ISIT'2013 1569716339

Large-System Analysis of the K-Hop AF MIMO Relay Channel with Arbitrary Inputs

Maksym A. Girnyk, Mikko Vehkaperä, Lars K. Rasmussen
School of Electrical Engineering and the ACCESS Linnaeus Center,
KTH Royal Institute of Technology, Stockholm, Sweden
email: {mgyr, mikkov, lkra}@kth.se

***Abstract*—The present paper investigates the achievable data rates of multi-hop amplify-and-forward multi-antenna relay channels with arbitrary number of hops K. Each multi-antenna terminal in the system amplifies the received signal and retransmits it upstream. To analyze the ergodic end-to-end mutual information of the system, one has to perform averaging over the fading coefficients. To overcome this difficulty we apply large-system analysis, based on the assumption that the number of antennas grows without bound at every terminal. Using the replica method, we derive an explicit asymptotic expression for the ergodic mutual information between the input and output of the K-hop channel with no restrictions on the channel inputs. Numerical results support the validity of the replica analysis and show that the result is tight even for small antenna arrays.**

I. Introduction

Multiple-input multiple-output (MIMO) relaying has been recognized in recent years as an attractive solution for improvement of the quality-of-service of communication systems. In spite of the considerable efforts, however, the capacity of MIMO relay systems remains an open problem, except for some special cases [1]. Various relaying strategies have been developed in the literature. Among those, *amplify-and-forward* (AF) [2], being simple in implementation and having low computational complexity, is of particular interest.

The present paper studies a general K-hop AF MIMO relay channel, depicted in Fig. 1, in terms of the ergodic mutual information. The two major obstacles in evaluating the average performance of such channels are

- computation of the averages of the mutual information over the channel realizations,
- computation of the averages over the input signals when non-Gaussian constellations are used.

To overcome these difficulties, several asymptotic approaches, based on the large-system assumption, have been recently proposed. For instance, in [3], various asymptotic limits are considered; namely, when numbers of antennas at the one side of the channel (source, destination or relay terminals) grows large, while at the other side the number of antennas stays fixed. Other authors apply techniques from large *random matrix theory* (RMT), assuming that the number of antennas at each terminal within the relay channel grow very large. In [4], an explicit expression for the achievable rate of a two-hop

The present research was supported by VR Grant 621-2011-1024.

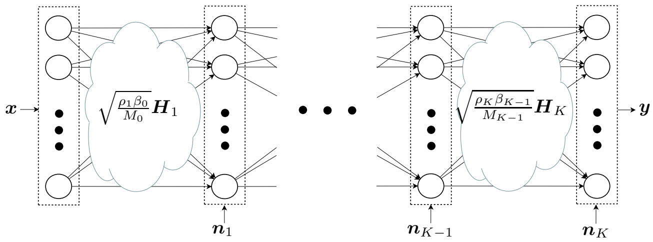

Fig. 1. K-hop AF MIMO relay channel.

AF MIMO channel with Gaussian inputs is obtained. Further, a similar approach is used in [5] to extend the analysis of multi-hop channels with noise present only at the destination terminal. Meanwhile, in [6] the capacity of a general K-hop AF MIMO channel with Gaussian inputs and noise at every hop is expressed in terms of the limiting eigenvalue distribution of a product of MIMO channel matrices, obtained via a set of recursive equations. In [7], by means of deterministic equivalents the authors derive a recursive expression for the ergodic mutual information between the input and output of a K-hop AF MIMO channel. Another recent paper, [8], considers a similar set-up while evaluating the average throughput under the SVD-based linear precoding.

All the prior results above are based on the assumption of Gaussian channel inputs. While being a helpful abstraction due to its mathematical convenience, Gaussian signaling is, however, impractical for real-world systems. Moreover, RMT does not provide sufficient tools to deal with practical signal constellations, e.g., QPSK and QAM. Therefore, there is a need for a method capable of analyzing the performance of the system under arbitrary channel inputs. To address this, we employ the *replica method* [9] from the field of statistical physics. The method allows to characterize the performance of large CDMA [10], [11], as well as MIMO [12] communication systems, including AF MIMO relay channels [13], under arbitrary signaling.

In the present paper, we follow the philosophy of [7] and derive an explicit expression for the ergodic mutual information of a K-hop AF MIMO relay channel by iteratively removing the randomness of the channel and noise at each hop. However, in contrast to [7], our analysis is not restricted to Gaussian channel inputs, which allows us, for instance, to evaluate the performance of the system under QPSK signaling.

II. System Model

Consider a K-hop channel, consisting of multi-antenna relay terminals that help a multi-antenna source to communicate with a multi-antenna destination. The corresponding setup is depicted in Fig. 1. There is no direct link between the source and the destination and the terminals operate under the time-division multiple-access (TDMA) protocol, so that a single node transmits at a given time. Furthermore, we assume that each relay receives only the signals from the preceding hop. Namely, a symbol sent by the source has to traverse K hops before it reaches the destination, where K is a fixed finite number. The source, the destination and the kth relay terminal are equipped with M_0, M_K and M_k antennas, respectively. AF relaying is employed at each relay, so that the relay amplifies and retransmits the received signal upstream without decoding.

