Correspondence

Large n Analysis of Amplify-and-Forward MIMO Relay Channels With Correlated Rayleigh Fading

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

***Abstract*—In this correspondence, the cumulants of the mutual information of the flat Rayleigh fading two-hop amplify-and-forward multiple-input multiple-output (MIMO) relay channel under independent and identically distributed (i.i.d.) Gaussian input vectors are derived in the large array limit. The analysis is based on the *replica trick* and covers both spatially independent and Kronecker correlated fading. Beamforming at all terminals is restricted to weight matrices that are independent of the channel realization and constant over time. Expressions for mean and variance of the mutual information are obtained. Their parameters are determined by a nonlinear equation system. All higher cumulants are shown to vanish as the number of antennas per terminal, n, grows to infinity. In conclusion, the distribution of the mutual information I becomes Gaussian in the large n limit. In this asymptotic regime, it is completely characterized by the expressions obtained for mean and variance of I, which are in $\Theta(n)$ and $\mathcal{O}(1)$, respectively. Comparisons with simulation results show that the asymptotic results serve as excellent approximations for systems with only few antennas at each terminal.**

***Index Terms*—Correlated channels, cumulants of mutual information, amplify-and-forward, multiple-input multiple-output (MIMO) relay channel, replica analysis.**

I. Introduction

Cooperative relaying has obtained major attention in the wireless communications community in recent years due to its various potentials regarding the enhancement of diversity, achievable rates and range. An important building block in this field is the multiple-input multiple-output *(MIMO) relay channel*. Such a channel consists of a source, a relay and a destination terminal, each equipped with multiple antennas.

Generally, there are different ways of including relays in the transmission between a source and a destination terminal. Most commonly, relays are introduced to either decode the noisy signal transmitted by the source or another relay, to reencode the signal and to transmit it to another relay (*multihop*) or the destination terminal (*two-hop*). Or the relay simply forwards a linearly modified version of the noisy signal. These relaying strategies are referred to as decode-and-forward and amplify-and-forward, respectively. Currently, the simple amplify-and-forward approach seems to be promising in many practical applications, e.g., since it is power efficient, does not introduce decoding delay and achieves optimal diversity in many settings. Another approach is the so called compress-and-forward strategy, which quantizes the received signal and reencodes the resulting samples efficiently.

Manuscript received December 22, 2006; revised March 01, 2008. Current version published November 21, 2008. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Nice, France, June 2007.

The authors are with the Swiss Federal Institute of Technology (ETH) Zurich, CH-8092 Zurich, Switzerland.

Communicated by A. J. Grant, Associate Editor for Communications.

Color versions of Figures 2–5 in this correspondence are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2008.2006385

We briefly give an overview over important contributions to the field of cooperative communications and relaying. The capability of relays to provide diversity for combating multipath fading has been studied in [1], [2], and [3]. In [4], the potential of spatial multiplexing gain enhancement in correlated fading channels by means of relays has been demonstrated. Tight upper and lower bounds on the capacity of the fading relay channel are provided in [5]–[9]. Furthermore, in [10] the capacity has been shown to scale like $N \log K$ for the fading MIMO relay channel, where N is the number of source and destination antennas and K is the number of relays.

In this paper, we focus on the two-hop amplify-and-forward MIMO relay channel with either independent and identically distributed (i.i.d.) or Kronecker correlated Rayleigh-fading channels. Our quantities of interest are the cumulant moments of the mutual information of this channel, when the input vector is complex Gaussian and i.i.d. over time. Of particular importance in this context are its mean and variance. While the mean completely determines the long term achievable rate in a fast fading communication channel, the variance is crucial for the characterization of the outage capacity of a channel, which is commonly the quantity of interest in slow fading channels. Seeking for closed form expressions of the cumulant moments of the mutual information in MIMO systems usually is a difficult task. For the conventional point-to-point MIMO channel, it turned out to be useful to defer the analysis to the regime of large antenna numbers. For the i.i.d. Rayleigh-fading MIMO channel, closed-form expressions for the mean were obtained in [11] and [12] in this regime. For correlated fading at either transmitter or receiver side, the mean was derived in [13]. Reference [14] derives various MIMO channel models based on different levels of *a priori* knowledge about the channel and provides expressions for the corresponding mean mutual information each. Reference [15] finally provided the mean for the case of i.i.d. Rayleigh fading and MIMO interference. All these results are obtained via the deterministic asymptotic eigenvalue spectra of the receive signal and noise covariance matrices.

