_l^{(\mathrm{inf})}\beta - \sigma_{l-1}^2\beta + \sigma_{l-1}^2\right) \\ & \times\Big(\sigma_{l-1}^4(\beta - \alpha) + \sigma_{l-1}^2(4P_L + 2\sigma_{\mathrm{w}}^2 + \Delta_l^{(\mathrm{inf})})(\beta - \alpha) \\ & + \sigma_{\mathrm{w}}^2(\sigma_{\mathrm{w}}^2 + \Delta_l^{(\mathrm{inf})})(\beta - \alpha) + P_L\left(4\sigma_{\mathrm{w}}^2(\beta - \alpha) + 4\Delta_l^{(\mathrm{inf})}\beta\right)\Big)\Big] \Big/ \\ & \Big[P_L(\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2)\alpha(1 + \alpha)(\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})})(\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2 + \Delta_l^{(\mathrm{inf})}) \\ & \times\frac{(\sigma_{l-1}^2 + 4P_L + \sigma_{\mathrm{w}}^2 + \Delta_l^{(\mathrm{inf})})(1 + \beta)^2}{(\log e)^2}\Big]\end{aligned}$$

where $\alpha = \sqrt{1 + \frac{4P_L}{\sigma_{\mathrm{w}}^2 + \sigma_{l-1}^2}}$ and $\beta = \sqrt{1\frac{4P_L}{\sigma_{\mathrm{w}}^2 + \sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})}}}$. This derivative is nonpositive for positive P_L, $\Delta_l^{(\mathrm{inf})}$, σ_{w}^2, and σ_{l-1}^2. This is seen as follows. The denominator is obviously positive. The numerator has zeros only at $\Delta_l^{(\mathrm{inf})} = 0$, $\Delta_l^{(\mathrm{inf})} = -\sigma_{\mathrm{w}}^2 - \sigma_{l-1}^2 - 4P_L$ and $\Delta_l^{(\mathrm{inf})} = -\sigma_{l-1}^2 - \frac{\sigma_{\mathrm{w}}^2 P_L}{P_L + \sigma_{\mathrm{w}}^2}$, which implies that P_L, $\Delta_l^{(\mathrm{inf})}$, σ_{w}^2, and σ_{l-1}^2 cannot be positive simultaneously at a zero, and, therefore, the numerator has the same sign for all positive tuples $(P_L, \Delta_l^{(\mathrm{inf})}, \sigma_{\mathrm{w}}^2, \sigma_{l-1}^2)$. It thus remains to show that the numerator is negative for an arbitrary choice of the positive tuple $(P_L, \Delta_l^{(\mathrm{inf})}, \sigma_{\mathrm{w}}^2, \sigma_{l-1}^2)$, e.g., for $P_L = 6, \sigma_{\mathrm{w}}^2 = 1, \sigma_{l-1}^2 = 2, \Delta_l^{(\mathrm{inf})} = 5$, it evaluates to -5760.

for all $l \in \{l_0, \dots, L\}$, we fix

$$c = \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}}\right) + \varepsilon \tag{89}$$

which for small ε leads to the coupling ($d > 1$ by assumption)

$$c \approx \frac{\sigma_{\mathrm{w}}^2 \cdot \sqrt{d}}{\sqrt{d} - \sigma_{\mathrm{w}}}. \tag{90}$$

Thus, we have shown that

$$\Delta_l^{(\mathrm{inf})} < \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}}\right) + \varepsilon \text{ for all } l > l_0. \tag{91}$$

It remains to show that there exists an L_0, such that for all $L > L_0$

$$\Delta_l^{(\mathrm{inf})} < \sigma_{\mathrm{w}}^2 \cdot \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}}\right) + \varepsilon \text{ for all } l \leq l_0. \tag{92}$$

To this end, we substitute $P_L = dL$ and $\sigma_{l-1}^2 = cl$ in (80) and take L to infinity. This results for $1 \leq l \leq l_0$ into the first order difference equation (with $\sigma_0^2 = 0$)

$$\lim_{L \longrightarrow \infty} \left(\sigma_l^{(\mathrm{inf})}\right)^2 = \sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2 \tag{93}$$

or, equivalently

$$\lim_{L \longrightarrow \infty} \Delta_l^{(\mathrm{inf})} = \sigma_{\mathrm{w}}^2. \tag{94}$$

Thus, we conclude that there is for every $\varepsilon > 0$ an L_0, such that for all $L \geq L_0$ there exists almost surely an $n_0(L)$, such that for all $n \geq n_0(L)$

$$\sigma_L^2 = Lc = L\left(\sigma_{\mathrm{w}}^2 \left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}}\right) + \varepsilon\right) > \sum_{l=1}^{L} \Delta_l^{(\mathrm{inf})} \tag{95}$$

is achievable. In order to sustain a certain value of $\mathrm{snr}_{\mathcal{D}}$ in the large n limit, it is thus sufficient that

$$\begin{aligned} P_L &= \mathrm{snr}_{\mathcal{D}}(\sigma_L^2 + \sigma_{\mathrm{w}}^2) \\ &= \mathrm{snr}_{\mathcal{D}}\sigma_{\mathrm{w}}^2 \left(1 + L\left(1 + \frac{1 + \sqrt{1 + \frac{4d}{c}}}{2\frac{d}{c}} + \varepsilon\right)\right). \end{aligned} \tag{96}$$

Since $d/c \longrightarrow \mathrm{snr}_{\mathcal{D}}$ as $L \longrightarrow \infty$, this leads to the following asymptotically linear coupling between P_L and L:

$$\lim_{L \longrightarrow \infty} \frac{P_L}{L} = \sigma_{\mathrm{w}}^2 \cdot \left((1 + \varepsilon) \cdot \mathrm{snr}_{\mathcal{D}} + \frac{1}{2} + \frac{\sqrt{1 + 4\mathrm{snr}_{\mathcal{D}}}}{2}\right). \tag{97}$$

For the converse, i.e., the proof that any P_L that scales slower than linearly with L cannot sustain a constant SNR at the destination stage, (79) serves as the starting point, again

$$\psi\left(\frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2}\right) = \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2}\right) + \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})}}\right). \tag{98}$$

The inequality[7]

$$\begin{aligned} &\log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2}\right) - \log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2}\right) \\ &\quad \geq \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2}\right) - \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2}\right). \end{aligned} \tag{101}$$

implies the inequality

$$\begin{aligned} &\log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2}\right) \geq \\ &\log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma \mathrm{w}^2}\right) + \log\left(1 + \frac{\sigma_{\mathrm{w}}^2}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})}}\right) \end{aligned} \tag{102}$$

which leads to the following lower-bounds on $\Delta_l^{(\mathrm{inf})}$ and $\left(\sigma_L^{(\mathrm{inf})}\right)^2$

$$\Delta_l^{(\mathrm{inf})} \geq \sigma_{\mathrm{w}}^2, \tag{103}$$

$$\left(\sigma_L^{(\mathrm{inf})}\right)^2 \geq L \cdot \sigma_{\mathrm{w}}^2. \tag{104}$$

Thus, the power required to sustain a constant SNR, $\mathrm{snr}_{\mathcal{D}}$, must fulfill

$$P_L = \mathrm{snr}_{\mathcal{D}} \cdot (\sigma_L^2 + \sigma_{\mathrm{w}}^2) \geq \mathrm{snr}_{\mathcal{D}} \cdot (L + 1) \cdot \sigma_{\mathrm{w}}^2. \tag{105}$$

C. *Amplify & Forward Networks*

The following theorem characterizes the scaling of the supremum of achievable sum-rates in multi-antenna amplify & forward multi-hop networks.

Theorem 4: Let $\mathbf{H}_1 \in \mathbb{C}^{n_{\mathcal{D}} \times n_{\mathcal{R}}}, \mathbf{H}_2, \dots, \mathbf{H}_L \in \mathbb{C}^{n_{\mathcal{R}} \times n_{\mathcal{R}}}$ and $\mathbf{H}_{L+1} \in \mathbb{C}^{n_{\mathcal{R}} \times n_{\mathcal{S}}}$ be statistically independent random matrices, each with i.i.d. elements of zero-mean, unit-variance and bounded fourth moment, and fix $\frac{n_{\mathcal{S}}}{n_{\mathcal{D}}} \triangleq \beta_{\mathcal{S}}$ and $\frac{n_{\mathcal{R}}}{n_{\mathcal{D}}} \triangleq \beta_{\mathcal{R}}$. Let $\mathrm{snr}_{\mathcal{D}}$ (signal-to-noise ratio at destination stage) be a positive constant. Fix

$$P_L = \sigma_{\mathrm{w}}^2 \cdot \left(\sqrt[L+1]{\frac{\mathrm{snr}_{\mathcal{D}} + 1}{\mathrm{snr}_{\mathcal{D}}}} - 1\right)^{-1} \text{ and } \alpha = \frac{P_L}{P_L + \sigma_{\mathrm{w}}^2}. \tag{106}$$

[7]Proof: Rewrite the inequality as

$$\begin{aligned} &\log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2}\right) - \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \sigma_{\mathrm{w}}^2}\right) \geq \\ &\log\left(1 + \frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma \mathrm{w}^2}\right) - \psi\left(\frac{P_L}{\sigma_{l-1}^2 + \Delta_l^{(\mathrm{inf})} + \sigma_{\mathrm{w}}^2}\right). \end{aligned} \tag{99}$$

This inequality holds, since $\Delta_l^{(\mathrm{inf})} \geq 0$ and $\log(1+x) + \psi(x)$ is monotonously increasing in x

$$\begin{aligned} &\frac{\partial(\log(1+x) - \psi(x))}{\partial x} = \\ &\frac{\log e \sqrt{1+4x}\left((1+2x) - \sqrt{1+4x}\right)}{x(1+x)\sqrt{1+4x}\left(1+\sqrt{1+4x}\right)} > 0 \text{ for all } x > 0. \end{aligned} \tag{100}$$

Let $(c_L^{\mathrm{AF}})_{L=0}^{\infty}$ be the sequence, such that $c_L^{\mathrm{AF}} = \lim_{n_{\mathcal{D}}\longrightarrow\infty} R_L^{\mathrm{AF}}/n_{\mathcal{D}}$ almost surely. Then,

$$\lim_{L\longrightarrow\infty} c_L^{\mathrm{AF}} = \begin{cases} 0, & \text{if } \beta_{\mathcal{R}} \in \mathcal{O}\left(L^{1-\varepsilon}\right), \\ \beta_{\mathcal{S}} \log\left(1+\mathrm{snr}_{\mathcal{D}} - \frac{1}{4}\chi(\mathrm{snr}_{\mathcal{D}},\beta_{\mathcal{S}})\right) \\ +\log\left(1+\mathrm{snr}_{\mathcal{D}}\beta_{\mathcal{S}} - \frac{1}{4}\chi(\mathrm{snr}_{\mathcal{D}},\beta_{\mathcal{S}})\right) \\ -\frac{\log e}{4\mathrm{snr}_{\mathcal{D}}}\chi(\mathrm{snr}_{\mathcal{D}},\beta_{\mathcal{S}}) \triangleq c_0, & \text{if } \beta_{\mathcal{R}} \in \Omega(L^{1+\varepsilon}) \end{cases} \tag{107}$$

where $\chi(x,z) = \left(\sqrt{x(1+\sqrt{z})^2+1} - \sqrt{x(1-\sqrt{z})^2+1}\right)^2$.

Remark 1: The per-stage transmit power P_L is chosen such that the destination SNR is constant with respect to L. This transmit power scales linearly with L

$$\lim_{L\longrightarrow\infty} \frac{P_L}{L} = \frac{\sigma_{\mathrm{w}}^2 \log e}{\log\left(1+\frac{1}{\mathrm{snr}_{\mathcal{D}}}\right)}. \tag{108}$$

Remark 2: Theorem 4 allows to conclude the following:

- R_L^{AF} does not scale linearly in $\min\{n_{\mathcal{S}}, n_{\mathcal{D}}\}$, if $\beta_{\mathcal{R}}$ scales less than linearly with L.
- The asymptotic sum-capacity of an $n_{\mathcal{D}} \times n_{\mathcal{S}}$ single-hop uplink channel (cf. [1], [16]) is approached for more than linear scaling of $\beta_{\mathcal{R}}$ with L.

Remark 3: The case of *exactly* linear scaling of $\beta_{\mathcal{R}}$ with L, is not considered for the sake of brevity. In this case, c_L^{AF} is still bound away from zero as $L\longrightarrow\infty$. More precisely, if $\beta_{\mathcal{R}} \in \Theta(L)$

$$\limsup_{L\longrightarrow\infty} c_L^{\mathrm{AF}} < c_0 \quad \text{and} \quad \liminf_{L\longrightarrow\infty} c_L^{\mathrm{AF}} > 0. \tag{109}$$

The corresponding proof is along the very same lines as the subsequent proofs for the two cases covered in Theorem 4.

For the proof of the theorem, the following notation is introduced. The empirical eigenvalue distribution (EED) of some Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is defined in terms of the indicator function $1\{\cdot\}$ as $F_{\mathbf{A}}^{(\gamma_1,\ldots,\gamma_K)}(x) \triangleq \frac{1}{n}\sum_{i=1}^{n} 1\{\lambda_i\{\mathbf{A}\} < x\}$. The superscripts $\gamma_1,\ldots,\gamma_K$ indicate parameters which the EED depends on. Whenever one of these parameters is taken to infinity, the respective superscript is dropped. If $\mathbf{H} \in \mathbb{C}^{n_{\mathcal{D}}\times n_{\mathcal{S}}}$ is a random matrix with i.i.d. entries of zero-mean and unit-variance, $n_{\mathcal{D}}\longrightarrow\infty$, $n_{\mathcal{S}}\longrightarrow\infty$ and $\beta_{\mathcal{S}} = n_{\mathcal{S}}/n_{\mathcal{D}}$ kept nonzero and finite, the EED $F_{\frac{1}{n_{\mathcal{S}}}\mathbf{H}\mathbf{H}^H}^{(n_{\mathcal{D}},\beta_{\mathcal{S}})}(x)$ converges uniformly and almost surely to an asymptotic EED $F_{\frac{1}{n_{\mathcal{S}}}\mathbf{H}\mathbf{H}^H}^{(\beta_{\mathcal{S}})}(x) \triangleq F_{\mathrm{MP}}^{(\beta_{\mathcal{S}})}(x)$, which is referred to as Marcenko-Pastur law [15], [19]. Hence, the supremum of the set of sum-rates that are achievable over a single-hop uplink channel with $n_{\mathcal{S}}$ single-antenna sources, an $n_{\mathcal{D}}$-antenna destination and per-node transmit power $P/n_{\mathcal{S}}$ fulfills

$$\lim_{n_{\mathcal{D}}\to\infty} \frac{1}{n_{\mathcal{D}}} R_0 = \int_0^{\infty} \log(1+\mathrm{snr}_{\mathcal{D}}\cdot x)\cdot dF_{\mathrm{MP}}^{(\beta_{\mathcal{S}})}(x) = c_0 \quad \text{almost surely}. \tag{110}$$

The closed form solution to this integral corresponds to the lower expression in (107).

The expression for R_L^{AF} in (37) can be written in terms of the EED of $\mathbf{R}_s\mathbf{R}_n^{-1}$ (normalized by $n_{\mathcal{D}}$) as follows:

$$\frac{1}{n_{\mathcal{D}}} R_L^{\mathrm{AF}} = \int_0^{\infty} \log(1+x)\cdot dF_{\mathrm{snr}_{\mathcal{D}}\cdot\tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x) \tag{111}$$

where

$$\tilde{\mathbf{R}}_s = P_L^{-1}\alpha^{-L}\cdot\mathbf{R}_s \quad \text{and} \quad \tilde{\mathbf{R}}_n = \sigma_{\mathrm{w}}^{-2}\cdot\frac{1-\alpha}{1-\alpha^{L+1}}\cdot\mathbf{R}_n. \tag{112}$$

Note that the choice of α and P_L in the theorem ensures that $\mathbf{R}_s\mathbf{R}_n^{-1} = \mathrm{snr}_{\mathcal{D}}\cdot\tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}$.

In the proof of the theorem, limits are taken in (111), first with respect to $n_{\mathcal{D}}$ and then with respect to L. The first step yields almost sure convergence to a deterministic limit, the second step drives this limit either to zero in the case $\beta_{\mathcal{R}} \in \mathcal{O}\left(L^{1-\varepsilon}\right)$ or to c_0 in the case $\beta_{\mathcal{R}} \in \Omega\left(L^{1+\varepsilon}\right)$. The essential effects behind these results are the following:

- If $\beta_{\mathcal{R}} \in \mathcal{O}\left(L^{1-\varepsilon}\right)$, (almost) all eigenvalues of $\tilde{\mathbf{R}}_s$ tend to zero, as L grows large.
- If $\beta_{\mathcal{R}} \in \Omega\left(L^{1+\varepsilon}\right)$, the asymptotic EED of $\tilde{\mathbf{R}}_s$ approaches the one that corresponds to the single-hop case, as L grows large, while $\tilde{\mathbf{R}}_n$ tends towards the identity matrix.

The following Propositions 2 and 3 make these effects precise and form the basis for the subsequent proof of Theorem 4.