A channel at hop k is described by the following input-output relation

$$\boldsymbol{y}_k = \sqrt{\frac{\rho_k \beta_{k-1}}{M_k}} \boldsymbol{H}_k \boldsymbol{y}_{k-1} + \boldsymbol{n}_k, \tag{1}$$

where $\boldsymbol{y}_{k-1}$ and $\boldsymbol{y}_k$ are the input and output of the kth channel, respectively. Moreover, ρ_k is the signal-to-noise ratio (SNR) at hop k and β_{k-1} is the normalization constant chosen so that the long-term transmit power constraint at terminal $k-1$ is satisfied, that is,

$$\beta_{k-1} \mathsf{E}\left\{ \mathrm{tr}\{\boldsymbol{y}_{k-1}\boldsymbol{y}_{k-1}^{\mathsf{H}}\} \right\} \leq M_{k-1}. \tag{2}$$

Finally, $\boldsymbol{H}_k$ is the channel matrix between terminals $k-1$ and k, assumed to be a standard complex Gaussian random matrix and $\boldsymbol{n}_k \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{M_k})$.

The end-to-end input-output relation of the whole channel can be written as follows:

$$\boldsymbol{y} = \boldsymbol{G}_0^{K-1}\boldsymbol{x} + \sum_{k=1}^{K-1} \boldsymbol{G}_k^{K-1}\boldsymbol{n}_k + \boldsymbol{n}_K, \tag{3}$$

where $\boldsymbol{x} \triangleq \boldsymbol{y}_0$ is the input to the relay channel, assumed for the sake of simplicity to have i.i.d. zero-mean unit-variance components distributed according to some probability density (mass) function $p(x)$, $\boldsymbol{y} \triangleq \boldsymbol{y}_K$ and the corresponding matrices are defined as[1]

$$\boldsymbol{G}_i^j \triangleq \prod_{k=i}^{j} \sqrt{\frac{\rho_{k+1}\beta_k}{M_k}} \boldsymbol{H}_{k+1}, \quad j \geq i. \tag{4}$$

III. Asymptotic Ergodic Mutual Information

In this section we study the achievable rates for the multi-hop AF MIMO relay channel under the assumption that the destination has full channel state information. In this case, the long-term maximum achievable rate is given by the average mutual information between the input and output of the channel. Define the set of channel matrices $\mathcal{H} \triangleq \{\boldsymbol{H}_1, \ldots, \boldsymbol{H}_K\}$ and the set of noise realizations $\mathcal{N} \triangleq \{\boldsymbol{n}_1, \ldots, \boldsymbol{n}_{K-1}\}$. The mutual information between $\boldsymbol{y}$ and $\boldsymbol{x}$ is then written as

$$I(\boldsymbol{y};\boldsymbol{x}) = \frac{1}{K}\left[h(\boldsymbol{y}|\mathcal{H}) - h(\boldsymbol{y}|\boldsymbol{x},\mathcal{H})\right], \tag{5}$$

where the differential entropy terms are given by

$$h(\boldsymbol{y}|\mathcal{H}) = -\,\mathsf{E}_{\boldsymbol{y},\mathcal{H}} \ln \mathsf{E}_{\boldsymbol{x},\mathcal{N}}\left\{p(\boldsymbol{y}|\boldsymbol{x},\mathcal{H},\mathcal{N})\right\}, \tag{6a}$$

$$h(\boldsymbol{y}|\boldsymbol{x},\mathcal{H}) = -\,\mathsf{E}_{\boldsymbol{y},\boldsymbol{x},\mathcal{H}} \ln \mathsf{E}_{\mathcal{N}}\left\{p(\boldsymbol{y}|\boldsymbol{x},\mathcal{H},\mathcal{N})\right\}, \tag{6b}$$

with the conditional distribution of the channel being

$$p(\boldsymbol{y}|\boldsymbol{x},\mathcal{H},\mathcal{N}) = \frac{1}{\pi^{M_K}} e^{-\|\boldsymbol{y} - \boldsymbol{G}_0^{K-1}\boldsymbol{x} - \sum_{k=1}^{K-1}\boldsymbol{G}_k^{K-1}\boldsymbol{n}_k\|^2} \tag{7}$$

Unfortunately, direct computation of the mutual information in (5) is prohibitive due to the expectation of the logarithmic term over the channel distribution. Hence, some simplifying assumptions have to be invoked in order for the problem to be solvable. In the present paper, we investigate the performance of the system in the *large-system limit* (LSL), meaning that for each intermediate hop k the number of transmit and receive antennas tend to infinity at some constant ratio, viz., $M_{k-1} = \alpha_k M_k \to \infty$ as $\alpha_k = \text{const}$. This allows us to obtain explicit expressions for differential entropy terms (6a) and (6b).