Higher moments were also considered, e.g., in [16], [17], [18], and [14], where the distribution of the mutual information in the large antenna limit was identified to be Gaussian each. Generally, these large array results turned out to be very tight approximations of the respective quantities in finite dimensional systems. Recently, also for two-hop amplify-and-forward MIMO relay channels progress has been achieved in the large array limit. The mean mutual information of Rayleigh fading amplify-and-forward MIMO relay channels in the large array limit has been studied in [10] for case of a forwarding matrix proportional to the identity matrix and channel matrices with i.i.d. elements of zero-mean. In this paper, a fourth order equation for the Stieltjes transform of the corresponding asymptotic eigenvalue spectrum is found, which allows for a numerical evaluation of the mean mutual information. Since even for this special case no analytic solution has been obtained so far via asymptotic eigenvalue spectra, we choose an alternative approach.

The key tool enabling the evaluation of the cumulant moments of the mutual information in the large array limit in this paper is the so called replica method. It was introduced by Edwards and Anderson in [19] and has its origins in physics, where it is applied to large random systems, as they arise, e.g., in statistical mechanics. In the context of channel capacity, it was applied by Tanaka in [20] for the first time. Moustakas *et al.* [18] finally used a framework utilizing the replica trick developed in [21] to evaluate the cumulant moments of the mutual information of

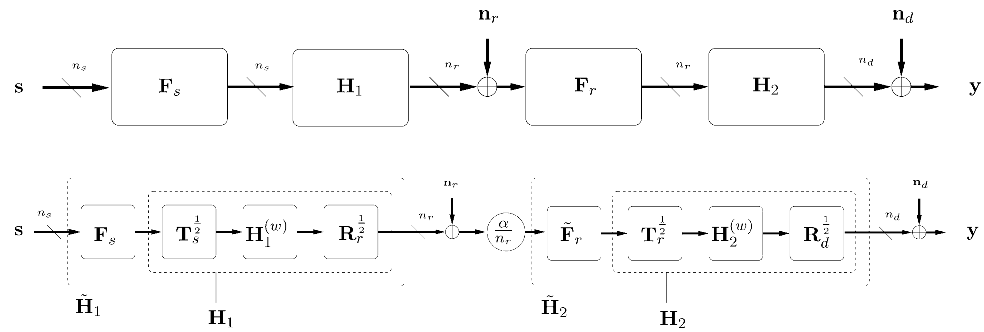

Fig. 1. Block diagrams of the channel.

the Rayleigh fading MIMO channel in the presence of correlated interference. The respective paper [18] is formulated in a very explicatory way, and our correspondence goes very much along the lines of this reference. Though not being proven in a rigorous way yet, the replica method is a particularly attractive tool when dealing with functions of large random matrices, since it allows for the evaluation of arbitrary moments.

There are also some large array results by Müller based on free probability theory that are of importance for amplify-and-forward MIMO relay channels. The considered concatenated fading vector channels in [22] (two hops) and [23] (limit of infinitely many hops) can be considered as multi-hop MIMO channels with noiseless relays.

The contributions of this paper are summarized as follows.

- In the large array limit, we derive mean and variance of the mutual information of the two-hop amplify-and-forward MIMO relay channel for Kronecker correlated Rayleigh fading and beamforming/forwarding matrices that are independent of the channel realization and constant over time. The obtained expressions depend on coefficients that are determined by a system of six nonlinear equations.
- We show that all higher cumulant moments are $\mathcal{O}(n^{-1})$ or smaller and thus vanish as n grows large. Accordingly, we conclude that the mutual information is Gaussian distributed with mean and variance given by our derived expressions in the large n limit.
- Considering that not all doubts about the replica method are dispelled yet, we verify the obtained expressions by means of computer simulations and thus confirm that the replica method indeed works out in our problem.