Proposition 2: Given the assumptions of Theorem 4, the EED of $\tilde{\mathbf{R}}_s$ converges uniformly as $n_{\mathcal{D}}\longrightarrow\infty$. That is, there is an asymptotic EED $F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x)$, such that

$$\lim_{n_{\mathcal{D}}\to\infty} \sup_x \left| F_{\tilde{\mathbf{R}}_s}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x) - F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x) \right| = 0 \quad \text{almost surely}. \tag{113}$$

Moreover, the following statements hold:

- If $\beta_{\mathcal{R}} \in \Omega(L^{1+\varepsilon})$, the asymptotic EED of $\tilde{\mathbf{R}}_s$ converges in the limit $L\longrightarrow\infty$ pointwise to the Marcenko-Pastur law, i.e.,

$$F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(x) \triangleq \lim_{L\longrightarrow\infty} F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x) = F_{\mathrm{MP}}^{(\beta_{\mathcal{S}})}(x) \text{ for all } x. \tag{114}$$

- If $\beta_{\mathcal{R}} \in \mathcal{O}(L^{1-\varepsilon})$, the asymptotic EED of $\tilde{\mathbf{R}}_s$ converges in the limit $L\longrightarrow\infty$ pointwise to a unit step at zero, i.e.,

$$F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(x) \triangleq \lim_{L\longrightarrow\infty} F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x) = \sigma(x) \text{ for all } x. \tag{115}$$

Proposition 3: Given the assumptions of Theorem 4, the EED of $\tilde{\mathbf{R}}_n^{-1}$ converges uniformly as $n_{\mathcal{D}}\longrightarrow\infty$: there is an asymptotic EED $F_{\tilde{\mathbf{R}}_n^{-1}}^{(\beta_{\mathcal{R}},L)}(x)$, such that

$$\lim_{n_{\mathcal{D}}\to\infty} \sup_x \left| F_{\tilde{\mathbf{R}}_n^{-1}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L)}(x) - F_{\tilde{\mathbf{R}}_n^{-1}}^{(\beta_{\mathcal{R}},L)}(x) \right| = 0 \text{ almost surely}. \tag{116}$$

Moreover, for $\beta_{\mathcal{R}} \in \Omega(L^{1+\varepsilon})$ and $L\longrightarrow\infty$, the asymptotic EED of $\tilde{\mathbf{R}}_n^{-1}$, $F_{\tilde{\mathbf{R}}_n^{-1}}^{(L,\beta_{\mathcal{R}})}$ (x), converges pointwise to

$$F_{\tilde{\mathbf{R}}_n^{-1}}(x) = \sigma(x-1). \tag{117}$$

For a proof of Theorem 4, it remains to bring these two propositions together in order to characterize the asymptotic EED of $\tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}$ in both cases. Specifically, the following is nontrivial:

- If $\beta_{\mathcal{R}} \in \mathcal{O}\left(L^{1-\varepsilon}\right)$, the eigenvalues of $\tilde{\mathbf{R}}_n^{-1}$ do not die out in the same way as those of $\tilde{\mathbf{R}}_s$.
- If $\beta_{\mathcal{R}} \in \Omega\left(L^{1+\varepsilon}\right)$, the asymptotic EED of $\tilde{\mathbf{R}}_s\tilde{\mathbf{R}}_n^{-1}$ converges pointwise to that of $\tilde{\mathbf{R}}_s$, i.e., to $F_{\mathrm{MP}}^{\beta_{\mathcal{S}}}(x)$ as $L \rightarrow \infty$.

Proof of Theorem 4: The two cases are considered separately. Case $\beta_{\mathcal{R}} \in \mathcal{O}\left(L^{1-\varepsilon}\right)$: It has to be shown that $\lim_{L\rightarrow\infty} c_L^{\mathrm{AF}} = 0$. Since $n_{\mathcal{D}}^{-1}R_L^{\mathrm{AF}}$ is known to converge to a nonrandom constant for every L almost surely as $n_{\mathcal{D}} \rightarrow \infty$ [3], this is implied by $\lim_{L\rightarrow\infty}\lim_{n_{\mathcal{D}}\rightarrow\infty} n_{\mathcal{D}}^{-1}\mathsf{E}\left[R_L^{\mathrm{AF}}\right] = 0$. This, in turn, can be shown based on the following lemma:

Lemma 4: $\mathsf{E}\left[R_L^{\mathrm{AF}}\right]$ is for all $L_0 \in \{1,\ldots,L\}$ upper-bounded according to

$$\mathsf{E}\left[R_L^{\mathrm{AF}}\right] \leq \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \frac{P_L s_1\cdots s_{L_0}\alpha^L}{\sigma_{\mathrm{w}}^2 n_{\mathcal{S}} n_{\mathcal{R}}^{L-L_0}\sum_{l=0}^{L_0} s_1\cdots s_l\alpha^l} \times \tilde{\mathbf{V}}_{L_0}^H \mathbf{H}_{L_0+1}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{L_0+1}^H\tilde{\mathbf{V}}_{L_0}\right)\right] \tag{118}$$

where $\tilde{\mathbf{V}}_{L_0}^H$ is an $n_{\mathcal{D}} \times n_{\mathcal{R}}$ matrix with orthonormal rows that is obtained through the following sequence of singular value decompositions:

$$\tilde{\mathbf{H}}_1 \triangleq \mathbf{H}_1 = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^H = \mathbf{U}_1\tilde{\mathbf{S}}_1\tilde{\mathbf{V}}_1^H \tag{119}$$
$$\tilde{\mathbf{H}}_2 \triangleq \tilde{\mathbf{V}}_1^H\mathbf{H}_2 = \mathbf{U}_2\mathbf{S}_2\mathbf{V}_2^H = \mathbf{U}_2\tilde{\mathbf{S}}_2\tilde{\mathbf{V}}_2^H \tag{120}$$
$$\tilde{\mathbf{H}}_3 \triangleq \tilde{\mathbf{V}}_2^H\mathbf{H}_3 = \mathbf{U}_3\mathbf{S}_3\mathbf{V}_3^H = \mathbf{U}_3\tilde{\mathbf{S}}_3\tilde{\mathbf{V}}_3^H. \tag{121}$$
$$\vdots$$

The matrices $\tilde{\mathbf{S}}_k$ and $\tilde{\mathbf{V}}_k$ correspond to the first $n_{\mathcal{D}}$ columns of $\mathbf{S}_k$ and $\mathbf{V}_k$, respectively, and $s_k \triangleq n_{\mathcal{R}}^{-1}\mathrm{Tr}\left[\tilde{\mathbf{H}}_k\tilde{\mathbf{H}}_k^H\right]$ for $k \in \{1,\ldots,L_0\}$.

Remark: This upper-bound corresponds to the supremum of sum-rates that are achievable by the amplify & forward strategy over an equivalent network with:

- noiseless relay stages $\mathcal{R}_{L_0+1},\ldots,\mathcal{R}_L$;
- the relay stage $\mathcal{R}_{L_0}$ replaced by an equivalent destination stage with $n_{\mathcal{D}}$ antennas, whose preceding hop has a channel matrix $\tilde{\mathbf{V}}_{L_0}^H\mathbf{H}_{L_0+1}$,
- white noise of power $\sigma_{\mathrm{w}}^2 \cdot \sum_{l=0}^{L_0}\alpha^l$ at the equivalent destination stage (corresponds to the noise power at the original destination stage due to the removed relay stages $\mathcal{R}_{L_0},\ldots,\mathcal{R}_1$, plus the noise power of the original destination stage).

This lemma is applied for $L_0 = \lfloor L/2\rfloor$ to obtain

$$\begin{aligned}\frac{1}{n_{\mathcal{D}}}\mathsf{E}\left[R_L^{\mathrm{AF}}\right] &\leq \frac{1}{n_{\mathcal{D}}}\mathsf{E}\left[\log\det\left(\mathbf{I}_{n_{\mathcal{D}}} + \rho_L\cdot\mathbf{A}\right)\right]\\ &= \frac{1}{n_{\mathcal{D}}}\mathsf{E}\left[\sum_{i:\lambda_i\{\mathbf{A}\}\leq\delta_1}\log\left(1+\rho_L\lambda_i\{\mathbf{A}\}\right)\right]\\ &\quad + \frac{1}{n_{\mathcal{D}}}\mathsf{E}\left[\sum_{i:\lambda_i\{\mathbf{A}\}>\delta_1}\log\left(1+\rho_L\lambda_i\{\mathbf{A}\}\right)\right]\end{aligned} \tag{122}$$

where

$$\mathbf{A} \triangleq \frac{1}{n_{\mathcal{S}} n_{\mathcal{R}}^{\lceil L/2\rceil}} \times \tilde{\mathbf{V}}_{\lfloor L/2\rfloor}^H\mathbf{H}_{\lfloor L/2\rfloor+1}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{\lfloor L/2\rfloor+1}^H\tilde{\mathbf{V}}_{\lfloor L/2\rfloor} \tag{123}$$

and

$$\rho_L \triangleq \frac{P_L s_1\cdots s_{\lfloor L/2\rfloor}\alpha^L}{\sigma_{\mathrm{w}}^2\cdot\sum_{l=0}^{\lfloor L/2\rfloor} s_1\cdots s_l\alpha^l} \tag{124}$$

with $\tilde{\mathbf{V}}_k$ and s_k constructed as in Lemma 4. Now, limits with respect to L and $n_{\mathcal{D}}$ can be taken on the right-hand side of (122). To this end, fix $\varepsilon > 0$ arbitrarily small, and fix $\delta_1 > 0$ sufficiently small, such that

$$\log(1 + 2\cdot\mathrm{snr}_{\mathcal{D}}\cdot\delta_1) < \frac{\varepsilon}{2} \tag{125}$$

and fix $\delta_2 > 0$ sufficiently small, such that

$$\delta_2\cdot\log\left(1 + 2\cdot\frac{\mathrm{snr}_{\mathcal{D}}}{\delta_2}\right) < \frac{\varepsilon}{2}. \tag{126}$$

The two sums in (122) are considered individually. The first sum fulfills independently of L

$$\lim_{n_{\mathcal{D}}\rightarrow\infty}\frac{1}{n_{\mathcal{D}}}\sum_{i:\lambda_i\{\mathbf{A}\}\leq\delta_1}\log\left(1+\rho_L\lambda_i\{\mathbf{A}\}\right) \tag{127}$$
$$\leq \log\left(1+2\cdot\mathrm{snr}_{\mathcal{D}}\cdot\delta_1\right) < \frac{\varepsilon}{2}\ \text{almost surely.} \tag{128}$$

This bound is obtained in two steps:

- Since the sum comprises no more than $n_{\mathcal{D}}$ terms that are upper-bounded by $\log\left(1+\rho_L\cdot\delta_1\right)$ each, one obtains
$$\frac{1}{n_{\mathcal{D}}}\sum_{i:\lambda_i\{\mathbf{A}\}\leq\delta_1}\log\left(1+\rho_L\lambda_i\{\mathbf{A}\}\right) \leq \log\left(1+\rho_L\cdot\delta_1\right). \tag{129}$$
- Since $\lim_{n_{\mathcal{D}}\rightarrow\infty} s_k = 1$ almost surely for all k, one obtains
$$\lim_{n_{\mathcal{D}}\rightarrow\infty}\rho_L = \frac{P_L\alpha^L}{\sigma_{\mathrm{w}}^2\cdot\sum_{l=0}^{\lfloor L/2\rfloor}\alpha^l} = \frac{P_L\alpha^L}{\sigma_{\mathrm{w}}^2\cdot\frac{1-\alpha^{\lfloor L/2\rfloor+1}}{1-\alpha}} \tag{130}$$
$$= \mathrm{snr}_{\mathcal{D}}\frac{1-\alpha^{L+1}}{1-\alpha^{\lfloor L/2\rfloor+1}} \leq \mathrm{snr}_{\mathcal{D}}\frac{1-\alpha^L}{1-\sqrt{\alpha^L}} \tag{131}$$
$$< 2\cdot\mathrm{snr}_{\mathcal{D}}\ \text{almost surely.} \tag{132}$$

The last inequality follows, since $0 \leq \alpha^L = \left(P_L/(\sigma_{\mathrm{w}}^2+P_L)\right)^L < 1$.

For the second sum, there exists for every $\varepsilon > 0$ an L_0, such that for all $L \geq L_0$

$$\lim_{n_{\mathcal{D}}\rightarrow\infty}\frac{1}{n_{\mathcal{D}}}\sum_{i:\lambda_i\{\mathbf{A}\}>\delta_1}\log\left(1+\rho_L\lambda_i\{\mathbf{A}\}\right) \leq \delta_2\cdot\log\left(1+2\cdot\frac{\mathrm{snr}_{\mathcal{D}}}{\delta_2}\right) < \frac{\varepsilon}{2}\ \text{almost surely.} \tag{133}$$

For establishing this upper-bound, the following lemmata are used:

Lemma 5: Let $\mathbf{H}_{\lfloor L/2\rfloor+1},\ldots,\mathbf{H}_{L+1}$ be as in Theorem 4. Moreover, define an $n_{\mathcal{D}}\times n_{\mathcal{R}}$ random matrix $\mathbf{X}$ whose elements follow a distribution independent of the elements of $\mathbf{H}_{\lfloor L/2\rfloor+1},\ldots,\mathbf{H}_{L+1}$ and fulfill $\frac{1}{n_{\mathcal{S}}}\mathbf{X}\mathbf{X}^H=\mathbf{I}_{n_{\mathcal{D}}}$. Then, the EEDs of $\mathbf{A}\triangleq\frac{1}{n_{\mathcal{R}}^{\lceil L/2\rceil}n_{\mathcal{D}}}\mathbf{X}\mathbf{H}_{\lfloor L/2\rfloor+1}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{\lfloor L/2\rfloor+1}^H\mathbf{X}^H$ and $\mathbf{B}\triangleq\frac{1}{n_{\mathcal{R}}^{\lceil L/2\rceil-1}n_{\mathcal{D}}}\tilde{\mathbf{H}}_{\lfloor L/2\rfloor+1}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\tilde{\mathbf{H}}_{\lfloor L/2\rfloor+1}^H$ converge, as $n_{\mathcal{D}}\longrightarrow\infty$ and $\beta_{\mathcal{S}}$ and $\beta_{\mathcal{R}}$ are fixed, uniformly and almost surely to the same asymptotic EED, where $\tilde{\mathbf{H}}_{\lfloor L/2\rfloor+1}$ is the matrix that contains the first $n_{\mathcal{D}}$ rows of $\mathbf{H}_{\lfloor L/2\rfloor+1}$.

Lemma 6: Let the matrices $\mathbf{H}_2,\mathbf{H}_3,\ldots,\mathbf{H}_{L+1}$ be as in Theorem 4. Moreover, let $\mathbf{H}_1$ be an arbitrary $n_{\mathcal{D}}\times n_{\mathcal{R}}$ random matrix that fulfills $\lim_{n_{\mathcal{R}}\longrightarrow\infty}n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[n_{\mathcal{R}}^{-1}\mathbf{H}_1^H\mathbf{H}_1\right]=1$ for every fixed ratio $n_{\mathcal{R}}/n_{\mathcal{D}}$ almost surely. Define $\mathbf{A}_l\triangleq\frac{1}{n_{\mathcal{S}}n_{\mathcal{R}}^L}\mathbf{H}_1\mathbf{H}_2\cdots\mathbf{H}_{l+1}\mathbf{H}_{l+1}^H\cdots\mathbf{H}_2^H\mathbf{H}_1^H$. Then, for every pair of fixed ratios $n_{\mathcal{R}}/n_{\mathcal{S}}$ and $n_{\mathcal{D}}/n_{\mathcal{S}}$ and any $l\in\{1,\ldots,L\}$

$$\lim_{n_{\mathcal{D}}\longrightarrow\infty}n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\mathbf{A}_l\right]=1\text{ almost surely.}\tag{134}$$

The bound (133) is obtained in four steps:

- Due to the concavity of the log-function, Jensen's inequality can be invoked to establish:

$$\begin{aligned}&\frac{n}{n_{\mathcal{D}}}\frac{1}{n}\sum_{i=1}^{n}\log\left(1+\rho_L\lambda_i\{\mathbf{A}\}\right) &(135)\\ &\le\frac{n}{n_{\mathcal{D}}}\log\left(1+\rho_L\frac{1}{n}\sum_{i=1}^{n}\lambda_i\{\mathbf{A}\}\right) &(136)\\ &\le\frac{n}{n_{\mathcal{D}}}\log\left(1+\rho_L\frac{1}{n}\sum_{i=1}^{n_{\mathcal{D}}}\lambda_i\{\mathbf{A}\}\right) &(137)\\ &=\frac{n}{n_{\mathcal{D}}}\log\left(1+\rho_L\frac{n_{\mathcal{D}}}{n}\frac{1}{n_{\mathcal{D}}}\mathrm{Tr}[\mathbf{A}]\right), &(138)\end{aligned}$$

  where we choose n as the maximal i that fulfills $\lambda_i\{\mathbf{A}\}>\delta_1$.
- From Lemma 5, we know that Proposition 2 applies also, if $\tilde{\mathbf{R}}_s$ is replaced by the random matrix $\mathbf{A}$. Note that $\beta_{\mathcal{R}}\in\mathcal{O}\left(L^{1-\varepsilon}\right)$, if and only if also $\lfloor L/2\rfloor\in\mathcal{O}\left(L^{1-\varepsilon}\right)$. Thus, there exists by Proposition 2 an L_0, such that there exists for all $L\ge L_0$ almost surely an $n_{\mathcal{D}}^{(0)}(L_0)$, such that for all $n_{\mathcal{D}}\ge n_{\mathcal{D}}^{(0)}(L_0)$ the fraction of eigenvalues of $\mathbf{A}$ larger than δ_1 fulfills

$$\frac{n}{n_{\mathcal{D}}}=1-F_{\mathbf{A}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L)}(\delta_1)<\delta_2.\tag{139}$$

- Moreover, Lemma 6 implies $\lim_{n_{\mathcal{D}}\longrightarrow\infty}n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\mathbf{A}\right]=1$ almost surely.
- Finally, we use again that $\lim_{n_{\mathcal{D}}\longrightarrow\infty}\rho_L<2\cdot\mathrm{snr}_{\mathcal{D}}$ almost surely.