Before presenting the main results, let us consider the following fixed MIMO channel

$$\boldsymbol{z} = \sqrt{\boldsymbol{A}}\boldsymbol{x} + \boldsymbol{w}, \tag{8}$$

where $\sqrt{\boldsymbol{A}}$ is a fixed channel matrix and the noise vector $\boldsymbol{w} \sim \mathcal{CN}(\boldsymbol{0}_{M_0}, \boldsymbol{I}_{M_0})$, so that

$$p\left(\boldsymbol{z}\middle|\boldsymbol{x},\sqrt{\boldsymbol{A}}\right) = \frac{1}{\pi^{M_0}} e^{-\|\boldsymbol{z}-\sqrt{\boldsymbol{A}}\boldsymbol{x}\|^2}. \tag{9}$$

The minimum mean-squared error (MMSE) estimate of $\boldsymbol{x}$ is denoted by

$$\langle \boldsymbol{x} \rangle \triangleq \mathsf{E}\left\{\boldsymbol{x}\middle|\boldsymbol{z},\sqrt{\boldsymbol{A}}\right\}, \tag{10}$$

where the expectation is taken over $p(\boldsymbol{x}|\boldsymbol{z},\sqrt{\boldsymbol{A}})$, which is obtained from the prior $p(\boldsymbol{x})$ and $p(\boldsymbol{z}|\boldsymbol{x},\sqrt{\boldsymbol{A}})$ through Bayes' formula. With the aforementioned definitions, we summarize our main finding in the following claim,[2] whose derivation is postponed to the Appendix.

Claim 1. *In the LSL, the differential entropy* $h(\boldsymbol{y}|\mathcal{H})$ *given in* (6a) *reads*

$$h(\boldsymbol{y}|\mathcal{H}) = I\left(\boldsymbol{z};\boldsymbol{x}|\sqrt{\boldsymbol{A}}\right) + M_K \ln(1+\varepsilon_K) - \sum_{k=1}^{K} M_{k-1}\xi_k\varepsilon_k + \sum_{k=1}^{K-1} M_k \ln\left[1 + \rho_{k+1}\beta_k\xi_{k+1}(\varepsilon_k+1)\right] + M_K(1+\ln\pi). \tag{11}$$

where $I(\boldsymbol{z};\boldsymbol{x}|\sqrt{\boldsymbol{A}})$ *is the mutual information between the input* $\boldsymbol{x}$ *and the output* $\boldsymbol{z}$ *of the fixed MIMO channel given*

[1]The matrix product notation is defined as $\prod_{k=i}^{j} \boldsymbol{H}_k \triangleq \boldsymbol{H}_j\boldsymbol{H}_{j-1}\cdots\boldsymbol{H}_i$.

[2]The result relies on the widely accepted assumptions that the *replica trick* is valid and the *replica symmetry* holds. For further discussion, vide [10], [14] and the Appendix.

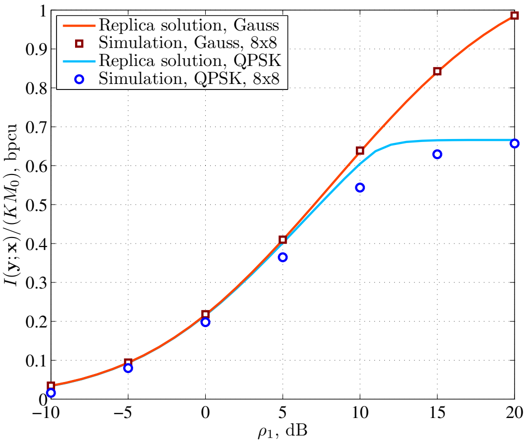

Fig. 2. Average mutual information per dimension versus the SNR at the first hop ρ_1. Solid curves denote our analytic results, markers denote the results of Monte-Carlo simulations.

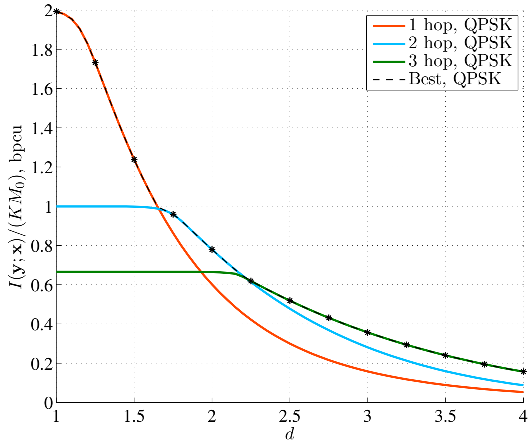

Fig. 3. Average mutual information per dimension versus the distance between the source and destination terminals for $K = 1, 2, 3$-hop networks. Solid curves denote analytic results, dashed curve denotes the best option.