II. The Channel and Its Mutual Information

The two-hop amplify-and-forward MIMO relay channel under consideration is defined as follows. Three terminals are equipped with n_s (source), n_r (relay), and n_d (destination) antennas, respectively. We allow for communication from source to relay and from relay to destination. Particularly, we do not allot a direct communication link between source and destination. Both the uplink (first hop from source to relay) and the downlink (second hop from relay to destination) are modeled as frequency-flat, i.e., the transmit symbol duration is much longer than the delay spread of up- and downlink. We denote the channel matrix of the uplink by $\mathbf{H}_1 \in \mathbb{C}^{n_r \times n_s}$, the one of the downlink by $\mathbf{H}_2 \in \mathbb{C}^{n_d \times n_r}$. Furthermore, we assume that the relays process the received signals linearly. The matrix performing this linear mapping is denoted $\mathbf{F}_r \in \mathbb{C}^{n_r \times n_r}$ and called the "forwarding matrix" in the following.

With $\mathbf{s}$ the transmit symbol vector, a precoding matrix $\mathbf{F}_s \in \mathbb{C}^{n_s \times n_s}$ and $\mathbf{n}_r$ and $\mathbf{n}_d$ the relay and destination noise vectors respectively, the end-to-end input–output-relation of this channel is then given by

$$\mathbf{y} = \mathbf{H}_2\mathbf{F}_r\mathbf{H}_1\mathbf{F}_s\mathbf{s} + \mathbf{H}_2\mathbf{F}_r\mathbf{n}_r + \mathbf{n}_d.$$

The system is depicted in the upper block diagram in Fig. 1.

The elements of the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are assumed to be zero-mean circular symmetric complex Gaussian (ZMCSCG) random variables with covariance matrices as defined in the Kronecker model [24]

$$\mathsf{E}[\mathrm{vec}(\mathbf{H}_1)\mathrm{vec}(\mathbf{H}_1)^H] = \mathbf{T}_s^T \otimes \mathbf{R}_r \quad (1)$$
$$\mathsf{E}[\mathrm{vec}(\mathbf{H}_2)\mathrm{vec}(\mathbf{H}_2)^H] = \mathbf{T}_r^T \otimes \mathbf{R}_d \quad (2)$$

where $\mathrm{vec}(\mathbf{X})$ stacks $\mathbf{X}$ into a vector columnwise, $\otimes$ denotes the Kronecker product, while $(\cdot)^H$ and $(\cdot)^T$ denote the Hermitian transpose and transpose operator, respectively. $\mathbf{T}_s \in \mathbb{C}^{n_s \times n_s}$, $\mathbf{R}_r \in \mathbb{C}^{n_r \times n_r}$, $\mathbf{T}_r \in \mathbb{C}^{n_r \times n_r}$ and $\mathbf{R}_d \in \mathbb{C}^{n_d \times n_d}$ are the (positive definite) spatial correlation matrices of the antenna arrays at the respective terminals. These matrices are required to have full rank for the analysis below. We remind the reader that matrices whose elements are ZMCSCG distributed such that (1) and (2) are fulfilled, can be generated from matrices $\mathbf{H}^{(w)}$ with i.i.d. ZMCSCG elements each—in the case at hand through the mappings

$$\mathbf{H}_1 = \mathbf{R}_r^{\frac{1}{2}}\mathbf{H}_1^{(w)}\mathbf{T}_s^{\frac{1}{2}}$$

and

$$\mathbf{H}_2 = \mathbf{R}_d^{\frac{1}{2}}\mathbf{H}_2^{(w)}\mathbf{T}_r^{\frac{1}{2}}.$$

The above described correlation model thus assumes separable transmit and receive correlations. This assumption is frequently made in the MIMO literature, since it usually simplifies analytic treatment significantly. However, we emphasize that it is not always supported in reality (cf. [25] for a discussion on the deficiencies of the model). A more general correlation model has been proposed by Weichselberger [26], for instance.

We assume all channel matrix elements to be constant during a certain interval and to change independently from interval to interval (block fading). The input symbols are chosen to be i.i.d. ZMCSCGs with variance ρ/n_s, i.e., $\mathsf{E}[\mathbf{s}\mathbf{s}^H] = \rho/n_s\mathbf{I}_{n_s}$, the additive noise at relay and destination is assumed to be white in both space and time and is modeled as ZMCSCG with unit variance, i.e., $\mathsf{E}[\mathbf{n}_r\mathbf{n}_r^H] = \mathbf{I}_{n_r}$ and $\mathsf{E}[\mathbf{n}_d\mathbf{n}_d^H] = \mathbf{I}_{n_d}$.