Thus, by combining (128) and (133), we have shown that there exists an L_0, such that there exists for all $L\ge L_0$ almost surely an $n_{\mathcal{D}}^{(0)}(L)$, such that for all $n_{\mathcal{D}}\ge n_{\mathcal{D}}^{(0)}(L_0)$(122) fulfills

$$\frac{1}{n_{\mathcal{D}}}\mathsf{E}\left[R_L^{\mathrm{AF}}\right]<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.\tag{140}$$

Case $\beta_{\mathcal{R}}\in\Omega\left(L^{1+\varepsilon}\right)$: Define the Hermitian matrix $\mathbf{\Theta}\triangleq\tilde{\mathbf{R}}_n-\mathbf{I}_{n_{\mathcal{D}}}$ and rewrite it in terms of its eigenvalue decomposition

$$\begin{aligned}\mathbf{\Theta}&=\sum_{i=1}^{n_{\mathcal{D}}}\lambda_i\{\Theta\}\mathbf{u}_i\mathbf{u}_i^H &(141)\\ &=\underbrace{\sum_{i:\lambda_i\{\Theta\}>0}\lambda_i\{\Theta\}\mathbf{u}_i\mathbf{u}_i^H}_{\mathbf{\Theta}^+}+\underbrace{\sum_{i:\lambda_i\{\Theta\}\le 0}\lambda_i\{\Theta\}\mathbf{u}_i\mathbf{u}_i^H}_{\mathbf{\Theta}^-} &(142)\end{aligned}$$

where $\mathbf{u}_i$ denotes the eigenvector that corresponds to the ith eigenvalue. Let $\mathbf{A}$ be an arbitrary positive semidefinite matrix with an asymptotic EED that depends on L and $\beta_{\mathcal{R}}$ and fulfills $\lim_{L\longrightarrow\infty}F_{\mathbf{A}}^{(L,\beta_{\mathcal{R}})}(x)=F_{\mathbf{A}}(x)$. We show that the asymptotic EEDs $F_{\mathbf{A}+\mathbf{\Theta}}^{(L,\beta_{\mathcal{R}})}(x)$ and $F_{\mathbf{A}}^{(L,\beta_{\mathcal{R}})}(x)$ coincide in the limit $L\longrightarrow\infty$

$$\begin{aligned}&\int_0^\infty\left|F_{\mathbf{A}+\mathbf{\Theta}}^{(L,\beta_{\mathcal{R}},n_{\mathcal{D}})}(x)-F_{\mathbf{A}}^{(L,\beta_{\mathcal{R}},n_{\mathcal{D}})}(x)\right|\cdot dx\\ &=\frac{1}{n_{\mathcal{D}}}\sum_{i=1}^{n_{\mathcal{D}}}\left|\lambda_i\{\mathbf{A}+\mathbf{\Theta}\}-\lambda_i\{\mathbf{A}\}\right| &(143)\\ &=\frac{1}{n_{\mathcal{D}}}\sum_{i:\lambda_i\{\mathbf{A}+\mathbf{\Theta}\}>\lambda_i\{\mathbf{A}\}}\left(\lambda_i\{\mathbf{A}+\mathbf{\Theta}\}-\lambda_i\{\mathbf{A}\}\right)\\ &\quad+\frac{1}{n_{\mathcal{D}}}\sum_{j:\lambda_j\{\mathbf{A}+\mathbf{\Theta}\}\le\lambda_j\{\mathbf{A}\}}\left(\lambda_j\{\mathbf{A}\}-\lambda_j\{\mathbf{A}+\mathbf{\Theta}\}\right) &(144)\\ &\le\frac{1}{n_{\mathcal{D}}}\sum_{i:\lambda_i\{\mathbf{A}+\mathbf{\Theta}\}>\lambda_i\{\mathbf{A}\}}\left(\lambda_i\{\mathbf{A}+\mathbf{\Theta}^+\}-\lambda_i\{\mathbf{A}\}\right)\\ &\quad+\frac{1}{n_{\mathcal{D}}}\sum_{j:\lambda_j\{\mathbf{A}+\mathbf{\Theta}\}\le\lambda_j\{\mathbf{A}\}}\left(\lambda_j\{\mathbf{A}\}-\lambda_j\{\mathbf{A}+\mathbf{\Theta}^-\}\right) &(145)\\ &\le\frac{1}{n_{\mathcal{D}}}\sum_{i=1}^{n_{\mathcal{D}}}\left(\lambda_i\{\mathbf{A}+\mathbf{\Theta}^+\}-\lambda_i\{\mathbf{A}\}\right)\\ &\quad+\frac{1}{n_{\mathcal{D}}}\sum_{j=1}^{n_{\mathcal{D}}}\left(\lambda_j\{\mathbf{A}\}-\lambda_j\{\mathbf{A}+\mathbf{\Theta}^-\}\right) &(146)\\ &=\frac{1}{n_{\mathcal{D}}}\sum_{i=1}^{n_{\mathcal{D}}}\left(\lambda_i\{\mathbf{A}+\mathbf{\Theta}^+\}-\lambda_i\{\mathbf{A}+\mathbf{\Theta}^-\}\right) &(147)\\ &=\frac{1}{n_{\mathcal{D}}}\mathrm{Tr}[\mathbf{A}+\mathbf{\Theta}^+-\mathbf{A}-\mathbf{\Theta}^-] &(148)\\ &=\frac{1}{n_{\mathcal{D}}}\mathrm{Tr}[\mathbf{\Theta}^+-\mathbf{\Theta}^-]\\ &=\frac{1}{n_{\mathcal{D}}}\sum_{i=1}^{n_{\mathcal{D}}}\left|\lambda_i\{\mathbf{\Theta}\}\right|=\frac{1}{n_{\mathcal{D}}}\|\mathbf{\Theta}\|_*. &(149)\end{aligned}$$

The first inequality follows, since removing the negative/positive definite part of $\mathbf{\Theta}$ can only increase/decrease each individual eigenvalue of $\mathbf{A}+\mathbf{\Theta}$. The second inequality follows, since all added terms are nonnegative. With Proposition 3 we can thus conclude that there exists for every $\varepsilon>0$ an L_0, such that for all $L\ge L_0$

$$\begin{aligned}&\lim_{n_{\mathcal{D}}\to\infty}\int_0^\infty\left|F_{\mathbf{A}+\mathbf{\Theta}}^{(L,\beta_{\mathcal{R}},n_{\mathcal{D}})}(x)-F_{\mathbf{A}}^{(L,\beta_{\mathcal{R}},n_{\mathcal{D}})}(x)\right|\cdot dx\\ &\quad\le\lim_{n_{\mathcal{D}}\to\infty}\frac{1}{n_{\mathcal{D}}}\|\mathbf{I}_{n_{\mathcal{D}}}-\tilde{\mathbf{R}}_n\|_*<\varepsilon\text{ almost surely} &(150)\end{aligned}$$

which implies that $F_{\mathbf{A}+\mathbf{\Theta}}(x)$ converges pointwise to $F_{\mathbf{A}}(x)$.

In the following, we use the identity

$$R_L^{\rm AF} = \log\det(\mathbf{I}_{n_\mathcal{D}} + {\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_n^{-1}\tilde{\mathbf{R}}_s) \quad (151)$$
$$= \log\det(\tilde{\mathbf{R}}_n + {\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_s) - \log\det(\tilde{\mathbf{R}}_n). \quad (152)$$

We consider the first term on the right-hand side and identify $\mathbf{A}$ with $\mathbf{I}_{n_\mathcal{D}} + {\rm snr}_\mathcal{D}\cdot\mathbf{R}_s$. Thus, we obtain

$$\frac{1}{n_\mathcal{D}}\log\det\left(\mathbf{I}_{n_\mathcal{D}} + \mathbf{\Theta} + {\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_s\right)$$
$$= \frac{1}{n_\mathcal{D}}\sum_{i=1}^{n_\mathcal{D}}\log(1+\lambda_i\{\mathbf{\Theta}+{\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_s\}) \quad (153)$$
$$= \int_0^{\left(1+\sqrt{\beta_\mathcal{S}^{-1}}\right)^2}\log(1+x)\cdot dF_{\mathbf{\Theta}+{\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_s}^{(n_\mathcal{D},\beta_\mathcal{S},L,\beta_\mathcal{R})}(x) + \frac{1}{n_\mathcal{D}}\sum_{i:\lambda_i\{\mathbf{\Theta}+\tilde{\mathbf{R}}_s\}>\left(1+\sqrt{\beta_\mathcal{S}^{-1}}\right)^2}\log(1+\lambda_i\{\mathbf{\Theta}+{\rm snr}_\mathcal{D}\tilde{\mathbf{R}}_s\}) \quad (154)$$

where the integration interval corresponds to the support of $\partial F_{\rm MP}^{(\beta_\mathcal{S})}(x)/\partial x$. We drop the sum, use the fact $F_{\mathbf{\Theta}+{\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_s}^{(\beta_\mathcal{S},L,\beta_\mathcal{R})}(x)$ converges pointwise to $F_{{\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_s}^{(\beta_\mathcal{S})}(x)$ as $L\rightarrow\infty$, and thus obtain with Proposition 2 that there exists for every $\varepsilon>0$ an L_0, such that for all $L\geq L_0$ almost surely

$$\lim_{n_\mathcal{D}\rightarrow\infty}\frac{1}{n_\mathcal{D}}\log\det\left(\mathbf{I}_{n_\mathcal{D}} + \mathbf{\Theta} + {\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_s\right) > \int_0^{\left(1+\sqrt{\beta_\mathcal{S}^{-1}}\right)^2}\log(1+{\rm snr}_\mathcal{D}\cdot x)\cdot dF_{\rm MP}^{(\beta_\mathcal{S})}(x) - \frac{\varepsilon}{2}. \quad (155)$$

Next, we investigate the second term in (152). Application of Jensen's inequality yields

$$\frac{1}{n_\mathcal{D}}\log\det(\mathbf{I}_{n_\mathcal{D}}+\mathbf{\Theta}) = \frac{1}{n_\mathcal{D}}\sum_{i=1}^{n_\mathcal{D}}\log(1+\lambda_i\{\mathbf{\Theta}\}) \quad (156)$$
$$\leq \log\left(1+\frac{1}{n_\mathcal{D}}\sum_i\lambda_i\{\mathbf{\Theta}\}\right) \quad (157)$$
$$= \log\left(1+\frac{1}{n_\mathcal{D}}{\rm Tr}[\mathbf{\Theta}]\right). \quad (158)$$

Since $|{\rm Tr}[\mathbf{\Theta}]|\leq\|\mathbf{\Theta}\|_*$, we obtain with (150) that there exists for every $\varepsilon>0$ an L_0, such that for all $L\geq L_0$

$$-\lim_{n_\mathcal{D}\longrightarrow\infty}\frac{1}{n_\mathcal{D}}\log\det(\mathbf{I}_{n_\mathcal{D}}+\mathbf{\Theta}) > -\frac{\varepsilon}{2}\text{ almost surely.} \quad (159)$$

We have shown that there exists for every $\varepsilon>0$ an L_0, such that for all $L\geq L_0$

$$\lim_{n_\mathcal{D}\longrightarrow\infty}\frac{1}{n_\mathcal{D}}\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\mathbf{\Theta}+{\rm snr}_\mathcal{D}\cdot\tilde{\mathbf{R}}_s\right) - \frac{1}{n_\mathcal{D}}\log\det(\mathbf{I}_{n_\mathcal{D}}+\mathbf{\Theta}) > \int_0^{\left(1+\sqrt{\beta_\mathcal{S}^{-1}}\right)^2}\log(1+{\rm snr}_\mathcal{D}\cdot x)\cdot dF_{\rm MP}^{(\beta_\mathcal{S})}(x) - \varepsilon \text{ almost surely.} \quad (160)$$

The integral evaluates to [16]

$$c_0 = \beta_\mathcal{S}\log\left(1+{\rm snr}_\mathcal{D} - \frac{1}{4}\chi({\rm snr}_\mathcal{D},\beta_\mathcal{S})\right) \quad (161)$$
$$+\log\left(1+{\rm snr}_\mathcal{D}\beta_\mathcal{S} - \frac{1}{4}\chi({\rm snr}_\mathcal{D},\beta_\mathcal{S})\right) - \frac{\log e}{4{\rm snr}_\mathcal{D}}\chi({\rm snr}_\mathcal{D},\beta_\mathcal{S}) \quad (162)$$

which is the sum-capacity of a single-hop uplink channel with $n_\mathcal{S}$ single-antenna sources and an $n_\mathcal{D}$-antenna destination.

It remains to prove that the single-hop sum-capacity cannot be exceeded. To this end, we can again consider $\mathsf{E}[R_L^{\rm AF}]$, since $c_L^{\rm AF} = \lim_{n_\mathcal{D}\longrightarrow\infty}\mathsf{E}[n_\mathcal{D}^{-1}R_L^{\rm AF}]$. We apply Lemma 4 for $L_0 = L$ to obtain

$$\frac{1}{n_\mathcal{D}}\mathsf{E}\left[R_L^{\rm AF}\right] \leq \frac{1}{n_\mathcal{D}}\mathsf{E}\left[\log\det(\mathbf{I}_{n_\mathcal{D}}+\rho_L\cdot\mathbf{A})\right] \quad (163)$$

where

$$\mathbf{A} = \frac{1}{n_\mathcal{S}n_\mathcal{R}}\tilde{\mathbf{V}}_L^H\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\tilde{\mathbf{V}}_L \quad (164)$$
$$\rho_L = \frac{P_L s_1\cdots s_L\alpha^L}{\sigma_{\rm w}^2\cdot\sum_{l=0}^L s_1\cdots s_l\alpha^l} \quad (165)$$

with $\tilde{\mathbf{V}}_k$ and s_k constructed as in Lemma 4.

Since $\lim_{n_\mathcal{D}\longrightarrow\infty}s_k = 1$ almost surely for all k, we obtain

$$\lim_{n_\mathcal{D}\longrightarrow\infty}\rho_L = {\rm snr}_\mathcal{D}\text{ almost surely.} \quad (166)$$

Moreover, according to Lemma 5, the asymptotic EED of $\mathbf{A}$ coincides with the asymptotic EED of $\frac{1}{n_\mathcal{S}}\mathbf{H}\mathbf{H}^H$, where $\mathbf{H}$ contains the first $n_\mathcal{D}$ rows of $\mathbf{H}_{L+1}$. Thus, we can write

$$\lim_{n_\mathcal{D}\longrightarrow\infty}\frac{1}{n_\mathcal{D}}\sum_{i=1}^{n_\mathcal{D}}\log(1+{\rm snr}_\mathcal{D}\lambda_i\{\mathbf{A}\}) \quad (167)$$
$$= \lim_{n_\mathcal{D}\longrightarrow\infty}\int_0^{\left(1+\sqrt{\beta_\mathcal{S}^{-1}}\right)^2}\log(1+{\rm snr}_\mathcal{D}\cdot x)dF_\mathbf{A}^{(\beta_\mathcal{S},n_\mathcal{D})}(x) + \lim_{n_\mathcal{D}\longrightarrow\infty}\frac{1}{n_\mathcal{D}}\sum_{i:\lambda_i\{\mathbf{A}\}>\left(1+\sqrt{\beta_\mathcal{S}^{-1}}\right)^2}\log(1+{\rm snr}_\mathcal{D}\cdot\lambda_i\{\mathbf{A}\}) \quad (168)$$
$$= \int_0^{\left(1+\sqrt{\beta_\mathcal{S}^{-1}}\right)^2}\log(1+{\rm snr}_\mathcal{D}\cdot x)dF_{\rm MP}(x)\text{ almost surely} \quad (169)$$

which is the sum-capacity of a single-hop uplink channel. Here, we have taken the limit inside the definite integral according to the bounded convergence theorem. The second term in (168) evaluates to zero due to the concavity of the log-function and Jensen's inequality: we choose n as the maxmimal i, such that $\lambda_i\{\mathbf{A}\} > \left(1+\sqrt{\beta_\mathcal{S}^{-1}}\right)^2$ and write

$$\frac{n}{n_\mathcal{D}}\frac{1}{n}\sum_{i=1}^{n}\log(1+{\rm snr}_\mathcal{D}\lambda_i\{\mathbf{A}\}) \quad (170)$$
$$\leq \frac{n}{n_\mathcal{D}}\log\left(1+{\rm snr}_\mathcal{D}\frac{1}{n}\sum_{i=1}^{n}\lambda_i\{\mathbf{A}\}\right) \quad (171)$$

TABLE III
ILLUSTRATION OF $l_{m,j}$: POSITION OF SET $\mathcal{M}_m$ IN THE DECODING ORDER OF THE $j\text{th}$ DECODING PHASE

set	time slot 1	time slot 2	time slot 3	...	time slot $M-1$	time slot M
$\mathcal{M}_1$	1	2	3	...	$M-1$	M
$\mathcal{M}_2$	M	1	2	...	$M-2$	$M-1$
$\mathcal{M}_3$	$M-1$	M	1	...	$M-3$	$M-2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\mathcal{M}_{M-1}$	3	4	5	...	1	2
$\mathcal{M}_M$	2	3	4	...	M	1

$$\leq \frac{n}{n_{\mathcal{D}}} \log\left(1 + \text{snr}_{\mathcal{D}} \frac{1}{n} \sum_{i=1}^{n_{\mathcal{D}}} \lambda_i\{\mathbf{A}\}\right) \tag{172}$$

$$= \frac{n}{n_{\mathcal{D}}} \log\left(1 + \text{snr}_{\mathcal{D}} \frac{n_{\mathcal{D}}}{n} \frac{1}{n_{\mathcal{D}}} \text{Tr}[\mathbf{A}]\right). \tag{173}$$

From Lemma 6, we know that $\lim_{n_{\mathcal{D}} \to \infty} n_{\mathcal{D}}^{-1} \text{Tr}\,[\mathbf{A}] = 1$ almost surely. Since $n/n_{\mathcal{D}} \to 0$ as $n_{\mathcal{D}} \to \infty$ almost surely—note that $\partial F_{\text{MP}}^{(\beta_{\mathcal{S}})}(x)/\partial x$ is not supported for $x > \left(1 + \sqrt{\beta_{\mathcal{S}}^{-1}}\right)^2$—the sum converges to zero almost surely. We have thus shown that an achievable sum-rate cannot exceed the sum-capacity of a single-hop uplink channel. □

V. CONCLUSIONS

In this work, it was shown that the choice of the relaying scheme is crucial in long multi-hop networks. While there is unsurprisingly an inherent performance gap between decode & forward and the investigated non-regenerative schemes due to noise accumulation, the key contribution of this work is the identification of the fundamental differences among the performances of non-regenerative relaying schemes in the regime of long multi-hop networks.