by (8), *with* $\boldsymbol{A} = \rho_1\beta_0\xi_1\boldsymbol{I}_{M_0}$. *The parameters* ξ_k, ε_k, $\forall k \in \{1, \ldots, K\}$ *satisfy the following set of fixed-point equations*[3]

$$\xi_K = \frac{M_K}{M_{K-1}(1+\varepsilon_K)}, \tag{12a}$$

$$\xi_k = \frac{M_k\rho_{k+1}\beta_k\xi_{k+1}}{M_{k-1}\left[1+\rho_{k+1}\beta_k\xi_{k+1}(1+\varepsilon_k)\right]}, \tag{12b}$$

$$\varepsilon_k = \frac{\rho_k\beta_{k-1}(\varepsilon_{k-1}+1)}{1+\rho_k\beta_{k-1}\xi_k(\varepsilon_{k-1}+1)}, \tag{12c}$$

$$\varepsilon_1 = \frac{\rho_1\beta_0}{M_0}\mathsf{E}_{\boldsymbol{z},\boldsymbol{x}}\{\|\boldsymbol{x} - \langle\boldsymbol{x}\rangle\|^2\}. \tag{12d}$$

At the same time, the conditional differential entropy $h(\boldsymbol{y}|\boldsymbol{x}, \mathcal{H})$*, given in* (6b)*, may be computed directly by* (11) *and* (12) *provided that* $\xi_1 = \varepsilon_1 = 0$ *beforehand.*

The expectation in (12d) is taken over $p(\boldsymbol{z}, \boldsymbol{x}|\sqrt{\boldsymbol{A}})$, and hence ε_1 can be seen as the MMSE of the MIMO channel (8). The entropy term $h(\boldsymbol{y}|\mathcal{H})$, given in (11), represents the amount of information contributed by the transmitted signal $\boldsymbol{x}$ and by noise components $\boldsymbol{n}_k$ added at each hop. Note that both the MMSE and mutual information terms are relatively easy to compute since $\boldsymbol{A}$ is a fixed diagonal matrix. Meanwhile, the differential entropy $h(\boldsymbol{y}|\boldsymbol{x}, \mathcal{H})$ represents the amount of information discarded at the destination terminal due to noise removal. It can also be shown that the normalization coefficients β_k, satisfying power constraint (2), are

$$\beta_k = \frac{1}{1+\rho_k}, \quad \forall k \in \{1, \ldots, K\}. \tag{13}$$

Knowing both differential entropy terms, one can directly compute the mutual information (5) as a function of SNRs ρ_k at each hop. For instance, the following examples represent two particular signal constellations.

Example 1 (*Gaussian inputs*)**.** When $p(x)$ is the standard complex Gaussian density, the MMSE term (12d) is given by $\varepsilon_1 = \frac{\rho_1\beta_0}{1+\rho_1\beta_0\xi_1}$ and the mutual information between the input and output of (8) reads

$$I\left(\boldsymbol{z}; \boldsymbol{x}|\sqrt{\boldsymbol{A}}\right) = M_0 \ln\left(1 + \rho_1\beta_0\xi_1\right). \tag{14}$$

Example 2 (*QPSK inputs*)**.** For the QPSK constellation we have $p(x) = 1/4$ for all $x = \pm\frac{1}{\sqrt{2}} \pm \frac{\mathrm{j}}{\sqrt{2}}$. The MMSE reads

$$\varepsilon_1 = \rho_1\beta_0 - \frac{\rho_1\beta_0}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{s^2}{2}} \tanh\left(\rho_1\beta_0\xi_1 + \sqrt{\rho_1\beta_0\xi_1}s\right)\,\mathrm{d}s, \tag{15}$$

and by the I-MMSE relation [15], the mutual information between the output and the input of (8) is evaluated as

$$I\left(\boldsymbol{z}; \boldsymbol{x}|\sqrt{\boldsymbol{A}}\right) = 2M_0\rho_1\beta_0 - \frac{2M_0}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{s^2}{2}} \ln\cosh(\rho_1\beta_0\xi_1 + \sqrt{\rho_1\beta_0\xi_1}s)\,\mathrm{d}s. \tag{16}$$

IV. Simulation Results

To support the main result, we simulate a network with $K = 3$ hops. We set $M_0 = \ldots = M_3 = 8$, so that each terminal has eight antennas, and fix SNRs $\rho_2 = \rho_3 = 20$ dB while varying the SNR of the first hop ρ_1. In Fig. 2, we plot the normalized ergodic mutual information for two types of channel inputs, Gaussian and QPSK signals. The rate loss of $1/K$ here is due to the TDMA protocol. Solid lines reflect the asymptotic results, while markers represent the numerical averaging via Monte-Carlo simulations. We note that for the case of Gaussian inputs the approximation matches the simulations perfectly. Moreover, it reproduces exactly the same result as in [7]. Meanwhile, for the case of QPSK inputs there is a slight gap in the middle- and high-SNR region, which decreases with increasing the actual number of antennas at each terminal.