The assumptions on the channel state information (CSI) are as follows. The destination perfectly knows the instantaneous channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ as well as $\mathbf{F}_s$ and $\mathbf{F}_r$. The source and the relay only know the second order statistics of $\mathbf{H}_1$ and $\mathbf{H}_2$, i.e., the corresponding covariance matrices. This implies, that $\mathbf{F}_s$ and $\mathbf{F}_r$—and thus beamforming and power loading—can only depend on the covariance matrices of $\mathbf{H}_1$ and $\mathbf{H}_2$, but not on the instantaneous channel realizations. Such a dependence could not be captured by our analysis, since it cannot be expressed as a Kronecker type of correlation. The elements of $\mathbf{F}_s$ and $\mathbf{F}_r$ thus are deterministic and remain constant over time. It is useful to decompose the forwarding matrix into a scaling factor $\sqrt{\alpha/n_r}$ and a matrix $\tilde{\mathbf{F}}_r$ fulfilling $\mathrm{Tr}\{\tilde{\mathbf{F}}_r\tilde{\mathbf{F}}_r^H\} = n_r$, where $\mathrm{Tr}(\cdot)$

$$
\mathbf{V}_1 = \begin{pmatrix}
v_1^{(1)} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -v_4^{(1)} & 0 & v_3^{(1)} & 0 & v_8^{(1)} & 0 & v_8^{(1)} & 0 & v_4^{(1)} \\
0 & 0 & v_2^{(1)} & 1 & v_2^{(1)} & 0 & v_2^{(1)} & 0 & v_2^{(1)} & 0 \\
0 & v_3^{(1)} & 1 & v_5^{(1)} & 0 & v_9^{(1)} & 0 & v_9^{(1)} & 0 & v_6^{(1)} \\
0 & 0 & v_2^{(1)} & 0 & v_2^{(1)} & 1 & v_2^{(1)} & 0 & v_2^{(1)} & 0 \\
0 & v_8^{(1)} & 0 & v_9^{(1)} & 1 & v_6^{(1)} & 0 & v_{10}^{(1)} & 0 & v_{11}^{(1)} \\
0 & 0 & v_2^{(1)} & 0 & v_2^{(1)} & 0 & v_2^{(1)} & 1 & v_2^{(1)} & 0 \\
0 & v_8^{(1)} & 0 & v_9^{(1)} & 0 & v_{10}^{(1)} & 1 & v_6^{(1)} & 0 & v_{11}^{(1)} \\
0 & 0 & v_2^{(1)} & 0 & v_2^{(1)} & 0 & v_2^{(1)} & 0 & v_2^{(1)} & 1 \\
0 & v_4^{(1)} & 0 & v_6^{(1)} & 0 & v_{11}^{(1)} & 0 & v_{11}^{(1)} & 1 & v_7^{(1)}
\end{pmatrix}
$$

$$
\mathbf{V}_2 = \begin{pmatrix} v_1^{(2)} & 1 \\ 1 & v_2^{(2)} \end{pmatrix}
$$

$$
\mathbf{V}_3 = \begin{pmatrix}
1 & -v_1^{(3)} & 0 & v_1^{(3)} \\
-v_4^{(3)} & 1 & v_2^{(3)} & 0 \\
0 & v_1^{(3)} & -1 & -v_1^{(3)} \\
v_2^{(3)} & 0 & -v_3^{(3)} & -1
\end{pmatrix} \quad (5)
$$

denotes the trace operator. We denote α as the power gain of the forwarding matrix.

With $\mathrm{Tr}\{\mathbf{F}_s\mathbf{F}_s^H\} = n_s$ the mutual information[1] conditioned on $\mathbf{H}_1$ and $\mathbf{H}_2$ in nats per channel use can be written as

$$
I = \ln\left(\frac{\det(\mathbf{R}_n + \mathbf{R}_s)}{\det(\mathbf{R}_n)}\right) \quad (3)
$$

where

$$
\mathbf{R}_n = \mathbf{I}_{n_d} + \frac{\alpha}{n_r}\mathbf{H}_2\tilde{\mathbf{F}}_r\tilde{\mathbf{F}}_r^H\mathbf{H}_2^H
$$

corresponds to the noise covariance matrix at the destination and

$$
\mathbf{R}_s = \frac{\rho\cdot\alpha}{n_s n_r}\mathbf{H}_2\tilde{\mathbf{F}}_r\mathbf{H}_1\mathbf{F}_s\mathbf{F}_s^H\mathbf{H}_1^H\tilde{\mathbf{F}}_{\mathrm{rH}}\mathbf{H}_2^H
$$

corresponds to the signal covariance matrix at the destination. Since the forwarding matrix does not depend on the instantaneous channel realizations by assumption, it can be incorporated into $\mathbf{T}_r$ according to

$$
\tilde{\mathbf{T}}_r \triangleq \tilde{\mathbf{F}}_r\mathbf{T}_r\tilde{\mathbf{F}}_r^H.
$$