APPENDIX A
PROOFS OF LEMMATA FOR THEOREMS 2 AND 3

Proof of Lemma 1: Consider the sum-capacity achieving minimum-mean-square error (MMSE) successive interference cancellation (SIC) receiver structure [2]. We modify this receiver as follows: rather than applying interference cancellation after each decoded codeword, we decode codewords of multiple transmit terminals simultaneously based on a given MMSE equalizer output signal. We group the transmit terminals into the sets $\mathcal{M}_m, m \in \{1, \ldots, M\}$. Codewords of transmit terminals within the same set $\mathcal{M}_m$ are decoded simultaneously based on the same MMSE equalizer output signal each, i.e., there are $M-1$ interference cancellation steps in total. Since n is not an integer multiple of M in general, we choose the cardinality of these sets as

$$|\mathcal{M}_m| = \begin{cases} \lceil n/M \rceil, & \text{if } 1 \leq m \leq n \bmod M \\ \lfloor n/M \rfloor, & \text{else.} \end{cases} \tag{174}$$

We introduce the one-to-one map $l : \{1, \ldots, M\} \longrightarrow \{1, \ldots, M\} : m \longrightarrow l_m$, where l_m corresponds to the position of set $\mathcal{M}_m$ in the decoding order. E.g., $l_3 = 4$ implies that set $\mathcal{M}_3$ is decoded based on the MMSE output of the fourth "decoding phase", when the transmit signals of three other sets are already canceled.

Let us denote the number of mutually interfering streams in the decoding phase of set $\mathcal{M}_m$ by n_m and consider the ratio

$$\frac{n_m}{n} = \frac{\sum_{k=l_m}^{M} \left|\mathcal{M}_{(l^{-1})_k}\right|}{n}. \tag{175}$$

Here, we denote by $\left(l^{-1}\right)_k$ the inverse function of l_m. This ratio is upper- and lower-bounded as follows:

$$\frac{n_m}{n} \leq \frac{(M - l_m + 1) \cdot \lceil \frac{n}{M} \rceil}{n} \tag{176}$$

$$< \frac{(M - l_m + 1) \cdot \left(\frac{n}{M} + 1\right)}{n} \tag{177}$$

$$\frac{n_m}{n} \geq \frac{(M - l_m + 1) \cdot \lfloor \frac{n}{M} \rfloor}{n} \tag{178}$$

$$> \frac{(M - l_m + 1) \cdot \left(\frac{n}{M} - 1\right)}{n}. \tag{179}$$

Since both bounds converge to the same value as n grows large, we conclude

$$\lim_{n \to \infty} \frac{n_m}{n} = \frac{M - l_m + 1}{M} \triangleq \beta_{l_m}. \tag{180}$$

In the following, we assume that transmission is divided into M time slots of N symbol durations each. Each transmit terminal transmits a sequence of M codewords of length N, one in each time slot. We specify the decoding order for the codewords of the set $\mathcal{M}_m$ that are transmitted in the $j\text{th}$ time slot through the function $l : \{1, \ldots, M\}^2 \longrightarrow \{1, \ldots, M\} : (m, j) \longrightarrow l_{m,j}$, where

$$l_{m,j} = ((M - m + j) \bmod M) + 1. \tag{181}$$

This function is one-to-one in j for each fixed m and illustrated in Table III. The codewords of the transmit terminals in set $\mathcal{M}_m$ for the $j\text{th}$ time slot are selected from a common code whose rate $R(l_{m,j})$ is fully determined by the value of $l_{m,j}$, i.e., the respective position in the decoding order for this time slot. The average code rate over the M codebooks of the transmit terminals in set $\mathcal{M}_m$ is given by

$$R^{(m)} = \frac{1}{M} \sum_{j=1}^{M} R(l_{m,j}). \tag{182}$$

Since (i) the codomain of $l_{m,j}$ with respect to j is the same for all m, and (ii) for each fixed m the function $l_{m,j}$ is one to one in j, this average rate $R^{(m)}$ is the *same* for all $\mathcal{M}_m$, i.e.,

$$R^{(1)} = R^{(2)} = \cdots = R^{(M)}. \tag{183}$$

In the following, we use a result [21, Lemma 3.1] on the t signal-to-interference-plus-noise ratios (SINRs), $(\mathsf{SINR}_k)_{k=1}^t$, at the output of an MMSE receiver in an $r \times t$ MIMO channel with channel coefficients as assumed in this lemma. If $t/r \longrightarrow \beta$ as $n \longrightarrow \infty$, then there exists a constant $\mathsf{SINR}^{(\infty)}(\beta)$, such that almost surely

$$\lim_{n \longrightarrow \infty} \max_{k \in \{1,\ldots,t\}} \left| \mathsf{SINR}_k - \mathsf{SINR}^{(\infty)}(\beta) \right| = 0. \quad (184)$$

We apply this result, for each of the M time slots, to each of the M decoding phases in the corresponding modified MMSE-SIC receiver. Specifically, since M is finite, we conclude that almost surely

$$\lim_{n\to\infty} \max_{(j,m)\in\{1,\ldots,M\}^2} \max_{k\in\mathcal{M}_m} \left| \mathsf{SINR}_{k,j} - \mathsf{SINR}^{(\infty)}\left(\beta_{l_{m,j}}\right) \right| = 0 \quad (185)$$

where $\mathsf{SINR}_{k,j}$ denotes the SINR for the signal of transmit terminal k in time slot j, and $\beta_{l_{m,j}}$ is defined analogously to (180). Thus, in the limit of large n, $R(l_{m,j})$ is achievable for all transmit terminals in $\mathcal{M}_m$ in time slot j almost surely, if

$$R(l_{m,j}) < \log\left(1 + \mathsf{SINR}^{(\infty)}\left(\beta_{l_{m,j}}\right)\right). \quad (186)$$

The average (over time slots) of the suprema of achievable rates for the transmit terminals in set $\mathcal{M}_m$ is given by

$$R_m = \frac{1}{M}\sum_{j=1}^{M} \log\left(1 + \min_{k\in\mathcal{M}_m} \mathsf{SINR}_{k,j}\right). \quad (187)$$

According to (185), there exists for every $\varepsilon > 0$ almost surely an n_0, such that for all $n \geq n_0$

$$\max_{m\in\{1,\ldots,M\}} \left| R_m - \tilde{\xi} \right| < \frac{\varepsilon}{2}, \quad (188)$$

where due to the choice of $l_{m,j}$

$$\tilde{\xi} = \frac{1}{M}\sum_{j=1}^{M} \log\left(1 + \mathsf{SINR}^{(\infty)}\left(\beta_{l_{1,j}}\right)\right) \quad (189)$$

$$= \ldots = \frac{1}{M}\sum_{j=1}^{M} \log\left(1 + \mathsf{SINR}^{(\infty)}\left(\beta_{l_{M,j}}\right)\right). \quad (190)$$

Next, we show that there is for each $\varepsilon > 0$ an M_0, such that for all $M \geq M_0$

$$\left| \xi - \tilde{\xi} \right| < \frac{\varepsilon}{2} \quad (191)$$

where ξ fulfills [15]

$$\lim_{n \longrightarrow \infty} \frac{1}{n} \log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) = \xi \text{ almost surely.} \quad (192)$$

Let us fix j and denote by $\overline{\mathsf{SINR}}_{k,j}$ the SINR of the kth transmit terminal in time slot j, when interference of each codeword is canceled individually, once it is decoded. That is, we consider the SINRs as seen in the capacity achieving structure. Thus, the following relation holds [2]:

$$\log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) = \sum_{k=1}^{n} \log(1+\overline{\mathsf{SINR}}_{k,j}). \quad (193)$$

We fix the decoding order in this case, such that the codewords of the transmit terminals in set $\mathcal{M}_m$ are decoded in the decoding phases

$$\left\{ \sum_{i<l_{m,j}} \left|\mathcal{M}_{\left(l_j^{-1}\right)_i}\right| + 1, \ldots, \sum_{i\leq l_{m,j}} \left|\mathcal{M}_{\left(l_j^{-1}\right)_i}\right| \right\}. \quad (194)$$

Here, we denote by $\left(l_j^{-1}\right)_i$ the inverse function of $l_{m,j}$ with respect to m. Since the SINR of each transmit terminal signal becomes the larger the more interference signals are canceled, we have $\overline{\mathsf{SINR}}_{k',j} \leq \mathsf{SINR}_{k,j}$ for all $k,\ k'$, such that $k \in \mathcal{M}_m$ and $k' \in \mathcal{M}_{m'}$ and $l_{m,j} > l_{m',j}$. Thus, we conclude

$$\begin{aligned}
&\sum_{k=1}^{n} \log\left(1 + \mathsf{SINR}_{k,j}\right) \\
&\quad = \sum_{i=1}^{M} \sum_{k\in\mathcal{M}_{\left(l_j^{-1}\right)_i}} \log\left(1 + \mathsf{SINR}_{k,j}\right) && (195)\\
&\quad \geq \sum_{i=2}^{M} \sum_{k\in\mathcal{M}_{\left(l_j^{-1}\right)_i}} \log\left(1 + \mathsf{SINR}_{k,j}\right) && (196)\\
&\quad \geq \sum_{k=1}^{n-\left|\mathcal{M}_{\left(l_j^{-1}\right)_M}\right|} \log\left(1 + \overline{\mathsf{SINR}}_{k,j}\right) && (197)\\
&\quad \geq \log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) \\
&\qquad - \left|\mathcal{M}_{\left(l_j^{-1}\right)_M}\right| \max_{k\in\mathcal{M}_{\left(l_j^{-1}\right)_M}} \log\left(1 + \overline{\mathsf{SINR}}_{k,j}\right). && (198)
\end{aligned}$$

Here, the last inequality follows from (193). Taking the limit with respect to n in (195), normalized by n^{-1}, yields almost surely

$$\begin{aligned}
&\lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{j=1}^{M} \sum_{k\in\mathcal{M}_{\left(l_j^{-1}\right)_j}} \log\left(1 + \mathsf{SINR}_{k,j}\right) \\
&\quad = \sum_{m=1}^{M} \left(\lim_{n \longrightarrow \infty} \frac{\left|\mathcal{M}_{l_{m,j}}\right|}{n} \right) \log\left(1 + \mathsf{SINR}^{(\infty)}\left(\beta_{l_{m,j}}\right)\right) && (199)\\
&\quad = \frac{1}{M}\sum_{m=1}^{M} \log\left(1 + \mathsf{SINR}^{(\infty)}\left(\beta_{l_{m,j}}\right)\right) = \tilde{\xi}. && (200)
\end{aligned}$$

Likewise, we obtain for (198) almost surely

$$\lim_{n\longrightarrow\infty}\frac{1}{n}\log\det\left(\mathbf{I}_n+\frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) - \frac{\left|\mathcal{M}_{(l_i^{-1})_M}\right|}{n}\max_{k\in\mathcal{M}_{(l_i^{-1})_M}}\log\left(1+\overline{\mathsf{SINR}}_{k,j}\right) \tag{201}$$

$$=\xi-\frac{1}{M}\log\left(1+\mathsf{SINR}^{(\infty)}(0)\right). \tag{202}$$

Since $\mathsf{SINR}^{(\infty)}(0)$ is finite almost surely according to [21, Lemma 3.1], there exists for every $\varepsilon>0$ an M_0, such that for all $M\geq M_0$ we have $\frac{1}{M}\log\left(1+\mathsf{SINR}^{(\infty)}(0)\right)<\varepsilon/2$. Thus, we eventually conclude

$$\xi-\frac{\varepsilon}{2}<\tilde{\xi}<\xi, \tag{203}$$

which establishes (191). The second inequality follows, since ξ is the normalized asymptotic sum-capacity of the channel.

We finally combine (188) and (191) in order to conclude that there exists for every $\varepsilon>0$ almost surely an n_0, such that for all $n\geq n_0$

$$\max_{m\in\{1,\ldots,M\}}|R_m-\xi|\leq\left|R_m-\tilde{\xi}\right|+\left|\xi-\tilde{\xi}\right|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}<\varepsilon \tag{204}$$

where we used the triangle inequality. Thus, the rate tuple $(R^{(1)},R^{(2)},\ldots,R^{(n)})=(R,R,\ldots,R)$ is achievable almost surely in the limit $n\longrightarrow\infty$, if $R<\xi$. □

Proof of Lemma 2: We decompose the matrix $\mathbf{H}\in\mathbb{C}^{n\times n}$ into the matrices $\tilde{\mathbf{H}}_m\in\mathbb{C}^{n_m\times n}$, $m\in\{1,\ldots,M\}$, where

$$n_m=\begin{cases}\lceil n/M\rceil, & \text{if } 1\leq m\leq n \bmod M\\ \lfloor n/M\rfloor, & \text{else.}\end{cases} \tag{205}$$

The matrices $\tilde{\mathbf{H}}_m$ contain disjoint sections of $\mathbf{H}$, such that

$$\mathbf{H}=\left(\tilde{\mathbf{H}}_1^T\ldots\tilde{\mathbf{H}}_M^T\right)^T. \tag{206}$$

With this notation, we obtain the following bounds on $\|\mathbf{h}_k^T\|^2$:

$$\frac{1}{n}\lambda_{\min}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\}\leq\min_{k:\lceil k/M\rceil=m}\frac{1}{n}\|\mathbf{h}_k^T\|^2 \tag{207}$$

$$\leq\max_{k:\lceil k/M\rceil=m}\frac{1}{n}\|\mathbf{h}_k^T\|^2 \tag{208}$$

$$\leq\frac{1}{n}\lambda_{\max}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\}. \tag{209}$$

Here, we used that the maximum/minimum Eucledian norm of any row of a matrix is upper-/lower-bounded by the maximum/minimum singular value of the matrix.

From [20], it is known that almost surely

$$\lim_{n\longrightarrow\infty}\lambda_{\max}\frac{1}{n}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\}=\lim_{n\longrightarrow\infty}\left(1+\sqrt{\frac{n_m}{n}}\right)^2=\left(1+\sqrt{\frac{1}{M}}\right)^2 \tag{210}$$

$$\lim_{n\longrightarrow\infty}\lambda_{\min}\frac{1}{n}\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\}=\lim_{n\longrightarrow\infty}\left(1-\sqrt{\frac{n_m}{n}}\right)^2=\left(1-\sqrt{\frac{1}{M}}\right)^2. \tag{211}$$

Since M is finite, we can conclude that also almost surely

$$\lim_{n\longrightarrow\infty}\max_{m\in\{1,\ldots,M\}}\frac{1}{n}\lambda\max\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\}=\left(1+\sqrt{\frac{1}{M}}\right)^2 \tag{212}$$

$$\lim_{n\longrightarrow\infty}\min_{m\in\{1,\ldots,M\}}\frac{1}{n}\lambda\min\left\{\tilde{\mathbf{H}}_m\tilde{\mathbf{H}}_m^H\right\}=\left(1-\sqrt{\frac{1}{M}}\right)^2. \tag{213}$$

In the following, we choose M for a given $\varepsilon>0$ sufficiently large, such that

$$P\left(1+\sqrt{\frac{1}{M}}\right)^2<P+\frac{\varepsilon}{2} \tag{214}$$

and

$$P\left(1-\sqrt{\frac{1}{M}}\right)^2>P-\frac{\varepsilon}{2}. \tag{215}$$

Then, there exists for every $\varepsilon>0$ almost surely an n_0, such that for all $n\geq n_0$

$$\max_{k\in\{1,\ldots,n\}}\frac{P}{n}\|\mathbf{h}_k^T\|^2-P\leq\max_{m\in\{1,\ldots,M\}}\frac{P}{n}\lambda_{\max}\{\mathbf{H}_m\mathbf{H}_m^H\}-P<P\left(1+\sqrt{\frac{1}{M}}\right)^2+\frac{\varepsilon}{2}-P<\frac{\varepsilon}{2}. \tag{216}$$

The first inequality follows from (207), the second one from (212), and the third one from (214). Likewise, there exists for every $\varepsilon>0$ almost surely an n_0, such that for all $n\geq n_0$.

$$\min_{k\in\{1,\ldots,n\}}\frac{P}{n}\|\mathbf{h}_k^T\|^2-P\geq\min_{m\in\{1,\ldots,M\}}\frac{P}{n}\lambda_{\min}\{\mathbf{H}_m\mathbf{H}_m^H\}-P \tag{217}$$

$$>P\left(1-\sqrt{1/M}\right)^2-\frac{\varepsilon}{2}-P>-\frac{\varepsilon}{2}. \tag{218}$$

Again, the first inequality follows from (207), the second one from (213), and the third one from (215). The combination of the bounds (216) and (218) yields, that there also exists for every $\varepsilon>0$ almost surely an n_0, such that for all $n\geq n_0$

$$\max_{k\in\{1,\ldots,n\}}\left|\frac{P}{n}\|\mathbf{h}_k^T\|^2-P\right|<\varepsilon \tag{219}$$

which establishes the Lemma. □

Proof of Lemma 3: We rewrite the considered conditional mutual information as follows:

$$\begin{aligned} I(\mathsf{Y}_{\mathcal{S}};\hat{\mathsf{Y}}_{\mathcal{S}}|\hat{\mathsf{Y}}_{\mathcal{S}^C}) &= h(\hat{\mathsf{Y}}_{\mathcal{S}}|\hat{\mathsf{Y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{Y}}_{\mathcal{S}}|\mathsf{Y}_{\mathcal{S}},\hat{\mathsf{Y}}_{\mathcal{S}^C}) &(220)\\ &= h(\hat{\mathsf{Y}}_{\mathcal{S}}|\hat{\mathsf{Y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{Y}}_{\mathcal{S}}|\mathsf{Y}_{\mathcal{S}}) &(221)\\ &= h(\hat{\mathsf{Y}}_{\mathcal{S}},\hat{\mathsf{Y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{Y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{Y}}_{\mathcal{S}}|\mathsf{Y}_{\mathcal{S}}) &(222)\\ &= h(\hat{\mathsf{Y}}) - h(\hat{\mathsf{Y}}_{\mathcal{S}^C}) - h(\hat{\mathsf{Y}}_{\mathcal{S}}|\mathsf{Y}_{\mathcal{S}}). &(223)\end{aligned}$$

Here, we used the fact that $\hat{\mathsf{Y}}_{\mathcal{S}}$ is independent of $\hat{\mathsf{Y}}_{\mathcal{S}^C}$ when conditioned on $\mathsf{Y}_{\mathcal{S}}$ in order to obtain (221). The three entropy expressions are evaluated as follows:

$$\begin{aligned} h(\hat{\mathsf{Y}}) &= n\log\left(2\pi e(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\right)\\ &\quad + \log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\mathbf{H}\mathbf{H}^H\right) &(224)\\ h(\hat{\mathsf{Y}}_{\mathcal{S}^C}) &= \left|\mathcal{S}^C\right|\log\left(2\pi e(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\right)\\ &\quad + \log\det\left(\mathbf{I}_{|\mathcal{S}^C|} + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\left(\mathbf{H}_{\mathcal{S}^C}\right)^H\mathbf{H}_{\mathcal{S}^C}\right) &(225)\\ &= \left|\mathcal{S}^C\right|\log\left(2\pi e(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\right)\\ &\quad + \log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\mathbf{H}_{\mathcal{S}^C}\left(\mathbf{H}_{\mathcal{S}^C}\right)^H\right) &(226)\\ h(\hat{\mathsf{Y}}_{\mathcal{S}}|\mathsf{Y}_{\mathcal{S}}) &= h(\mathsf{z}) = |\mathcal{S}|\log(2\pi e\sigma_{\mathrm{q}}^2). &(227)\end{aligned}$$

Here, $\mathbf{H}_{\mathcal{S}^C}$ denotes the matrix that contains the $|\mathcal{S}^C|$ columns of $\mathbf{H}$ whose indexes are contained in $\mathcal{S}^C$. Thus, we obtain according to (223)

$$\begin{aligned} I(\mathsf{Y}_{\mathcal{S}};\hat{\mathsf{Y}}_{\mathcal{S}}|\hat{\mathsf{Y}}_{\mathcal{S}^C}) &=\\ &\log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\mathbf{H}\mathbf{H}^H\right)\\ &- \log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\mathbf{H}_{\mathcal{S}^C}\left(\mathbf{H}_{\mathcal{S}^C}\right)^H\right)\\ &+ |\mathcal{S}|\log\left(1+\frac{\sigma_{\mathrm{w}}^2}{\sigma_{\mathrm{q}}^2}\right). &(228)\end{aligned}$$

Next, we use the following corollary that follows from Lemma 1:

Corollary 1: Let the elements of the random matrix $\mathbf{H} \in \mathbb{C}^{n\times n}$ be distributed according to the assumptions of Theorem 2. Let P and σ^2 be positive constants. Then, there exists almost surely an n_0, such that for all $n \geq n_0$:

$$\begin{aligned} \frac{|\mathcal{S}|}{n}&\left(\frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right)\right)\\ &\leq \frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}_{\mathcal{S}}\left(\mathbf{H}_{\mathcal{S}}\right)^H\right)\\ &\qquad \text{for all } \mathcal{S}\subseteq\{1,\ldots,n\} &(229)\end{aligned}$$

where $\mathbf{H}_{\mathcal{S}}$ denotes the matrix that contains the $|\mathcal{S}|$ columns of $\mathbf{H}$ whose indexes are contained in $\mathcal{S}$.