For the second setup we incorporate pathloss $\gamma_k = d_k^{-\alpha}$ into SNRs $\rho_k = 10$ dB, $\forall k$, where d_k is the distance between terminals $k-1$ and k, and $\alpha = 4$ is the pathloss exponent. In Fig. 3, we plot the normalized ergodic mutual information as a function of distance d between the source and destination

[3] The fixed-point equations are solved iteratively. In case of non-Gaussian inputs, the system of equations may have several solutions among which the one minimizing the differential entropy of interest should be picked.

terminals. The three curves correspond to $K = 1, 2, 3$. The relaying terminals are added in such way that all the terminals of the network are equidistant. The dashed line represents the maximum value of the mutual information at each point. Notably, for different values of d, different number of hops provides higher data rate. Thus, one could use this information to optimize the number of relays or their positions. This is, however, outside the scope of the present paper.

V. Conclusions

In this paper, we have analyzed the asymptotic performance of the multi-hop AF MIMO relay channel in the large system limit. With help of the replica method, we have derived an explicit deterministic approximation for the ergodic mutual information for arbitrary number of hops. The expression is presented as a function of the SNRs at each hop, as well as a set of parameters obtained by iterative solution of a set of fixed-point equations. Advantageous to the random matrix-based approaches, our analysis is not restricted to Gaussian distributed channel inputs. For instance, we have effectively computed the asymptotic ergodic mutual information with practical QPSK inputs. The numerical results corroborate the validity of the derived asymptotic approximation. Moreover, simulations suggest that the approximation is very accurate even for small numbers of antennas at the terminals. The presented results are extendable to more sophisticated channel models of interest (e.g., Kronecker model or Rician fading) allowing for further performance optimization.

Appendix
Derivation of Claim 1

Let us define the following statistical mechanical quantity

$$Z(\boldsymbol{y}, \mathcal{H}) \triangleq \mathsf{E}_{\boldsymbol{x},\mathcal{N}} \left\{ \frac{1}{\pi^{M_K}} e^{-\|\boldsymbol{y}-\boldsymbol{G}_0^{K-1}\boldsymbol{x}-\sum_{k=1}^{K-1}\boldsymbol{G}_k^{K-1}\boldsymbol{n}_k\|^2} \right\}, \tag{17}$$

which designates the *partition function* of the system. The differential entropy of interest (6a) can thus be rewritten as

$$h(\boldsymbol{y}|\mathcal{H}) = -\,\mathsf{E}_{\boldsymbol{y},\mathcal{H}} \ln Z(\boldsymbol{y}, \mathcal{H}) \tag{18a}$$

$$= -\lim_{u\to 0} \frac{\partial}{\partial u} \ln \mathsf{E}_{\boldsymbol{y},\mathcal{H}} \left\{ Z^u(\boldsymbol{y}, \mathcal{H}) \right\}. \tag{18b}$$

Henceforth, instead of averaging the logarithm we have to compute the uth moment of the partition function. However, for real-valued u computing $\mathsf{E}_{\boldsymbol{y},\mathcal{H}}\{Z^u(\boldsymbol{y}, \mathcal{H})\}$ is very difficult. Thus, we make a non-rigorous assumption that the latter can be evaluated for integer u as follows

$$\mathsf{E}_{\boldsymbol{y},\mathcal{H}} \left\{ Z^u(\boldsymbol{y}, \mathcal{H}) \right\} = \mathsf{E} \left\{ \int \frac{1}{\pi^{M_K}} \prod_{a=0}^{u} e^{-\|\boldsymbol{y}-\boldsymbol{G}_0^{K-1}\boldsymbol{x}^{(a)}-\sum\limits_{k=1}^{K-1}\boldsymbol{G}_k^{K-1}\boldsymbol{n}_k^{(a)}\|^2} \mathrm{d}\boldsymbol{y} \right\}, \tag{19}$$

and then generalize to real values outside the logarithm in (18b). In the literature, this step is referred to as the *replica trick* and its rigorousness remains an open problem (cf. [16]).

In (19), $\boldsymbol{x}^{(a)}$ and $\boldsymbol{n}_k^{(a)}$ denote the ath replica vectors that are assumed to be i.i.d., having the same distributions as their original counterparts $\boldsymbol{x}^{(0)}$ and $\boldsymbol{n}_k^{(0)}$, representing the signal vector $\boldsymbol{x}$ and noise vector $\boldsymbol{n}_k$ at the kth hop. For ease of exposition, we group the corresponding vectors into $\boldsymbol{X} \triangleq [\boldsymbol{x}^{(0)\mathsf{T}}, \ldots, \boldsymbol{x}^{(u)\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{M_0(u+1)}$ and $\boldsymbol{N}_k \triangleq [\boldsymbol{n}_k^{(0)\mathsf{T}}, \ldots, \boldsymbol{n}_k^{(u)\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{M_k(u+1)}$.