Similarly $\mathbf{F}_s$ can be incorporated into $\mathbf{T}_s$ according to

$$
\tilde{\mathbf{T}}_s \triangleq \mathbf{F}_s\mathbf{T}_s\mathbf{F}_s^H.
$$

Refer to the lower block diagram in Fig. 1 for an illustration.

In terms of the respective equivalent channel matrices $\tilde{\mathbf{H}}_1 \triangleq \mathbf{R}_r^{\frac{1}{2}}\mathbf{H}_1^{(w)}\tilde{\mathbf{T}}_s^{\frac{1}{2}}$ and $\tilde{\mathbf{H}}_2 \triangleq \mathbf{R}_d^{\frac{1}{2}}\mathbf{H}_2^{(w)}\tilde{\mathbf{T}}_r^{\frac{1}{2}}$, (3) can be rewritten as

$$
I = \ln\left(\frac{\det\left(\mathbf{I}_{n_d} + \frac{\alpha}{n_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H + \frac{\rho\cdot\alpha}{n_s n_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_2^H\right)}{\det\left(\mathbf{I}_{n_d} + \frac{\alpha}{n_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H\right)}\right). \quad (4)
$$

Due to the randomness in $\tilde{\mathbf{H}}_1$ and $\tilde{\mathbf{H}}_2$, also I is a random variable. The theorem stated in the following section fully characterizes the distribution of I in the limit of large antenna numbers.

III. Results

We formulate our results in the subsequent theorem. Whenever we use the notation $\mathcal{O}(f(n))$ in the following, we assume that n_s, n_r and n_d are proportional to n and thus grow to infinity with all ratios among them fixed.

Theorem 1: Assuming that the moment generating function of I can be analytically continued at least in the positive vicinity of zero, as well as $n_s \in \Theta(n)$, $n_r \in \Theta(n)$, and $n_d \in \Theta(n)$ as $n \to \infty$, the mutual information I as stated in (4)

- has a mean which is $\Theta(n)$ and given by

$$
\begin{aligned}
\mathsf{E}[I] = &\ln(\det(\mathbf{I}_{n_s} + \rho s_1\tilde{\mathbf{T}}_s)) \\
&+ \ln(\det(\mathbf{I}_{n_r} + \alpha s_2\mathbf{R}_d)) \\
&- \ln(\det(\mathbf{I}_{n_r} + \alpha s_3\tilde{\mathbf{T}}_r)) \\
&- \ln(\det(\mathbf{I}_{n_d} + t_3\mathbf{R}_d)) \\
&+ \ln(\det(\mathbf{I}_{n_r} + t_2\mathbf{T}_r + t_1t_2\mathbf{R}_r\mathbf{T}_r)) \\
&- (n_s s_1 t_1 + n_r s_2 t_2 - n_r s_3 t_3) + \mathcal{O}(n^{-1})
\end{aligned}
$$

  with

$$
\begin{aligned}
t_1 &= \frac{1}{n_s}\mathrm{Tr}\left\{\rho\tilde{\mathbf{T}}_s\left[\mathbf{I}_{n_s} + \rho s_1\tilde{\mathbf{T}}_s\right]^{-1}\right\} \\
t_2 &= \frac{1}{n_r}\mathrm{Tr}\left\{\alpha\mathbf{R}_d\left[\mathbf{I}_{n_r} + \alpha s_2\mathbf{R}_d\right]^{-1}\right\} \\
t_3 &= \frac{1}{n_r}\mathrm{Tr}\left\{\alpha\tilde{\mathbf{T}}_r\left[\mathbf{I}_{n_r} + \alpha s_3\tilde{\mathbf{T}}_r\right]^{-1}\right\} \\
s_1 &= \frac{1}{n_s}\mathrm{Tr}\left\{t_2\mathbf{R}_r\tilde{\mathbf{T}}_r \right. \\
&\quad\left. \times\left[\mathbf{I}_{n_r} + t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r + t_2\tilde{\mathbf{T}}_r\right]^{-1}\right\} \\
s_2 &= \frac{1}{n_r}\mathrm{Tr}\left\{(\tilde{\mathbf{T}}_r + t_2\mathbf{R}_r\tilde{\mathbf{T}}_r) \right. \\
&\quad\left. \times\left[\mathbf{I}_{n_r} + t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r + t_2\tilde{\mathbf{T}}_r\right]^{-1}\right\} \\
s_3 &= \frac{1}{n_r}\mathrm{Tr}\left\{\mathbf{R}_d\left[\mathbf{I}_{n_d} + t_3\mathbf{R}_d\right]^{-1}\right\};
\end{aligned}
$$