The proof of this corollary is provided subsequent to this proof. According to Corollary 1, there exists almost surely an n_0, such that for all $n \geq n_0$

$$\begin{aligned} \frac{1}{n}&\log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\mathbf{H}_{\mathcal{S}^C}\left(\mathbf{H}_{\mathcal{S}^C}\right)^H\right)\\ &\geq \frac{\left|\mathcal{S}^C\right|}{n}\left(\frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\mathbf{H}\mathbf{H}^H\right)\right)\\ &\qquad \text{for all } \mathcal{S}\subseteq\{1,\ldots,n\}. &(230)\end{aligned}$$

We apply this result to (228), and conclude that there is almost surely an n_0, such that for all $n \geq n_0$

$$\begin{aligned} \frac{1}{n}&I(\mathsf{Y};\hat{\mathsf{Y}}_{\mathcal{S}}|\hat{\mathsf{Y}}_{\mathcal{S}^C})\\ &\leq \frac{1}{n}\cdot\left(1-\frac{\left|\mathcal{S}^C\right|}{n}\right)\cdot\log\det\left(\mathbf{I}_n + \frac{P}{(\sigma_{\mathrm{w}}^2+\sigma_{\mathrm{q}}^2)\cdot n}\mathbf{H}\mathbf{H}^H\right)\\ &\quad + \frac{|\mathcal{S}|}{n}\log\left(1+\frac{\sigma_{\mathrm{w}}^2}{\sigma_{\mathrm{q}}^2}\right) &(231)\\ &= \frac{|\mathcal{S}|}{n}\cdot\frac{1}{n}\log\det\left(\left(1+\frac{\sigma_{\mathrm{w}}^2}{\sigma_{\mathrm{q}}^2}\right)\mathbf{I}_n + \frac{P}{n}\mathbf{H}\mathbf{H}^H\right)\\ &\qquad \text{for all } \mathcal{S}\subseteq\{1,\ldots,n\} &(232)\end{aligned}$$

where (232) converges to $\frac{|\mathcal{S}|}{n}\zeta$ almost surely as $n\rightarrow\infty$, i.e., there exists almost surely for every tuple of rates of compressed quantization codebooks $(R^{(1)},R^{(2)},\ldots,R^{(n)}) = (R,R,\ldots,R)$ with $R > \zeta$ an n_0, such that for all $n \geq n_0$ the quantization noise variance σ_{q}^2 is achievable in the sense of (30). □

Proof of Corollary 1: Lemma 1 implies that there is for all $R < \xi$ almost surely an n_0, such that for all $n \geq n_0$

$$\begin{aligned} |\mathcal{S}|R < \log\det&\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\cdot\mathbf{H}_{\mathcal{S}}(\mathbf{H}_{\mathcal{S}})^H\right)\\ &\text{for all } \mathcal{S}\subseteq\{1,\ldots,n\}. &(233)\end{aligned}$$

Since almost surely

$$\xi = \lim_{n\longrightarrow\infty}\frac{1}{n}\log\det\left(\mathbf{I}_n + \frac{P}{n\cdot\sigma^2}\mathbf{H}\mathbf{H}^H\right) \quad (234)$$

this is equivalent to (229). □

APPENDIX B
PROOFS OF PROPOSITIONS AND LEMMATA FOR THEOREM 4

Proof of Proposition 2: The uniform and almost sure convergence of $F_{\tilde{\mathbf{R}}_s}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x)$ to $F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(x)$ for $n_{\mathcal{D}}\rightarrow\infty$ follows from the following result in [7]:

Lemma 7: Let $\mathbf{M}_1 \in \mathbb{C}^{k_0\times k_1},\ldots,\mathbf{M}_N \in \mathbb{C}^{k_{N-1}\times k_N}$ be statistically independent random matrices that fulfill the conditions for the Marcenko-Pastur law and have elements of unit-variance. Define $\beta_n = \frac{k_n}{k_0}$. Then, the EED of

$$\mathbf{A} \triangleq 1/(k_1\cdots k_N)\mathbf{M}_1\cdots\mathbf{M}_N\mathbf{M}_N^H\cdots\mathbf{M}_1^H \quad (235)$$

converges uniformly and almost surely as $k_0 \longrightarrow \infty$: there exists an asymptotic $F_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(x)$, such that

$$\lim_{k_0\to\infty} \sup_x \left| F_{\mathbf{A}}^{(k_0,\beta_1,\ldots,\beta_N}(x) - F_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(x) \right| = 0 \quad \text{almost surely.} \tag{236}$$

Moreover, the Stieltjes transform of this asymptotic EED[8] fulfills the implicit equation

$$\frac{G_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(s)}{\beta_N} \prod_{n=0}^{N-1} \frac{sG_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(s) - 1 + \beta_{n+1}}{\beta_n} + sG_{\mathbf{A}}^{(\beta_1,\ldots,\beta_N)}(s) = 1. \tag{237}$$

Remark: Note that $\mathbf{A}$ is normalized with respect to k_N here, while it is normalized with respect to k_0 in [7]. The Stieltjes transform $\tilde{G}(s)$ therein relates to $G(s)$ as $G(s) = \beta_N \tilde{G}(\beta_N s)$.

For the setting considered in this work, N is identified with $L+1$, $\mathbf{M}_n$ with $\mathbf{H}_l$, k_N with $n_{\mathcal{S}}$, and k_n, with $n_{\mathcal{R}}$ for all $n < N$. This yields for our setting the implicit equation

$$\Psi_{\tilde{\mathbf{R}}_s} \cdot \frac{G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s)}{\beta_{\mathcal{S}}} \cdot \frac{sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) - 1 + \beta_{\mathcal{S}}}{\beta_{\mathcal{R}}} \times \left(sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) - 1 + \beta_{\mathcal{R}} \right) + sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) = 1 \tag{238}$$

where

$$\Psi_{\tilde{\mathbf{R}}_s} = \left(\frac{sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) - 1}{\beta_{\mathcal{R}}} + 1 \right)^{L-1}. \tag{239}$$

The asymptotic analysis for $L \longrightarrow \infty$ can now be performed in the Stieltjes domain, since the EED $F_{\mathbf{A}}^{(\gamma)}(x)$ converges to $F_{\mathbf{A}}(x)$ pointwise with respect to γ, if and only if $G_{\mathbf{A}}^{(\gamma)}(s)$ converges to $G_{\mathbf{A}}(s)$ pointwise [18, Corollary 1]. The basis for this asymptotic analysis is the following lemma, whose proof is provided subsequently in this Appendix:

Lemma 8: Let $\varepsilon > 0$ and g be some function $g : \mathbb{N} \longrightarrow \mathbb{Q}^+ : \kappa \longrightarrow g(\kappa)$.

- Then, for all $c \in \mathbb{C}$ and $g(\kappa) \in \Omega(\kappa^{1+\varepsilon})$

$$\lim_{\kappa\longrightarrow\infty} \left(\frac{c}{g(\kappa)} + 1 \right)^{\kappa} = 1. \tag{240}$$

- Then, for all negative c and $g(\kappa) \in \mathcal{O}(\kappa^{1-\varepsilon})$

$$\lim_{\kappa\longrightarrow\infty} g(\kappa) \cdot \left(\frac{c}{g(\kappa)} + 1 \right)^{\kappa} = 0. \tag{241}$$

Lemma 8 is applied to $\Psi_{\tilde{\mathbf{R}}_s}$, where L is identified with κ and $\beta_{\mathcal{R}}$ with $g(\kappa)$. In the limit $L \longrightarrow \infty$, the implicit (238) simplifies as follows:

1) If $\beta_{\mathcal{R}} \in \Omega(L^{1+\varepsilon})$, then $\Psi_{\tilde{\mathbf{R}}_s} \longrightarrow 1$, and thus

$$\beta_{\mathcal{S}}^{-1} s G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})^2}(s) + \left(s + 1 - \beta_{\mathcal{S}}^{-1}\right) G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(s) = 1. \tag{242}$$

The solution to this equation is the Stieltjes transform of the Marcenko-Pastur law with parameter $\beta_{\mathcal{S}}$.

2) If $\beta_{\mathcal{R}} \in \mathcal{O}(L^{1-\varepsilon})$, then $\Psi_{\tilde{\mathbf{R}}_s} \longrightarrow 0$ for $s > 0$. Lemma 8 applies, since $0 \leq sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) < 1$ for $s > 0$. Note that the Stieltjes transform is positive for positive s, and thus the left-hand side of (238) would be larger than one, if $sG_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{R}},L,\beta_{\mathcal{S}})}(s) > 1$, cf. proof of [7, Theorem 4]). Thus

$$G_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(s) = \frac{1}{s}. \tag{243}$$

Since the Stieltjes transform is an analytic function on its full domain, (243) holds for all s. The corresponding asymptotic EED is $F_{\tilde{\mathbf{R}}_s}^{(\beta_{\mathcal{S}})}(x) = \sigma(x)$.

□

Proof of Proposition 3: The matrix $\mathbf{R}_{n,l}$ is defined as $\mathbf{R}_{n,l} \triangleq \frac{\alpha^l}{n_{\mathcal{R}}^l} \mathbf{H}_1 \cdots \mathbf{H}_l \mathbf{H}_l^H \cdots \mathbf{H}_1^H$, where $l \in \{1, \ldots, L\}$. The EED of $\mathbf{R}_{n,l}$ converges uniformly and almost surely to an asymptotic EED. This and the corresponding implicit equation for the Stieltjes transform of the asymptotic EED of $\mathbf{R}_{n,l}$ follows again from Lemma 7, where N is identified with l, $\mathbf{M}_n$ with $\mathbf{H}_l$ and k_n with $n_{\mathcal{R}}$. Thus, we obtain the following equation for the Stieltjes transform:

$$G_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(s) \cdot \left(\Psi_{\mathbf{R}_{n,l}} + s \right) = 1 \tag{244}$$

with

$$\Psi_{\mathbf{R}_{n,l}} = \left(\frac{sG_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(s) - 1}{\beta_{\mathcal{R}}} + 1 \right)^{l}. \tag{245}$$

Again, we assume $s > 0$. Once more, the obtained limiting Stieltjes transform generalizes to its full domain, since it is an analytic function. Since $0 < sG_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(s) < 1$, the following statements apply:

- For every fixed $\beta_{\mathcal{R}}$ and L, the following inequality holds for all $l \in \{1, \ldots, L\}$ and $s > 0$:

$$\left| \Psi_{\mathbf{R}_{n,l}} - 1 \right| < \left| \left(-\frac{1}{\beta_{\mathcal{R}}} + 1 \right)^{L} - 1 \right|. \tag{246}$$

- If $\beta_{\mathcal{R}} \in \Omega\left(L^{1+\varepsilon}\right)$, then

$$\lim_{L\longrightarrow\infty} \left(-\frac{1}{\beta_{\mathcal{R}}} + 1 \right)^{L} = 1. \tag{247}$$

This implies that the L terms *converge uniformly* as L and $\beta_{\mathcal{R}}$ tend to infinity:

$$\lim_{L\longrightarrow\infty} \max_{l\in\{1,\ldots,L\}} \left| \Psi_{\mathbf{R}_{n,l}} - 1 \right| = 0, \tag{248}$$

which in turn yields

$$\lim_{L\longrightarrow\infty} \max_{l\in\{1,\ldots,L\}} \left| G_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(s) - \frac{1}{s+1} \right| = 0. \tag{249}$$

In a next step, it is concluded that also $F_{\tilde{\mathbf{R}}_n}(x) = \sigma(x-1)$, if $\beta_{\mathcal{R}} \in \Omega\left(L^{1+\varepsilon}\right)$. To this end, the following Lemmata 9 and 10 are stated. The corresponding proofs are provided subsequently in this Appendix.

[8]The Stieltjes transform of some EED $F(\cdot)$ is defined as $G(s) \triangleq \int_{-\infty}^{\infty} (s+x)^{-1} \cdot dF(x)$[17]. This definition is adopted from [7] here, while it is generally more common to define the Stieltjes transform with a minus sign in the denominator. The EED is uniquely determined by its Stieltjes transform.

Lemma 9: Let $\mathbf{A}^{(n)}(\gamma) \in \mathbb{C}^{n\times n}$ be a sequence of positive semidefinite random matrices with parameter γ whose EED converges uniformly to the asymptotic EED $F_{\mathbf{A}}^{(\gamma)}(x)$ almost surely. Assume for all γ i) $\lim_{n\to\infty} \frac{1}{n}\mathrm{Tr}[\mathbf{A}(\gamma)] = 1$ and ii) $\lambda_{\max}^{(\gamma)} \triangleq \lim_{n\to\infty} \lambda_{\max}\{\mathbf{A}(\gamma)\} < \infty$ almost surely. Then, the following types of convergence are equivalent[9]:

1) $\lim_{\gamma\to\infty} d(\gamma) = 0$, where $d(\gamma)$ fulfills for every γ:

$$\lim_{n\to\infty} \frac{1}{n}\|\mathbf{I}_n - \mathbf{A}(\gamma)\|_* = d(\gamma) \text{ almost surely.}$$

2) $\lim_{\gamma\to\infty} |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)| = 0$ for all x.

3) $\lim_{\gamma\to\infty} \int_0^\infty |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)| \cdot dx = 0$.

Lemma 10: $\tilde{\mathbf{R}}_n$ fulfills the assumptions of Lemma 9 when the parameter γ is identified with L.