In (19), the averaging should be performed over all possible channel inputs $\boldsymbol{X}$, channel gains $\mathcal{H}$ and noise realizations $\mathcal{N}$. According to the Fubini theorem [17], provided that the expectation in (19) exists, the multiple integral can be computed via repeated integrals. In other words, averaging over the channel matrices and noise vectors of all hops can be done iteratively hop-by-hop. That is, at each hop, we average out the randomness of the channel matrix and the noise vector, while keeping the variables related to other hops fixed. To do this, define the following set of vectors

$$\boldsymbol{v}_k^{(a)} \triangleq \boldsymbol{G}_0^{k-1}\boldsymbol{x}^{(a)} + \sum_{i=1}^{k-1} \boldsymbol{G}_i^{k-1}\boldsymbol{n}_i^{(a)} \in \mathbb{C}^{M_k}, \tag{20}$$

containing the randomness of the first k hops in the network. For each k, stack these vectors as $\boldsymbol{V}_k \triangleq [\boldsymbol{v}_k^{(0)\mathsf{T}}, \ldots, \boldsymbol{v}_k^{(u)\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{M_k(u+1)}$. Note that each vector $\boldsymbol{v}_k^{(a)}$ can be recursively expressed in terms of vector $\boldsymbol{v}_{k-1}^{(a)}$, containing the randomness of first $k-1$ hops, as follows

$$\boldsymbol{v}_k^{(a)} = \sqrt{\frac{\rho_k \beta_{k-1}}{M_{k-1}}} \boldsymbol{H}_k \boldsymbol{v}_{k-1}^{(a)} \in \mathbb{C}^{M_{k-1}}. \tag{21}$$

It can be shown that for $k \in \{2, \ldots, K-1\}$, conditioned on $\{\boldsymbol{H}_1, \ldots, \boldsymbol{H}_{k-1}\}$ and $\{\boldsymbol{n}_1^{(a)}, \ldots, \boldsymbol{n}_{k-1}^{(a)}\}$, $a \in \{1, \ldots, u\}$, vector $\boldsymbol{V}_k$ in the LSL converges to a complex Gaussian random vector (vide [10]) with the covariance matrix given by $\boldsymbol{K}_k = \boldsymbol{Q}_k \otimes \boldsymbol{I}_{M_k} + \boldsymbol{I}_{k(u+1)} \in \mathbb{C}^{M_k(u+1)\times M_k(u+1)}$, where

$$[\boldsymbol{Q}_k]_{a,b} \triangleq \frac{\rho_k \beta_{k-1}}{M_{k-1}} \boldsymbol{v}_{k-1}^{(b)\mathsf{H}} \boldsymbol{v}_{k-1}^{(a)} \in \mathbb{C}^{(u+1)\times(u+1)}. \tag{22}$$

Starting from hop K, we follow the standard approach used in [10], [11] and write (19) as

$$\mathsf{E}_{\boldsymbol{y},\mathcal{H}} \left\{ Z^u(\boldsymbol{y}, \mathcal{H}) \right\} = \int e^{G_K^{(u)}(\boldsymbol{Q}_K)} \,\mathrm{d}\mu_K^{(u)}(\boldsymbol{Q}_K), \tag{24}$$

where we have omitted all the vanishing terms and the probability measure of $\boldsymbol{Q}_K$ is given by

$$\mu_K^{(u)}(\boldsymbol{Q}_K) = \mathsf{E} \left\{ \prod_{a,b=0}^{u} \delta\left(\rho_K \beta_{K-1} \boldsymbol{v}_{K-1}^{(b)\mathsf{H}} \boldsymbol{v}_{K-1}^{(a)} - M_{K-1} Q_K^{a,b} \right) \right\}, \tag{25}$$

with $\delta(\cdot)$ being the Dirac function.

Based on the Gärtner-Ellis theorem [18], it can be further shown that in the LSL

$$\mathsf{E}_{\boldsymbol{y},\mathcal{H}} \left\{ Z^u(\boldsymbol{y}, \mathcal{H}) \right\} = \min_{\mathcal{Q}_K} \max_{\tilde{\mathcal{Q}}_K} \left\{ T_K^{(u)}(\mathcal{Q}_K^K, \tilde{\mathcal{Q}}_K^K) \right\}, \tag{26}$$

where $\mathcal{Q}_i^j \triangleq \{\boldsymbol{Q}_i, \ldots, \boldsymbol{Q}_j\}$, $\tilde{\mathcal{Q}}_i^j \triangleq \{\tilde{\boldsymbol{Q}}_K, \ldots, \tilde{\boldsymbol{Q}}_j\}$, and

$$T_K^{(u)}(\mathcal{Q}_K^K, \tilde{\mathcal{Q}}_K^K) = M_{K-1}\mathrm{tr}\{\boldsymbol{Q}_K\tilde{\boldsymbol{Q}}_K\} - \ln M_K^{(u)}(\tilde{\boldsymbol{Q}}_K) + M_K \ln \pi(u+1) + \ln\det(\boldsymbol{I}_{M_K(u+1)} + \boldsymbol{Q}_K\boldsymbol{\Sigma} \otimes \boldsymbol{I}_{M_K}), \tag{27}$$