- has a variance which is $\mathcal{O}(1)$ and given by

$$
\mathsf{Var}[I] = -\ln|\det(\mathbf{V}_1)| - \ln|\det(\mathbf{V}_2)| + 2\ln|\det(\mathbf{V}_3)| + \mathcal{O}(n^{-2})
$$

  with $\mathbf{V}_1$, $\mathbf{V}_2$, and $\mathbf{V}_3$ given in (5) at the top of the page and

$$
\begin{aligned}
v_1^{(1)} &= \frac{1}{n_s^2}\mathrm{Tr}\left\{\left(\rho\tilde{\mathbf{T}}_s\left[\mathbf{I}_{n_s} + \rho s_1\tilde{\mathbf{T}}_s\right]^{-1}\right)^2\right\} \\
v_2^{(1)} &= \frac{1}{n_r^2}\mathrm{Tr}\left\{\left(\alpha\mathbf{R}_d\left[\mathbf{I}_{n_d} + \alpha s_2\mathbf{R}_d\right]^{-1}\right)^2\right\}
\end{aligned}
$$

[1]If the relay is assumed to be half-duplex, that is, not to be able to transmit and receive at the same time, a pre-log factor $1/2$ would be required in order to account for the use of two time slots.

$$v_3^{(1)} = -\mathrm{Tr}\left\{\mathbf{R}_r\tilde{\mathbf{T}}_r\left(\left[\mathbf{I}_{n_r}+t_2\tilde{\mathbf{T}}_r\right]\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_4^{(1)} = -\mathrm{Tr}\left\{\left(t_2\mathbf{R}_r\tilde{\mathbf{T}}_r\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_5^{(1)} = \mathrm{Tr}\left\{\left(t_1\mathbf{R}_r\tilde{\mathbf{T}}_r\left[\mathbf{I}_{n_r}+t_2\tilde{\mathbf{T}}_r\right]\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_6^{(1)} = \mathrm{Tr}\left\{\left(t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r^2\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_7^{(1)} = \mathrm{Tr}\left\{\left(\left[\tilde{\mathbf{T}}_r+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r^2\right]\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_8^{(1)} = \mathrm{Tr}\left\{t_2\mathbf{R}_r\tilde{\mathbf{T}}_r^2\left[\mathbf{I}_{\mathrm{n_r}}+t_2\tilde{\mathbf{T}}_r\right]\times\left(\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_9^{(1)} = -\mathrm{Tr}\left\{t_2\tilde{\mathbf{T}}_r\left[\mathbf{I}_{n_r}+t_2\tilde{\mathbf{T}}_r\right]\left(t_1\mathbf{R}_r\tilde{\mathbf{T}}_r\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_{10}^{(1)} = \mathrm{Tr}\left\{t_1\mathbf{R}_r\tilde{\mathbf{T}}_r^2\left[\mathbf{I}_{n_r}+t_2\tilde{\mathbf{T}}_r\right]\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r\right]\times\left(\left[\mathbf{I}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_{11}^{(1)} = -\mathrm{Tr}\left\{t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r^3\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r\right]\times\left(\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_1^{(2)} = \frac{1}{n_r^2}\mathrm{Tr}\left\{\left(\alpha\tilde{\mathbf{T}}_r\left[\mathbf{I}_{n_r}+\alpha s_3\tilde{\mathbf{T}}_r\right]^{-1}\right)^2\right\}$$

$$v_2^{(2)} = \mathrm{Tr}\left\{\left(\mathbf{R}_d\left[\mathbf{I}_{n_d}+t_3\mathbf{R}_d\right]^{-1}\right)^2\right\}$$