As a consequence of that, there exists for every $\varepsilon > 0$ an L_0, such that there exists for all $L \geq L_0$ almost surely an $n_{\mathcal{D}}^{(0)}(L_0)$, such that for all $n_{\mathcal{D}} \geq n_{\mathcal{D}}^{(0)}(L_0)$

$$\begin{aligned}
&\frac{1}{n_{\mathcal{D}}}\left\|\mathbf{I}_{n_{\mathcal{D}}} - \tilde{\mathbf{R}}_n\right\|_* \\
&= \frac{1}{n_{\mathcal{D}}}\left\|\mathbf{I}_{n_{\mathcal{D}}} - \frac{1-\alpha}{1-\alpha^{L+1}}\cdot \mathbf{R}_n\right\|_* && (250)\\
&= \frac{1}{n_{\mathcal{D}}}\left\|\frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\left(\alpha^l\cdot \mathbf{I}_{n_{\mathcal{D}}} - \mathbf{R}_{n,l}\right)\right\|_* && (251)\\
&= \frac{1}{n_{\mathcal{D}}}\frac{1-\alpha}{1-\alpha^{L+1}} \\
&\quad \times \left\|\sum_{l=0}^{L}\alpha^l\left(\mathbf{I}_{n_{\mathcal{D}}} - \frac{1}{n_{\mathcal{R}}^l}\mathbf{H}_1\cdots\mathbf{H}_l\mathbf{H}_l^H\cdots\mathbf{H}_1^H\right)\right\|_* && (252)\\
&\leq \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\frac{\alpha^l}{n_{\mathcal{D}}}\left\|\mathbf{I}_{n_{\mathcal{D}}} - \frac{1}{n_{\mathcal{R}}^l}\mathbf{H}_1\cdots\mathbf{H}_l\mathbf{H}_l^H\cdots\mathbf{H}_1^H\right\|_* && (253)\\
&= \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l\int_0^\infty\left|F_{\mathbf{R}_{n,l}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},l)}(x) - \sigma(x-1)\right|dx && (254)\\
&\leq \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l\left(\int_0^\infty\left|F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x) - \sigma(x-1)\right|dx\right. \\
&\quad \left. + \int_0^\infty\left|F_{\mathbf{R}_{n,l}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},l)}(x) - F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x)\right|dx\right) && (255)\\
&\leq \max_{l\in\{1,\ldots,L\}}\left\{\int_0^\infty\left|F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x) - \sigma(x-1)\right|dx\right\} \\
&\quad + \max_{l\in\{1,\ldots,L\}}\left\{\int_0^\infty\left|F_{\mathbf{R}_{n,l}}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},l)}(x) - F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x)\right|dx\right\} && (256)\\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. && (257)
\end{aligned}$$

Expression (253) is obtained by applying the triangle inequality and using the homogeneity of the nuclear norm. Equality between (253) and (254) is established by the following chain of identities for a positive semidefinite matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$:

$$\begin{aligned}
\frac{1}{n}\|\mathbf{I}_n - \mathbf{A}(\gamma)\|_* &= \frac{1}{n}\sum_{i=1}^{n}\sigma_i\{\mathbf{I}_n - \mathbf{A}(\gamma)\} \\
&= \frac{1}{n}\sum_{i=1}^{n}|\lambda_i\{\mathbf{I}_n - \mathbf{A}(\gamma)\}| && (258)\\
&= \frac{1}{n}\sum_{i=1}^{n}|1 - \lambda_i\{\mathbf{A}(\gamma)\}| && (259)\\
&= \frac{1}{n}\sum_{i:\lambda_i\{\mathbf{A}(\gamma)\}\leq 1}(1 - \lambda_i\{\mathbf{A}(\gamma)\}) \\
&\quad + \frac{1}{n}\sum_{i:\lambda_i\{\mathbf{A}(\gamma)\}>1}(\lambda_i\{\mathbf{A}(\gamma)\} - 1) && (260)\\
&= \int_0^1 |F_{\mathbf{A}}^{(n,\gamma)}(x)|dx \\
&\quad + \int_1^\infty |F_{\mathbf{A}}^{(n,\gamma)}(x) - 1|dx && (261)\\
&= \int_0^\infty |F_{\mathbf{A}}^{(n,\gamma)}(x) - \sigma(x-1)|dx. && (262)
\end{aligned}$$

Equation (255) follows by adding and substracting $F_{\mathbf{R}_{n,l}}^{(\beta_{\mathcal{R}},l)}(x)$ and repeated application of the triangle inequality. In (256), the individual integrals are upper-bounded by the largest ones. In the final step, both terms in (256) are upper-bounded by $\varepsilon/2$. For the first term, one can fix an L_0, such that this upper-bound holds for all $L \geq L_0$ by (249) and Lemma 9. For fixed L (and thus $\beta_{\mathcal{R}}$), one can finally choose an $n_{\mathcal{D}}^{(0)}(L)$ large enough, such that for all $n_{\mathcal{D}} \geq n_{\mathcal{D}}^{(0)}(L)$ the second term is smaller than $\varepsilon/2$. Note that $\lim_{n_{\mathcal{D}}\to\infty} \lambda_{\max}\{\tilde{\mathbf{R}}_n\} < \infty$ for fixed L almost surely according to Lemma 10. Therefore, the limit can be taken inside the integral due to the uniform convergence of $F_{\tilde{\mathbf{R}}_n}^{(n_{\mathcal{D}},\beta_{\mathcal{R}},L)}$ to $F_{\tilde{\mathbf{R}}_n}^{(\beta_{\mathcal{R}},L)}$). We have thus shown that $F_{\tilde{\mathbf{R}}_n}^{(\beta_{\mathcal{R}},L)}(x)$ converges pointwise to $\sigma(x-1)$ as $L\to\infty$. Since the eigenvalues of the inverse of $\tilde{\mathbf{R}}_n$ are the inverse eigenvalues of $\tilde{\mathbf{R}}_n$, i.e., $\lambda_k\{\tilde{\mathbf{R}}_n^{-1}\} = \lambda_k^{-1}\{\tilde{\mathbf{R}}_n\}$, one can conclude that also $F_{\tilde{\mathbf{R}}_n^{-1}}^{(\beta_{\mathcal{R}},L)}(x)$ converges pointwise to $\sigma(x-1)$. □

Proof of Lemma 4: We define the matrices

$$\mathbf{A}_l \triangleq \left(\frac{\alpha}{n_{\mathcal{R}}}\right)^{L-l+2}\mathbf{H}_l\cdots\mathbf{H}_L\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\mathbf{H}_L^H\cdots\mathbf{H}_l^H \quad (263)$$

and

$$\begin{aligned}
\mathbf{B}_l \triangleq \mathbf{I}_{n_{\mathcal{R}}} &+ \frac{\alpha}{n_{\mathcal{R}}}\mathbf{H}_l\left(\mathbf{I}n_{\mathcal{R}} + \cdots\right.\\
&\left. + \frac{\alpha}{n_{\mathcal{R}}}\mathbf{H}_{L_0-1}\left(\mathbf{I}_{n_{\mathcal{R}}} + \frac{\alpha}{n_{\mathcal{R}}}\mathbf{H}_{L_0}\mathbf{H}_{L_0}^H\right)\mathbf{H}_{L_0-1}^H\cdots\right)\mathbf{H}_l^H.
\end{aligned} \quad (264)$$

Note that $\sigma_{\mathrm{w}}^2 \cdot (\mathbf{I}_{n_{\mathcal{D}}} + \mathbf{H}_1\mathbf{B}_2\mathbf{H}_1^H)$ corresponds to the noise covariance matrix at the destination antennas under the assumption

[9] $\|\mathbf{M}\|_* = \sum_i \sigma_i\{\mathbf{M}\}$ denotes the nuclear norm.

of noiseless relay stages $\mathcal{R}_L$ to $\mathcal{R}_{L-L_0+1}$. Therefore, we obtain the following upper-bound:

$$\begin{aligned}
&\mathsf{E}\left[R_L^{\mathrm{AF}}\right]\\
&\leq \mathsf{E}\left[\log\det\left(\mathbf{I}_{\mathbf{n}_\mathcal{D}}+\frac{P_L}{\sigma_\mathrm{w}^2 n_\mathcal{S}}\mathbf{H}_1\mathbf{A}_2\mathbf{H}_1^H \times\left(\mathbf{I}_{n_\mathcal{D}}+\mathbf{H}_1\mathbf{B}_2\mathbf{H}_1^H\right)^{-1}\right)\right]\\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{P_L}{\sigma_\mathrm{w}^2 n_\mathcal{S}}\mathbf{U}_1\tilde{\mathbf{S}}_1\tilde{\mathbf{V}}_1^H\mathbf{A}_2\tilde{\mathbf{V}}_1\tilde{\mathbf{S}}_1\mathbf{U}_1^H \times\left(\mathbf{I}_{n_\mathcal{D}}+\mathbf{U}_1\tilde{\mathbf{S}}_1\tilde{\mathbf{V}}_1^H\mathbf{B}_2\tilde{\mathbf{V}}_1\tilde{\mathbf{S}}_1\mathbf{U}_1^H\right)^{-1}\right)\right] &(265)\\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{P_L}{\sigma_\mathrm{w}^2 n_\mathcal{S}}\tilde{\mathbf{V}}_1^H\mathbf{A}_2\tilde{\mathbf{V}}_1\tilde{\mathbf{S}}_1^2 \times\left(\mathbf{I}_{n_\mathcal{D}}+\tilde{\mathbf{V}}_1^H\mathbf{B}_2\tilde{\mathbf{V}}_1\tilde{\mathbf{S}}_1^2\right)^{-1}\right)\right] &(266)\\
&\leq \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{P_L}{\sigma_\mathrm{w}^2 n_\mathcal{S}}\tilde{\mathbf{V}}_1^H\mathbf{A}_2\tilde{\mathbf{V}}_1\frac{\mathrm{Tr}[\tilde{\mathbf{S}}_1^2]}{n_\mathcal{R}} \times\left(\mathbf{I}_{n_\mathcal{D}}+\tilde{\mathbf{V}}_1^H\mathbf{B}_2\tilde{\mathbf{V}}_1\frac{\mathrm{Tr}[\tilde{\mathbf{S}}_1^2]}{n_\mathcal{R}}\right)^{-1}\right)\right] &(267)\\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{s_1\alpha P_L}{\sigma_\mathrm{w}^2 n_\mathcal{S}}\tilde{\mathbf{H}}_2\mathbf{A}_3\tilde{\mathbf{H}}_2^H \times\left((1+s_1\alpha)\mathbf{I}_{n_\mathcal{D}}+s_1\alpha\tilde{\mathbf{H}}_2\mathbf{B}_3\tilde{\mathbf{H}}_2^H\right)^{-1}\right)\right]. &(268)
\end{aligned}$$

In (265), we use the singular value decomposition $\mathbf{H}_1 = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^H = \mathbf{U}_1\tilde{\mathbf{S}}_1\tilde{\mathbf{V}}_1^H$, where $\tilde{\mathbf{S}}_1$ is an $n_\mathcal{D}\times n_\mathcal{D}$ diagonal matrix with all nonzero singular values on its diagonal and $\tilde{\mathbf{V}}_1^H$ an $n_\mathcal{D}\times n_\mathcal{R}$ matrix derived from $\tilde{\mathbf{V}}_1$ by deleting the rows that correspond to the zero singular values.

Equation (266) follows from the following chain of identities, where we define $\mathbf{X}\triangleq\mathbf{U}_1\tilde{\mathbf{S}}_1$:

$$\begin{aligned}
&\mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\left(\mathbf{I}_{n_\mathcal{D}}+\mathbf{X}\tilde{\mathbf{B}}_2\mathbf{X}^H\right)^{-1}\\
&\quad=\mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\mathbf{X}^{-1}\left(\mathbf{I}_{n_\mathcal{D}}+\mathbf{X}\tilde{\mathbf{B}}_2\mathbf{X}^H\right)^{-1} &(269)\\
&\quad=\mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\left(\mathbf{X}+\mathbf{X}\tilde{\mathbf{B}}_2\mathbf{X}^H\mathbf{X}\right)^{-1} &(270)\\
&\quad=\mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\left(\mathbf{I}_{n_\mathcal{D}}+\tilde{\mathbf{B}}_2\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^{-1} &(271)
\end{aligned}$$

and the fact that

$$\lambda_i\left\{\mathbf{X}\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\left(\mathbf{I}_{n_\mathcal{D}}+\tilde{\mathbf{B}}_2\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^{-1}\right\} = \lambda_i\left\{\tilde{\mathbf{A}}_2\mathbf{X}^H\mathbf{X}\left(\mathbf{I}_{n_\mathcal{D}}+\tilde{\mathbf{B}}_2\mathbf{X}^H\mathbf{X}\right)^{-1}\mathbf{X}^{-1}\mathbf{X}\right\}. \quad (272)$$

Equation (267) follows from the fact that $\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\tilde{\mathbf{A}}_2\tilde{\mathbf{S}}_1(\mathbf{I}_{n_\mathcal{D}}+\tilde{\mathbf{B}}_2\tilde{\mathbf{S}}_1)^{-1}\right)$ is concave on the set of positive definite matrices $\tilde{\mathbf{S}}_1$ according to Lemma 11 on p. 42, and, thus, for a given trace maximized by the respective matrix proportional to the identity [1]. In (268), we introduce the matrix $\tilde{\mathbf{H}}_2=\tilde{\mathbf{V}}_1^H\mathbf{H}_2$, which fulfills

$$\mathsf{E}\left[\mathrm{Tr}\left[\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H\right]\right]=\frac{1}{n_\mathcal{R}}\mathsf{E}\left[\mathrm{Tr}\left[\mathbf{H}_2\mathbf{H}_2^H\right]\right]=\mathsf{E}\left[\mathrm{Tr}\left[\mathbf{H}_1\mathbf{H}_1^H\right]\right].$$

We finally obtain by induction

$$\begin{aligned}
&\mathsf{E}\left[R_L^{\mathrm{AF}}\right]\\
&\leq \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{s_1s_2\alpha P_L}{\sigma_\mathrm{w}^2 n_\mathcal{S}}\tilde{\mathbf{V}}_2^H\mathbf{A}_3\tilde{\mathbf{V}}_2 \times\left((1+s_1\alpha)\mathbf{I}_{n_\mathcal{D}}+s_1s_2\alpha^2\tilde{\mathbf{V}}_2^H\mathbf{B}_3\tilde{\mathbf{V}}_2\right)^{-1}\right)\right] &(273)\\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{s_1s_2\alpha^2 P_L}{\sigma_\mathrm{w}^2 n_\mathcal{S}}\tilde{\mathbf{H}}_3\mathbf{A}_4\tilde{\mathbf{H}}_3^H \times\left((1+s_1\alpha+s_1s_2\alpha^2)\mathbf{I}_{n_\mathcal{D}}+s_1s_2\alpha^2\tilde{\mathbf{H}}_3\mathbf{B}_4\tilde{\mathbf{H}}_3^H\right)^{-1}\right)\right] &(274)\\
&\ \vdots\\
&\leq \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{s_1\cdots s_{L_0}\alpha^{L_0-1}P_L}{\sigma_\mathrm{w}^2\cdot n_\mathcal{S}}\tilde{\mathbf{V}}_{L_0}^H\mathbf{A}_{L_0+1}\tilde{\mathbf{V}}_{L_0} \times\left(\sum_{l=0}^{L_0}s_1\cdots s_l\alpha^l\cdot\mathbf{I}_{n_\mathcal{D}}\right)^{-1}\right)\right] &(275)\\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{P_L}{\sigma_\mathrm{w}^2\cdot n_\mathcal{S}n_\mathcal{R}^{L-L_0}}\frac{s_1\cdots s_{L_0}\alpha^{L_0-1}}{\sum_{l=0}^{L_0}s_1\cdots s_l\alpha^l} \times\tilde{\mathbf{V}}_{L_0}^H\mathbf{A}_{L_0+1}\tilde{\mathbf{V}}_{L_0}\right)\right] &(276)\\
&= \mathsf{E}\left[\log\det\left(\mathbf{I}_{n_\mathcal{D}}+\frac{P_L}{\sigma_\mathrm{w}^2\cdot n_\mathcal{S}n_\mathcal{R}^{L-L_0}}\frac{s_1\cdots s_{L_0}\alpha^{L}}{\sum_{l=0}^{L_0}s_1\cdots s_l\alpha^l} \times\tilde{\mathbf{V}}_{L_0}^H\mathbf{H}_{L_0+1}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{L_0+1}^H\tilde{\mathbf{V}}_{L_0}\right)\right]. &(277)
\end{aligned}$$

□

Proof of Lemma 5: We prove the statement by showing that the S-transforms of the asymptotic EEDs of $\mathbf{A}$ and $\mathbf{B}$ coincide. The S-transform of $\mathbf{B}$ is given by [7, Theorem 2]

$$S_\mathbf{B}(z)=\begin{cases}\dfrac{\frac{n_\mathcal{D}}{n_\mathcal{S}}}{\left(1+\frac{n_\mathcal{D}}{n_\mathcal{S}}z\right)}, & \text{if } L=1,\\[2ex] \left(\dfrac{1}{1+\frac{n_\mathcal{D}}{n_\mathcal{R}}z}\right)^{L-2}\dfrac{\frac{n_\mathcal{D}}{n_\mathcal{S}}}{\left(1+\frac{n_\mathcal{D}}{n_\mathcal{S}}z\right)\left(\frac{n_\mathcal{D}}{n_\mathcal{R}}z+1\right)}, & \text{if } L\geq 2.\end{cases} \quad (278)$$

We derive the S-transform of $\mathbf{A}$ in the following. First, we obtain the S-transform of $\tilde{\mathbf{A}}\triangleq\frac{1}{n_\mathcal{R}^L}\mathbf{H}_2\mathbf{H}_3\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_3^H\mathbf{H}_2^H$ from [7, Theorem 2] as

$$S_{\tilde{\mathbf{A}}}(z)=\left(\frac{1}{1+z}\right)^{L-1}\frac{1}{\frac{n_\mathcal{S}}{n_\mathcal{R}}+z}. \quad (279)$$

Using that the S-transform of $\frac{1}{n_{\mathcal{D}}}\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1$ is given by $S_{\frac{1}{n_{\mathcal{D}}}\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1}(z)=\frac{z+1}{1+n_{\mathcal{R}}/n_{\mathcal{D}}z}$ (e.g., [17]) we obtain the S-transform of $\tilde{\tilde{\mathbf{A}}} \triangleq \frac{1}{n_{\mathcal{R}}^L n_{\mathcal{D}}}\mathbf{H}_2\mathbf{H}_3\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_3^H\mathbf{H}_2^H\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1$ as [7, Theorem 1]

$$S_{\tilde{\tilde{\mathbf{A}}}}(z)=S_{\frac{1}{n_{\mathcal{D}}}\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1}(z)S_{\tilde{\mathbf{A}}}(z) \tag{280}$$

$$=\frac{z+1}{1+\frac{n_{\mathcal{R}}}{n_{\mathcal{D}}}z}\left(\frac{1}{1+z}\right)^{L-1}\frac{1}{\frac{n_{\mathcal{S}}}{n_{\mathcal{R}}}+z}. \tag{281}$$

Finally, we obtain the S-transform of the asymptotic EED of $\mathbf{A}$ as [7, Eq. (15)] as

$$S_{\mathbf{A}}(z)=\frac{z+1}{z+\frac{n_{\mathcal{R}}}{n_{\mathcal{D}}}}S_{\tilde{\tilde{\mathbf{A}}}}\left(\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}z\right) \tag{282}$$

$$=\begin{cases}\frac{\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}}{1+\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}z}, & \text{if } L=1,\\ \left(\frac{1}{1+\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}z}\right)^{L-2}\frac{\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}}{\left(1+\frac{n_{\mathcal{D}}}{n_{\mathcal{S}}}z\right)\left(\frac{n_{\mathcal{D}}}{n_{\mathcal{R}}}z+1\right)}, & \text{if } L\geq 2\end{cases} \tag{283}$$

which coincides with (278). □

Proof of Lemma 6: The proof relies on a result of [25], [26] (also [17, Theorem 2.39]): Let $\mathbf{Y}=\frac{1}{k}\mathbf{X}\mathbf{T}\mathbf{X}^H$ be a random matrix, where:

- $\mathbf{T}$ is an arbitrary nonnegative definite $k\times k$ random matrix whose EED converges uniformly to a nonrandom distribution almost surely as $k\longrightarrow\infty$ and satisfies $\lim_{k\longrightarrow\infty}k^{-1}\mathrm{Tr}[\mathbf{T}]=T$ almost surely;
- $\mathbf{X}$ is an $n\times k$ random matrix, independent of $\mathbf{T}$, with i.i.d. elements of unit variance.