$$
\begin{aligned}
-\ln M_k^{(u)}(\tilde{\boldsymbol{Q}}_k) = &-\ln M_{k-1}^{(u)}(\tilde{\boldsymbol{Q}}_{k-1}) + M_{k-2}\mathrm{tr}\{\boldsymbol{Q}_{k-1}\tilde{\boldsymbol{Q}}_{k-1}\} \\
&+ \ln\det\left(\boldsymbol{I}_{M_{k-1}(u+1)} + \rho_{k-1}\beta_{k-2}(\tilde{\boldsymbol{Q}}_k \otimes \boldsymbol{I}_{M_{k-1}})\left[\boldsymbol{Q}_{k-1} \otimes \boldsymbol{I}_{M_k} + \boldsymbol{I}_{M_{k-2}(u+1)}\right]\right)
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
T_1^{(u)}(\mathcal{Q}_1, \tilde{\mathcal{Q}}_1) = &M_K \ln \pi(u+1) + \ln\det\left(\boldsymbol{I}_{M_K(u+1)} + \boldsymbol{Q}_K \boldsymbol{\Sigma} \otimes \boldsymbol{I}_{M_K}\right) - \ln \mathsf{E}_{\boldsymbol{X}}\left\{\exp\left(\rho_1\beta_0 \boldsymbol{X}^{\mathsf{H}}(\tilde{\boldsymbol{Q}}_1 \otimes \boldsymbol{I}_{M_0})\boldsymbol{X}\right)\right\} \\
&+ \sum_{k=1}^{K-1} \ln\det\left(\boldsymbol{I}_{M_k(u+1)} + \rho_{k+1}\beta_k(\tilde{\boldsymbol{Q}}_{k+1} \otimes \boldsymbol{I}_{M_k})[\boldsymbol{Q}_k \otimes \boldsymbol{I}_{M_{k+1}} + \boldsymbol{I}_{M_k(u+1)}]\right) + \sum_{k=1}^{K} M_{k-1}\mathrm{tr}\{\boldsymbol{Q}_k\tilde{\boldsymbol{Q}}_k\}
\end{aligned} \tag{31}
$$

$$
-\ln \frac{1}{\pi^{M_0}} \int \mathsf{E}_{\boldsymbol{X}}\left\{\exp\left(-\|\boldsymbol{z} - \sqrt{\boldsymbol{A}}\boldsymbol{x}\|^2 + \boldsymbol{x}^{\mathsf{H}}\boldsymbol{B}\boldsymbol{x}\right)\right\}\left[\mathsf{E}_{\boldsymbol{X}}\left\{\exp\left(\boldsymbol{z}^{\mathsf{H}}\sqrt{\boldsymbol{A}}\boldsymbol{x} + \boldsymbol{x}^{\mathsf{H}}\sqrt{\boldsymbol{A}}\boldsymbol{z} - \boldsymbol{x}^{\mathsf{H}}(\boldsymbol{A} - \boldsymbol{B})\boldsymbol{x}\right)\right\}\right]^u \mathrm{d}\boldsymbol{z} \tag{32}
$$

$$
\sum_{k=1}^{K-1} uM_k \ln\left(1 - \rho_{k+1}\beta_k\omega_{k+1}[1+\varepsilon_k]\right) + \sum_{k=1}^{K-1} M_k \ln\left(1 - \rho_{k+1}\beta_k[\omega_{k+1} + (u+1)\xi_{k+1}][1 + \varepsilon_k + (u+1)\zeta_k]\right) \tag{33}
$$

where $\boldsymbol{\Sigma} \triangleq \boldsymbol{I}_{u+1} - \frac{1}{u+1}\mathbf{1}_{u+1}\mathbf{1}_{u+1}^{\mathsf{T}}$ and $M_K^{(u)}(\tilde{\boldsymbol{Q}}_K)$ is the moment-generating function for $\boldsymbol{Q}_K$, given by

$$
M_K^{(u)}(\tilde{\boldsymbol{Q}}_K) = \mathsf{E}_{\boldsymbol{V}_{K-1}}\left\{e^{\rho_K\beta_{K-1}\boldsymbol{V}_{K-1}^{\mathsf{H}}(\tilde{\boldsymbol{Q}}_K \otimes \boldsymbol{I}_{M_{K-1}})\boldsymbol{V}_{K-1}}\right\}. \tag{28}
$$

We now need to evaluate the second term in (27) following the same procedure as above. Namely, it can be shown via induction that for $k \in \{2, \ldots, K\}$ this term can be written as (29) on the top of the page. Thus, we iteratively evaluate $M_k^{(u)}(\tilde{\boldsymbol{Q}}_k)$, $\forall k$ and arrive at

$$
h(\boldsymbol{y}|\mathcal{H}) = -\lim_{u\to 0}\frac{\partial}{\partial u}\min_{\mathcal{Q}_1}\max_{\tilde{\mathcal{Q}}_1}\left\{T_1^{(u)}(\mathcal{Q}_1, \tilde{\mathcal{Q}}_1)\right\}, \tag{30}
$$

where $T_1^{(u)}(\mathcal{Q}_1, \tilde{\mathcal{Q}}_1)$ is given by (31) on the top of the page.