Then, for every fixed ratio k/n

$$\lim_{n\longrightarrow\infty}\frac{1}{n}\mathrm{Tr}[\mathbf{Y}]=\lim_{n\longrightarrow\infty}\frac{1}{k}\mathrm{Tr}[\mathbf{T}]=T \text{ almost surely.} \tag{284}$$

Since the trace of a product is invariant under cyclic permutation of the factors, we can write

$$n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}^2}\mathbf{H}_1\mathbf{H}_2\mathbf{H}_2^H\mathbf{H}_1^H\right]=n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_2^H\left(\frac{1}{n_{\mathcal{R}}}\mathbf{H}_1^H\mathbf{H}_1\right)\mathbf{H}_2\right]. \tag{285}$$

If $n_{\mathcal{R}}^{-1}\mathbf{H}_1^H\mathbf{H}_1$ is identified with $\mathbf{T}$, where $n=k=n_{\mathcal{R}}$ and $T=n_{\mathcal{D}}/n_{\mathcal{R}}$, we obtain

$$\lim_{n_{\mathcal{D}}\to\infty}n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_2^H\left(\frac{1}{n_{\mathcal{R}}}\mathbf{H}_1^H\mathbf{H}_1\right)\mathbf{H}_2\right]=\lim_{n_{\mathcal{D}}\to\infty}n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_1^H\mathbf{H}_1\right]=1 \text{ almost surely.} \tag{286}$$

Repeating the argument for

$$n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}^3}\mathbf{H}_1\mathbf{H}_2\mathbf{H}_3\mathbf{H}_3^H\mathbf{H}_2^H\mathbf{H}_1^H\right]=n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_3^H\left(\frac{1}{n_{\mathcal{R}}^2}\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2\right)\mathbf{H}_3\right] \tag{287}$$

with $\mathbf{T}=n_{\mathcal{R}}^{-2}\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2$ yields

$$\lim_{n_{\mathcal{D}}\longrightarrow\infty}n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_3^H\left(\frac{1}{n_{\mathcal{R}}^2}\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2\right)\mathbf{H}_3\right]=\lim_{n_{\mathcal{D}}\longrightarrow\infty}n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}^2}\mathbf{H}_2^H\mathbf{H}_1^H\mathbf{H}_1\mathbf{H}_2\right] \tag{288}$$

$$=\lim_{n_{\mathcal{D}}\longrightarrow\infty}n_{\mathcal{D}}^{-1}\mathrm{Tr}\left[\frac{1}{n_{\mathcal{R}}}\mathbf{H}_1^H\mathbf{H}_1\right]=1 \text{ almost surely.} \tag{289}$$

Finally, we obtain with $\mathbf{T}=n_{\mathcal{R}}^{-L}\mathbf{H}_L^H\cdots\mathbf{H}_1^H\mathbf{H}_1\cdots\mathbf{H}_L$

$$\lim_{n_{\mathcal{D}}\to\infty}n_{\mathcal{D}}^{-1}\times\mathrm{Tr}\left[\frac{1}{n_{\mathcal{S}}}\mathbf{H}_{L+1}^H\left(\frac{1}{n_{\mathcal{R}}^L}\mathbf{H}_L^H\cdots\mathbf{H}_1^H\mathbf{H}_1\cdots\mathbf{H}_L\right)\mathbf{H}_{L+1}\right]=1 \text{ almost surely} \tag{290}$$

which completes the proof. □

Proof of Lemma 8: The lemma follows from the fact that the limit can be taken inside a continuous function, which allows us to write

$$\lim_{\kappa\longrightarrow\infty}\left(\frac{c}{g(\kappa)}+1\right)^{\kappa}=\exp\left(\lim_{\kappa\longrightarrow\infty}\mathrm{Log}\left(\frac{c}{g(\kappa)}+1\right)^{\kappa}\right) \tag{291}$$

$$=\exp\left(\lim_{\kappa\longrightarrow\infty}\kappa\log\left|\frac{c}{g(\kappa)}+1\right|+\lim_{\kappa\longrightarrow\infty}\imath\arg\left\{\frac{c}{g(\kappa)}+1\right\}\right) \tag{292}$$

$$=\exp\left(\lim_{\kappa\longrightarrow\infty}\kappa\log\left|\frac{c}{g(\kappa)}+1\right|\right) \tag{293}$$

$$=\exp\left(\lim_{\kappa\longrightarrow\infty}\Re\left\{\kappa\mathrm{Log}\left(\frac{c}{g(\kappa)}+1\right)\right\}\right). \tag{294}$$

From the rule of Bernoulli-l'Hospital, we know that

$$\lim_{\kappa\longrightarrow\infty}\kappa\cdot\mathrm{Log}\left(\frac{c}{M\kappa^{\gamma}}+1\right)=\lim_{\kappa\longrightarrow\infty}\frac{c\gamma\kappa}{c+M\kappa^{\gamma}} \tag{295}$$

where γ and M are positive constants.

If $g(\kappa)\in\Omega(\kappa^{1+\varepsilon})$, there exists by definition some $M>0$, such that the absolute value of the argument of the exponential function in (294) can be upper-bounded according to

$$\lim_{\kappa\to\infty}\left|\kappa\cdot\mathrm{Log}\left(\frac{c}{g(\kappa)}+1\right)\right|\leq\lim_{\kappa\to\infty}\left|\kappa\cdot\mathrm{Log}\left(\frac{c}{M\kappa^{1+\varepsilon}}+1\right)\right|. \tag{296}$$

Evaluating (295) for $\gamma=1+\varepsilon$ renders this upper-bound zero, which establishes that also the left-hand side of (296) becomes zero and (294) evaluates to one in this case.

For the proof of the other two cases (note that c is real and negative then), we write analogously to (294)

$$\lim_{\kappa\longrightarrow\infty}\left(\frac{c}{g(\kappa)}+1\right)^{\kappa}=\exp\left(\lim_{\kappa\longrightarrow\infty}\kappa\log\left(\frac{c}{g(\kappa)}+1\right)\right). \tag{297}$$

If $g(\kappa) \in \mathcal{O}(\kappa^{1-\varepsilon})$, there exists some $M > 0$, such that the argument of the exponential function in (297) can be upper-bounded according to

$$\lim_{\kappa \longrightarrow \infty} \kappa \cdot \log\left(\frac{c}{g(\kappa)} + 1\right) \leq \lim_{\kappa \longrightarrow \infty} \kappa \cdot \log\left(\frac{c}{M\kappa^{1-\varepsilon}} + 1\right). \tag{298}$$

The limit (295) does not exist for $\gamma = 1 - \varepsilon$ implying that both sides of (298) evaluate to minus infinity, and (294) to zero in turn. □

Proof of Lemma 9: We establish the following chain of identities:

$$\lim_{n \longrightarrow \infty} \frac{1}{n} \left\| \mathbf{I}_n - \mathbf{A}^{(n)}(\gamma) \right\|_*$$

$$= \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=1}^{n} \sigma_i\{\mathbf{I}_n - \mathbf{A}^{(n)}(\gamma)\} \tag{299}$$

$$= \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=1}^{n} \left| \lambda_i\{\mathbf{I}_n - \mathbf{A}^{(n)}(\gamma)\} \right| \tag{300}$$

$$= \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=1}^{n} \left| 1 - \lambda_i\{\mathbf{A}^{(n)}(\gamma)\} \right| \tag{301}$$

$$= \lim_{n \longrightarrow \infty} \frac{1}{n} \left[\sum_{i:\lambda_i\{\mathbf{A}^{(n)}(\gamma)\} \leq 1} \left(1 - \lambda_i\{\mathbf{A}^{(n)}(\gamma)\}\right) + \sum_{i:\lambda_i\{\mathbf{A}^{(n)}(\gamma)\} > 1} \left(\lambda_i\{\mathbf{A}^{(\gamma)}\} - 1\right) \right] \tag{302}$$

$$= \lim_{n \longrightarrow \infty} \left[\int_0^1 \left| F_{\mathbf{A}}^{(n,\gamma)}(x) \right| \cdot dx + \int_1^{\infty} \left| F_{\mathbf{A}}^{(n,\gamma)}(x) - 1 \right| \cdot dx \right] \tag{303}$$

$$= \lim_{n \longrightarrow \infty} \int_0^{\infty} \left| F_{\mathbf{A}}^{(n,\gamma)}(x) - \sigma(x-1) \right| \cdot dx \tag{304}$$

$$= \int_0^{\infty} \left| F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1) \right| \cdot dx \text{ almost surely.} \tag{305}$$

In (300), we write the nuclear norm of the matrix $\mathbf{I}_n - \mathbf{A}$ in terms of its singular values σ_i. Since the matrix $\mathbf{I}_n - \mathbf{A}^{(\gamma)}$ is normal, i.e., $(\mathbf{I}_n - \mathbf{A}^{(\gamma)})^H(\mathbf{I}_n - \mathbf{A}^{(\gamma)}) = (\mathbf{I}_n - \mathbf{A}^{(\gamma)})(\mathbf{I}_n - \mathbf{A}^{(\gamma)})^H$, its singular values coincide with the absolute values of its eigenvalues. In (302), we arrange the terms, such that they can be related to the EED of $\mathbf{A}^{(\gamma)}$. In (305), we take the limit inside the integral. This is justified, since the maximum eigenvalue of $\mathbf{A}^{(\gamma)}$ converges to some bounded constant by assumption. Thus, the integration is over the compact interval $[0, \lambda_{\max}^{(\gamma)}]$, where we integrate over a uniformly convergent sequence (in n) of functions. This establishes the equivalence between 1) and 3). It remains to establish that

$$\lim_{\gamma \longrightarrow \infty} \int_0^{\lambda_{\max}^{(\gamma)}} |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)| \cdot dx = 0$$
$$\Longleftrightarrow \lim_{\gamma \longrightarrow \infty} |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)| = 0 \text{ for all } x. \tag{306}$$

For the forward part, consider $\epsilon(x) \triangleq |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)|$ for $x \in [0, 1)$, i.e., $\epsilon(x) = |F_{\mathbf{A}}^{(\gamma)}(x)|$. Fix any $\Delta \in [-1, 0)$. Since $\epsilon(x)$ is monotonically increasing on the interval of interest, we can write

$$\int_{1+\Delta}^{1} |F_{\mathbf{A}}^{(\gamma)}(x) - \sigma(x-1)| \cdot dx > |\Delta| \cdot \epsilon(1 + \Delta). \tag{307}$$

Thus, if $\epsilon(1 + \Delta)$ does not tend to zero for all Δ, the integral cannot tend to zero. The same reasoning can be applied for the interval $\Delta \in [1, \lambda_{\max}^{(\gamma)}]$.

For the backward part, we break the integration in (305) into two parts. The first integral is from zero to some constant d, $1 < d < \lambda_{\max}^{(\gamma)}$. $F_{\mathbf{A}}^{(\gamma)}(x)$ is a sequence of Riemann integrable functions that is uniformly bounded and pointwise convergent. By the bounded convergence theorem for the Riemann integral (e.g., [24]) we can take the limit inside the integral. Thus, the limit of this first integral is zero. The second part of the integral is from d to $\lambda_{\max}^{(\gamma)}$. Here, the limit cannot be taken inside the integral in general (note that $\lim_{\gamma \longrightarrow \infty} \lambda_{\max}^{(\gamma)}$ might be unbounded). However, we can write

$$\lim_{\gamma \longrightarrow \infty} \int_d^{\lambda_{\max}^{(\gamma)}} 1 - F_{\mathbf{A}}^{(\gamma)}(x) \cdot dx$$
$$= \lim_{\gamma \longrightarrow \infty} \int_0^{\lambda_{\max}^{(\gamma)}} 1 - F_{\mathbf{A}}^{(\gamma)}(x) \cdot dx - \lim_{\gamma \longrightarrow \infty} \int_0^{d} 1 - F_{\mathbf{A}}^{(\gamma)}(x) \cdot dx = 0. \tag{308}$$

The second term on the right-hand side of (308) evaluates to one, since the limit can be taken inside the integral. Again, this is justified by the bounded convergence theorem (e.g., [24]). The first integral on the right-hand side evaluates to one for every γ, since otherwise there would be a contradiction between the following two statements:

$$\lim_{n \longrightarrow \infty} \int_0^{\lambda_1\{\mathbf{A}(\gamma)\}} 1 - F_{\mathbf{A}}^{(n,\gamma)}(x) \cdot dx = 1 \text{ almost surely for each } \gamma \tag{309}$$

$$\lim_{n \longrightarrow \infty} \int_0^{\lambda_1\{\mathbf{A}(\gamma)\}} 1 - F_{\mathbf{A}}^{(n,\gamma)}(x) \cdot dx = \int_0^{\lambda_{\max}^{(\gamma)}} 1 - F_{\mathbf{A}}^{(\gamma)}(x) \cdot dx \text{ almost surely for each } \gamma. \tag{310}$$

The first statement corresponds to the assumption $\lim_{n \longrightarrow \infty} n^{-1} \mathrm{Tr}[\mathbf{A}(\gamma)] = 1$ almost surely for each γ, the second statement follows by taking the limit inside the integral, which is fine for every fixed value of $\lambda_{\max}^{(\gamma)}$. □

Proof of Lemma 10: We go through each of the assumptions required for Lemma 9:

1) To show that the EED of $\tilde{\mathbf{R}}_n$ converges uniformly to a nonrandom distribution irrespective of the values of L and $\beta_{\mathcal{R}}$, we use a result from [25], [26] (also [17, Theorem 2.39]):

Let $\mathbf{Y} = \frac{1}{k}\mathbf{X}\mathbf{T}\mathbf{X}^H$, where

- $\mathbf{T}$ is an arbitrary nonnegative definite $k \times k$ random matrix whose EED for every fixed ratio k/n converges uniformly to a nonrandom distribution almost surely as $n \longrightarrow \infty$;

- $\mathbf{X}$ is an $n \times k$ random matrix, independent of $\mathbf{T}$, with i.i.d. elements of unit variance.

Then, the EED of $\mathbf{Y}$ converges uniformly to a nonrandom distribution function almost surely as $n \longrightarrow \infty$. If $\mathbf{X}$ is identified with $\mathbf{H}_{l+1}$ and $\mathbf{T}$ with

$$\mathbf{I}_{n_\mathcal{D}} + \frac{\alpha}{n_\mathcal{R}}\mathbf{H}_l \times \left(\mathbf{I}_{n_\mathcal{D}} + \cdots + \frac{\alpha}{n_\mathcal{R}}\mathbf{H}_2 \times \left(\mathbf{I}_{n_\mathcal{D}} + \frac{\alpha}{n_\mathcal{R}}\mathbf{H}_1\mathbf{H}_1^H\right)\mathbf{H}_2^H \cdots\right)\mathbf{H}_l^H \tag{311}$$

uniform and almost sure convergence of the EED of $\tilde{\mathbf{R}}_n$ follows by induction.

2) For the trace condition, we obtain for arbitrary L and $\beta_\mathcal{R}$ by application of Lemma 6:

$$\lim_{n_\mathcal{D}\longrightarrow\infty}\frac{1}{n_\mathcal{D}}\mathrm{Tr}\left[\frac{1-\alpha}{1-\alpha^{L+1}}\cdot\mathbf{R}_n\right] = \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l \times \lim_{n_\mathcal{D}\longrightarrow\infty}\frac{1}{n_\mathcal{D}}\mathrm{Tr}\left[\frac{1}{n_\mathcal{R}^l}\mathbf{H}_1\cdots\mathbf{H}_l\mathbf{H}_l^H\cdots\mathbf{H}_1^H\right] \tag{312}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l = 1 \text{ almost surely.} \tag{313}$$

Note that $\beta_\mathcal{R}$ and L are fixed here, and thus the sum is finite.

3) For the condition on the maximum eigenvalue, we use the triangle inequality and the submultiplicativity of the spectral norm:

$$\lim_{n_\mathcal{D}\longrightarrow\infty}\lambda_{\max}\left\{\frac{1-\alpha}{1-\alpha^{L+1}}\mathbf{R}_n\right\} = \frac{1-\alpha}{1-\alpha^{L+1}}\times \lim_{n_\mathcal{D}\longrightarrow\infty}\lambda_{\max}\left\{\sum_{l=0}^{L}\alpha^l\frac{1}{n_\mathcal{R}^l}\mathbf{H}_1\cdots\mathbf{H}_l\mathbf{H}_l^H\cdots\mathbf{H}_1^H\right\} \tag{314}$$

$$\leq \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l\times \lim_{n_\mathcal{D}\longrightarrow\infty}\lambda_{\max}\left\{\frac{1}{n_\mathcal{R}^l}\mathbf{H}_1\cdots\mathbf{H}_l\mathbf{H}_l^H\cdots\mathbf{H}_1^H\right\} \tag{315}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l\times \lim_{n_\mathcal{D}\longrightarrow\infty}\lambda_{\max}\left\{\frac{1}{n_\mathcal{R}^l}\mathbf{H}_2\cdots\mathbf{H}_l\mathbf{H}_l^H\cdots\mathbf{H}_1^H\mathbf{H}_1\right\} \tag{316}$$

$$\leq \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l \times \lim_{n_\mathcal{D}\longrightarrow\infty}\lambda_{\max}\left\{\frac{1}{n_\mathcal{R}^{l-1}}\mathbf{H}_2\cdots\mathbf{H}_l\mathbf{H}_l^H\cdots\mathbf{H}_2^H\right\} \times \lim_{n_\mathcal{D}\longrightarrow\infty}\lambda_{\max}\left\{\frac{1}{n_\mathcal{R}}\mathbf{H}_1\mathbf{H}_1^H\right\}, \tag{317}$$

where we used that the nonzero eigenvalues of the matrix products in (315) and (316) coincide. Repeated application of steps (316) and (317) finally yields

$$\lim_{n_\mathcal{D}\longrightarrow\infty}\lambda_{\max}\left\{\frac{1-\alpha}{1-\alpha^{L+1}}\mathbf{R}_n\right\} \leq \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l\prod_{l'=1}^{l}\lambda_{\max}\left\{\frac{1}{n_\mathcal{R}}\mathbf{H}_{l'}\mathbf{H}_{l'}^H\right\} \tag{318}$$

$$= \frac{1-\alpha}{1-\alpha^{L+1}}\sum_{l=0}^{L}\alpha^l\cdot 4^{l-1}\left(1+\sqrt{\frac{n_\mathcal{D}}{n_\mathcal{R}}}\right)^2 \tag{319}$$

$$< \infty \text{ almost surely} \tag{320}$$

where each of the multipliers in (318) is bounded according to [20]. The ratio $n_\mathcal{D}/n_\mathcal{R}$ and L are fixed, such that also the resulting product is finite.

□

Lemma 11: Let $\mathbf{A}$ and $\mathbf{B}$ be positive definite matrices. Then, the function

$$f(\mathbf{Q}) = \log\det\left(\mathbf{I}+\mathbf{A}\mathbf{Q}(\mathbf{I}+\mathbf{B}\mathbf{Q})^{-1}\right) \tag{321}$$

is concave on the set of positive semidefinite matrices $\mathbf{Q}$.

Proof of Lemma 11: Define $\mathbf{Q} = \mathbf{Q}_a + t\cdot\mathbf{Q}_b$, where $\mathbf{Q}_a$ is positive semidefinite, $\mathbf{Q}_b$ is symmetric, and $t \in \mathbb{C}$. We proof concavity of $f(\cdot)$ in $\mathbf{Q}$ by proving that $f(\mathbf{Q}_a + t\cdot\mathbf{Q}_b)$ is convex in t for all t such that $\mathbf{Q}$ is positive semidefinite. We compute the first and second derivative of $f(\cdot)$ with respect to t as

$$\frac{\partial}{\partial t}\log\det\left(\mathbf{I}+\mathbf{A}(\mathbf{Q}_a+t\mathbf{Q}_b)(\mathbf{I}+\mathbf{B}(\mathbf{Q}_a+t\mathbf{Q}_b))^{-1}\right) \tag{322}$$

$$= \mathrm{Tr}\left[\left[(\mathbf{A}+\mathbf{B})(\mathbf{Q}_a+t\mathbf{Q}_b)+\mathbf{I}\right]^{-1}(\mathbf{A}+\mathbf{B})\mathbf{Q}_b\right] - \mathrm{Tr}\left[\left[\mathbf{B}(\mathbf{Q}_a+t\mathbf{Q}_b)+\mathbf{I}\right]^{-1}\mathbf{B}\mathbf{Q}_b\right] \tag{323}$$

and

$$\frac{\partial^2}{\partial t^2}\log\det(\mathbf{I}+\mathbf{A}(\mathbf{Q}_a+t\mathbf{Q}_b)(\mathbf{I}+\mathbf{B}(\mathbf{Q}_a+t\mathbf{Q}_b))^{-1}) \tag{324}$$

$$= -\mathrm{Tr}\left[\left(\left[(\mathbf{A}+\mathbf{B})(\mathbf{Q}_a+t\mathbf{Q}_b)+\mathbf{I}\right]^{-1}(\mathbf{A}+\mathbf{B})\mathbf{Q}_b\right)^2\right] + \mathrm{Tr}\left[\left(\left[\mathbf{B}(\mathbf{Q}_a+t\mathbf{Q}_b)+\mathbf{I}\right]^{-1}\mathbf{B}\mathbf{Q}_b\right)^2\right] \tag{325}$$

$$= -\mathrm{Tr}\left[\left(\left[\mathbf{Q}_a+t\mathbf{Q}_b+(\mathbf{A}+\mathbf{B})^{-1}\right]^{-1}\mathbf{Q}_b\right)^2\right] + \mathrm{Tr}\left[\left(\left[\mathbf{Q}_a+t\mathbf{Q}_b+\mathbf{B}^{-1}\right]^{-1}\mathbf{Q}_b\right)^2\right] \leq 0. \tag{326}$$

This expression is nonpositive due to [27, Lemma 2.3], which states that for a positive definite matrix $\mathbf{Z}$, a positive semidefinite matrix $\mathbf{W}$, and a Hermitian matrix $\mathbf{X}$

$$\mathrm{Tr}[(\mathbf{Z}^{-1}\mathbf{X})^2] \geq \mathrm{Tr}[\left((\mathbf{Z}+\mathbf{W})^{-1}\mathbf{X}\right)^2]. \quad (327)$$

In (326), we identify $\mathbf{X}$ with $\mathbf{Q}_b$, $\mathbf{Z}$ with $\mathbf{Q}_a + t\mathbf{Q}_b + (\mathbf{A}+\mathbf{B})^{-1}$ and $\mathbf{W}$ with $\mathbf{B}^{-1} - (\mathbf{A}+\mathbf{B})^{-1}$. □

Appendix C
Proof of Proposition 1

The sum-transmit power of stage $\mathcal{R}_l$ is given by

$$\begin{aligned}\sum_{k=1}^{n_\mathcal{R}} P_{\mathrm{R}_k^{(l)}} &= \frac{P_L}{n_\mathcal{S}} \frac{\alpha^{L-l+1}}{n_\mathcal{R}^{L-l+1}} \mathrm{Tr}\left[\mathbf{H}_{l+1}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{l+1}^H\right] \\ &+ \sigma_\mathrm{w}^2 \frac{\alpha}{n_\mathcal{R}}\left(1 + \sum_{l'=l}^{L-1}\frac{\alpha^{l'-l}}{n_\mathcal{R}^{l'-l}} \times \mathrm{Tr}\left[\mathbf{H}_{l+1}\cdots\mathbf{H}_{l'+1}\mathbf{H}_{l'+1}^H\cdots\mathbf{H}_{l+1}^H\right]\right). \quad (328)\end{aligned}$$

Each of the traces converges to one almost surely. This follows by repeated application of the following result from [25], [26]: Let $\mathbf{Y} = \frac{1}{m}\mathbf{X}\mathbf{R}\mathbf{X}^H$ be a random matrix, where:

- $\mathbf{R}$ is an arbitrary nonnegative definite $m \times m$ random matrix whose EED converges uniformly to a nonrandom distribution almost surely as $m \to \infty$ and satisfies $\lim_{m\to\infty} m^{-1}\mathrm{Tr}[\mathbf{R}] = 1$ almost surely;
- $\mathbf{X}$ is an $p \times m$ random matrix, independent of $\mathbf{R}$, with i.i.d. elements of zero-mean and unit variance.

Then, for every fixed ratio m/p, $\lim_{p\to\infty}\frac{1}{p}\mathrm{Tr}[\mathbf{Y}] = 1$ almost surely.

Due to the convergence of the traces, the following holds:

$$\begin{aligned}\lim_{n_\mathcal{R}\to\infty}\sum_{k=1}^{n_\mathcal{R}} P_{\mathrm{R}_k^{(l)}} &= P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2 \cdot \sum_{l'=l}^{L-1}\alpha^{L-l} \quad (329)\\ &= P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2\cdot\alpha\frac{1-\alpha^{L-l+1}}{1-\alpha} \\ &\quad \text{almost surely} \quad (330)\end{aligned}$$

where the right-hand side fulfills for $\alpha = P_L/(P_L+\sigma_\mathrm{w}^2)$

$$P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2\cdot\alpha\frac{1-\alpha^{L-l+1}}{1-\alpha} = P_L \text{ for all } l \in \{1,\ldots,L\}. \quad (331)$$

For the second part, the transmit power of relay $\mathrm{R}_k^{(l)}$ is written as

$$P_{\mathrm{R}_k^{(l)}} = \frac{\alpha}{n_\mathcal{R}}\mathbf{h}_{l+1}^{(k)\,H}\mathbf{T}_{l+1}\mathbf{h}_{l+1}^{(k)} \quad (332)$$

where $\mathbf{h}_k^{(l)}$ denotes the kth column of $\mathbf{H}_l$ and

$$\mathbf{T}_{l+1} = \begin{cases}\frac{P_L}{n_\mathcal{S}}\mathbf{I}_{n_\mathcal{R}}, \text{ if } l = L,\\ \frac{P_L}{n_\mathcal{S}}\frac{\alpha^{L-l}}{n_\mathcal{R}^{L-l}}\mathbf{H}_{l+2}\cdots\mathbf{H}_{L+1}\mathbf{H}_{L+1}^H\cdots\mathbf{H}_{l+2}^H \\ +\sigma_\mathrm{w}^2\cdot\Big(\mathbf{I}_{n_\mathcal{R}} + \sum_{l'=l}^{L-2}\frac{\alpha^{l'-l+1}}{n_\mathcal{R}^{l'-l+1}}\times \\ \quad \mathbf{H}_{l+2}\cdots\mathbf{H}_{l'+1}\mathbf{H}_{l'+1}^H\cdots\mathbf{H}_{l+2}^H\Big),\\ \quad \text{if } l \in \{1,\ldots,L-1\}.\end{cases}$$

First, a result from [28, Theorem 1] yields, that for every $k \in \{1,\ldots,n_\mathcal{R}\}$

$$\begin{aligned}\lim_{n_\mathcal{R}\to\infty} P_{\mathrm{R}_k^{(l)}} &= \lim_{n_\mathcal{R}\to\infty}\frac{\alpha}{n_\mathcal{R}}\mathsf{E}\left[\mathbf{h}_{l+1}^{(k)\,H}\mathbf{T}_{l+1}\mathbf{h}_{l+1}^{(k)}\right] \quad (333)\\ &= P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2\cdot\sum_{l'=l}^{L-1}\alpha^{L-l'+1} \quad (334)\\ &= P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2\cdot\alpha\frac{1-\alpha^{L-l+1}}{1-\alpha} \\ &\quad \text{almost surely} \quad (335)\end{aligned}$$

where the right-hand side fulfills

$$P_L\alpha^{L-l+1} + \sigma_\mathrm{w}^2\cdot\alpha\frac{1-\alpha^{L-l+1}}{1-\alpha} = P_L \text{ for all } l \in \{1,\ldots,L\} \quad (336)$$

for $\alpha = P_L/(P_L+\sigma_\mathrm{w}^2)$. [28, Theorem 1] requires the $\mathbf{T}_{l+1}$ to be positive semi-definite (obvious) and to fulfill almost surely $\lim_{n_\mathcal{R}\to\infty}\frac{1}{n_\mathcal{R}}\mathrm{Tr}[\mathbf{T}_{l+1}] < \infty$ (follows by the almost sure convergence of the traces in $\mathbf{T}_{l+1}$).

Thus, there exists for every relay $\mathrm{R}_k^{(l)}$ an $n_0^{(k)}$, such that for all $n_\mathcal{R} \geq n_0^{(k)}$

$$\left|n_\mathcal{R}P_{\mathrm{R}_k^{(l)}} - P_L\right| < \varepsilon. \quad (337)$$

Next, consider for each relay $\mathrm{R}_k^{(l)}$ the smallest such $n_0^{(k)}$ as a random variable. The distribution of these i.i.d. random variables fulfills for all $k \in \{1,\ldots,n_\mathcal{R}\}$

$$\lim_{n_\mathcal{R}\to\infty}\mathsf{Pr}\left[n_0^{(k)} < n_\mathcal{R}\right] = 1. \quad (338)$$

Now, fix for γ arbitrarily close to one, $n_\mathcal{R}^{(0)}$ sufficiently large, such that $\mathsf{Pr}\left[n_0^{(k)} < n_\mathcal{R}^{(0)}\right] > \gamma$. Then, for all $n_\mathcal{R} > n_\mathcal{R}^{(0)}$

$$\begin{aligned}&\mathsf{Pr}\left[\frac{1}{n_\mathcal{R}}\sum_{k=1}^{n_\mathcal{R}} 1\left\{n_0^{(k)} < n_\mathcal{R}\right\} > \gamma\right]\\ &> \mathsf{Pr}\left[\frac{1}{n_\mathcal{R}}\sum_{k=1}^{n_\mathcal{R}} 1\left\{n_0^{(k)} < n_\mathcal{R}^{(0)}\right\} > \gamma\right]. \quad (339)\end{aligned}$$

Consider now the inequality for all $n_\mathcal{R} > n_\mathcal{R}^{(0)}$

$$\frac{1}{n_\mathcal{R}}\sum_{k=1}^{n_\mathcal{R}} 1\left\{n_0^{(k)} < n_\mathcal{R}\right\} > \frac{1}{n_\mathcal{R}}\sum_{k=1}^{n_\mathcal{R}} 1\left\{n_0^{(k)} < n_\mathcal{R}^{(0)}\right\} \quad (340)$$

where the lower-bound fulfills due to (339) and the strong law of large numbers

$$\lim_{n_{\mathcal{R}} \rightarrow \infty} \frac{1}{n_{\mathcal{R}}} \sum_{k=1}^{n_{\mathcal{R}}} 1 \left\{ n_0^{(k)} < n_{\mathcal{R}}^{(0)} \right\} > \gamma \text{ almost surely.} \quad (341)$$

This establishes the proposition. □

REFERENCES

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[2] M. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel," in *Proc. 31st Asilomar Conf. Signals, Systems and Computers*, Nov. 1997, vol. 2, pp. 1405–1409.

[3] S. Yeh and O. Lévêque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007.

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

[5] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471–480, 1973.

[6] V. I. Morgenshtern and H. Bölcskei, "Crystallization in large wireless networks," *IEEE Trans. Inform. Theory (Special Issue on Relaying and Cooperation in Communication Networks)*, vol. 53, no. 10, pp. 3319–3349, Oct. 2007.

[7] R. R. Müller, "On the asymptotic eigenvalue distribution of concatenated vector-valued fading channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 2086–2091, Jul. 2002.

[8] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, "Asymptotic capacity and optimal precoding in MIMO multi-hop relay networks," *CoRR*, vol. Abs/0906.0037, 2009.

[9] S. Borade, L. Zheng, and R. Gallager, "Amplify and forward in wireless relay networks: Rate, diversity and network size," *IEEE Trans. Inform. Theory (Special Issue on Relaying and Cooperation in Communication Networks)*, Oct. 2007.

[10] D. Gündüz, A. Khojastepour, A. Goldsmith, and V. Poor, "Multi-hop MIMO relay networks: Diversity-multiplexing trade-off analysis," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 1738–1747, May 2010.

[11] S. Yang and J.-C. Belfiore, "Diversity of MIMO multihop relay channels," *IEEE Trans. Inform. Theory.*, unpublished.

[12] S.-W. Jeon, S.-Y. Chung, and S. Jafar, "Degrees of freedom of multi-source relay networks," in *Proc. 47th Annu. Allerton Conf.*, Sep. 2009.

[13] J. Wagner, M. Kuhn, and A. Wittneben, "Multihop-enabled orthogonalization in distributed MIMO networks," in *Proc. IEEE Personal, Indoor and Mobile Radio Conf.*, Cannes, France, Sep. 2008.

[14] A. Sanderovich, S. Shamai (Shitz), Y. Steinberg, and M. Peleg, "Decentralized receiver in a MIMO system," in *Proc. IEEE Int. Symposium on Inf. Theory*, Seattle, WA, Jul. 2006, pp. 6–10.

[15] V. A. Marcenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math. USSR-Sbornik*, vol. 1, pp. 457–483, 1967.

[16] S. Verdu and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.

[17] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. New York: Now Publishers, 2004.

[18] J. Geronimo and T. Hill, "Necessary and sufficient condition that the limit of stieltjes transforms is a stieltjes transform," *J. Approx. Theory*, vol. 121, no. 1, Mar. 2003.

[19] Z. D. Bai, "Methodologies in spectral analysis of large dimensional random matrices," *Statistica Sinica*, vol. 9, pp. 611–677, 1999.

[20] Z. Bai and J. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *IEEE Trans. Inform. Theory*, vol. 26, pp. 316–345, 1998.

[21] Y.-C. Liang, G. Pan, and Z. Bai, "Asymptotic performance of MMSE receivers for large systems using random matrix theory," *IEEE Trans. Inform. Theory*, vol. 53, no. 11, pp. 4173–4190, Nov. 2007.

[22] D. Voiculescu, "Multiplication of certain non-commuting random variables," *J. Oper. Theory*, vol. 18, pp. 223–235, 1987.

[23] R. R. Müller, "Random matrices, free probability and the replica method," in *Proc. European Signal Processing Conference (Eusipco)*, Vienna, Austria, Sep. 2004.

[24] R. Gordon, "A convergence theorem for the Riemann integral," *Mathemat. Mag.*, vol. 73, no. 2, Apr. 2000.

[25] Q. Y. Yin, "Limiting spectral distribution for a class of random matrices," *J. Multivariate Anal.*, vol. 29, pp. 50–68, 1986.

[26] J. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *J. Multivariate Anal.*, vol. 55, pp. 331–339, Nov. 1995.

[27] E.-V. Belmega, L. Samson, and M. Debbah, "A trace inequality for positive definite matrices," *J. Inequal. Pure and Applied Math.*, vol. 10, pp. 4–8, 2009.

[28] D. E. Varberg, "Almost sure convergence of quadratic forms in independent random variable," *Annal. Mathema. Statist,*, vol. 39, no. 5, pp. 1502–1506, 1968.

Jörg Wagner received both his M.Sc. degree in Information Technology and Electrical Engineering and his Ph.D. degree (Dr. Scientiae) from ETH Zurich, Switzerland, in 2006 and 2011, respectively.

From 2006 to 2011, he was a research and teaching assistant with the Communication Technology Laboratory of ETH Zurich, Switzerland. Since 2011 he has been with Swisscom (Schweiz) AG.

Armin Wittneben received his M.Sc. and Ph.D. degrees in Electrical Engineering from the Technical University of Darmstadt, Germany, in 1983 and 1989 respectively. In 1989 he joined the central research unit ASCOM Tech of ASCOM Ltd., Switzerland, as postdoctoral researcher. In 1992 he became head of the Wireless Technology and Security Department. In 1996 he also served as co-director of ASCOM Tech. In 1997 he received the Venia Legendi (Habilitation) in Communication Technology from the TU-Darmstadt. In 1998 he accepted a position as full professor of Telecommunication at the Saarland University, Germany. Since 2002 he has been full professor of Wireless Technology at ETH Zurich, Switzerland and is serving as head of the Communication Technology Laboratory. His research interests include all aspects of wireless communication with special emphasis on communication theory, signal processing, cooperative wireless signaling and protocols, distributed MIMO systems and Ultra-Wideband/Ultra-Low Power communication, imaging and position location.