Finding the fixed point of (30) is a complicated task and may not be realizable directly. Hence, a simplifying *replica symmetry* (RS) assumption is made in order to proceed. Namely, $\boldsymbol{Q}_k = \zeta_k \mathbf{1}_{u+1}\mathbf{1}_{u+1}^{\mathsf{T}} + \varepsilon_k \boldsymbol{I}_{u+1}$ and $\tilde{\boldsymbol{Q}}_k = \xi_k \mathbf{1}_{u+1}\mathbf{1}_{u+1}^{\mathsf{T}} + \omega_k \boldsymbol{I}_{u+1}$ for all $k \in \{1, \ldots, K\}$. The assumption has been widely accepted in literature due to the fact that the physics of the whole system should not depend on the artificially introduced replica indices. However, there have been reported cases where *replica symmetry breaking* appears [14] and the RS-based approach fails.

With the RS assumption, the second term of (31) is simplified to $uM_K \ln(1 + \varepsilon_K)$, the third term reduces to (32) on the top of the page, where $\boldsymbol{A} \triangleq \rho_1\beta_0\xi_1\boldsymbol{I}_{M_0}$ and $\boldsymbol{B} \triangleq \rho_1\beta_0(\omega_1 + \xi_1)\boldsymbol{I}_{M_0}$. Furthermore, the fourth term of (31) is reduced to (33) and, finally, the fifth term can be rewritten as $\sum_{k=1}^{K} M_{k-1}(u+1)\left[(\varepsilon_k + \zeta_k)(\omega_k + \xi_k) + u\zeta_k\xi_k\right]$.

Now, to evaluate the differential entropy $h(\boldsymbol{y}|\mathcal{H})$ we have to find the saddle-point conditions for which parameters have to satisfy in order to solve the min-max problem in (30). This is done by setting to zero all partial derivatives of (31) w.r.t. the $4K$ parameters. Thus, we find that, at $u = 0$, $\omega_k = -\xi_k$, $\forall k \in \{1, \ldots, K\}$ and the set of saddle-point conditions is formed precisely by (12). Further, taking the derivative of (31) w.r.t. u, at $u = 0$, we obtain the differential entropy (11).

The second entropy term, (6b), is computed by following a similar procedure, thereby concluding the proof.

References

[1] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[2] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[3] K.-J. Lee, J.-S. Kim, G. Caire, and I. Lee, "Asymptotic ergodic capacity analysis for MIMO amplify-and-forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2712–2717, Sep. 2010.

[4] S. Jin, R. McKay, C. Zhong, and K.-K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2204–2224, May 2010.

[5] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, "Asymptotic capacity and optimal precoding in MIMO multi-hop relay networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2050–2069, Apr. 2011.

[6] S.-P. Yeh and O. Lévêque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*. IEEE, 2007, pp. 1436–1440.

[7] J. Hoydis, R. Couillet, and M. Debbah, "Iterative deterministic equivalents for the capacity analysis of communication systems." [Online]. Available: http://arxiv.org/abs/1112.4167

[8] C.-K. Wen, J.-C. Chen, , and P. Ting, "Efficient approach for evaluating throughput of multi-hop MIMO amplify-and-forward relays," *IEEE Wireless Commun. Lett.*, vol. 1, no. 6, pp. 601–604, Dec. 2012.

[9] S. F. Edwards and P. W. Anderson, "Theory of spin glasses," *J. Phys. F: Metal Phys.*, vol. 5, pp. 965–974, May 1975.

[10] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inf. Theory*, vol. 48, no. 11, pp. 2888–2910, Nov. 2002.

[11] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1983–2010, Jun. 2005.

[12] R. Müller, "Channel capacity and minimum probability of error in large dual antenna array systems with binary modulation," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2821–2828, Nov. 2003.

[13] C.-K. Wen, K.-K. Wong, and C. Ng, "On the asymptotic properties of amplify-and-forward MIMO relay channels," *IEEE Trans. Commun.*, vol. 59, no. 99, pp. 1–13, Feb. 2011.

[14] B. M. Zaidel, R. R. Müller, A. L. Moustakas, and R. De Miguel, "Vector precoding for Gaussian MIMO broadcast channels: Impact of replica symmetry breaking," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1413–1440, Mar. 2012.

[15] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[16] T. Tanaka, "Moment problem in replica method," *Interdisc. Inf. Scienc.*, vol. 13, no. 1, pp. 17–23, 2007.

[17] P. Billingsley, *Probability and Measure*, 3rd ed. Hoboken, NJ: John Wiley and Sons, Inc., 1995.

[18] J. A. Bucklew, *Large deviations techniques in decision, simulation, and estimation.* John Wiley and Sons, Inc., Aug. 1990.