$$v_1^{(3)} = \frac{1}{n_r}\mathrm{Tr}\left\{\alpha\mathbf{R}_d^2\left[\mathbf{I}_{n_d}+t_3\mathbf{R}_d\right]^{-1}\times\left[\mathbf{I}_{n_d}+\alpha s_2\mathbf{R}_d\right]^{-1}\right\}$$

$$v_2^{(3)} = \frac{1}{n_r}\mathrm{Tr}\left\{\alpha t_1t_2\tilde{\mathbf{T}}_r^3\mathbf{R}_r\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\times\left[\mathbf{I}_{n_r}+\alpha s_3\tilde{\mathbf{T}}_r\right]^{-1}\right\}$$

$$v_3^{(3)} = \frac{1}{n_r}\mathrm{Tr}\left\{\alpha\tilde{\mathbf{T}}_r^2\left[\mathbf{I}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r\right]\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\times\left[\mathbf{I}_{n_r}+\alpha s_3\tilde{\mathbf{T}}_r\right]^{-1}\right\}$$

$$v_4^{(3)} = \frac{1}{n_r}\mathrm{Tr}\left\{\alpha\tilde{\mathbf{T}}_r^2\left[t_1\mathbf{R}_r+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r\right]\times\left[\mathbf{I}_{n_r}+t_1t_2\mathbf{R}_r\tilde{\mathbf{T}}_r+t_2\tilde{\mathbf{T}}_r\right]^{-1}\times\left[\mathbf{I}_{n_r}+\alpha s_3\tilde{\mathbf{T}}_r\right]^{-1}\right\};$$

- is a Gaussian distributed random variable (and thus fully determined by mean and variance) in the large n limit.

Note that the asymptotic Gaussianity of I is an immediate consequence of a result in [27] in the special case that $\tilde{\mathbf{T}}_s = \mathbf{I}_{n_s}$, $\tilde{\mathbf{T}}_r = \tilde{\mathbf{R}}_r = \mathbf{I}_{n_r}$ and $\tilde{\mathbf{R}}_d = \mathbf{I}_{n_d}$. Interestingly, the expression obtained for the mean can be written in closed form in the the special case of a forwarding matrix proportional to the identity matrix and channel matrices with i.i.d. elements.

IV. Mathematical Tools

In this section we briefly repeat the mathematical tools we use in the proof of the theorem. These are (cumulant) moment generating functions, the replica method and saddle point integration. At the same time we shall give a brief outline of the proof, which we provide in full detail in Section V.

A. Generating Functions

We define the *moment generating function* of the mutual information I as follows:

$$g_I(\nu) = \mathsf{E}[e^{-\nu I}]. \tag{6}$$

This definition differs from the standard definition in the sign of the argument of the exponential function. The minus sign used in the definition above will simplify notation later on. Given that the moment generating function exists in the vicinity of $\nu = 0$, we may expand (6) into a series as follows:

$$g_I(\nu) = 1 - \nu\cdot\mathsf{E}[I] + \frac{\nu^2}{2}\cdot\mathsf{E}[I^2] - \frac{\nu^3}{6}\cdot\mathsf{E}[I^3] + \dots.$$

We will also consider the *cumulant generating function* of I, which is defined as $\ln g_I(\nu)$ and can be expanded into the following series:

$$\ln(g_I(\nu)) = -\nu\cdot\mathsf{E}[I] + \frac{\nu^2}{2}\cdot\mathsf{Var}[I] + \sum_{p=3}^{\infty}\frac{(-\nu)^p}{p!}\mathcal{C}_p \tag{7}$$

with $\mathcal{C}_p$ the pth cumulant moment of I. Once we have found this series, it is easy to extract mean and variance of I by a simple comparison of coefficients. Furthermore, since a Gaussian random variable has the unique property that only a finite number of its cumulants are nonzero (more precisely its mean and variance), we will be able to prove the asymptotic Gaussianity of I by showing that the cumulants $\mathcal{C}_p$ vanish in the large n limit for all $p > 2$.

B. Integral Identities

We will need some useful integral identities in order to evaluate the moment generating function. Before stating them, we introduce a compact notation for products of differentials arising when integration over elements of matrices is performed. With $\imath = \sqrt{-1}$ as well as $\Re\{Z\}$ and $\Im\{Z\}$ the real and imaginary parts of a complex variable Z, we define the following integral measures, which coincide with those introduced in [18]:

$$d_c\mathbf{Z} \triangleq \frac{1}{2\pi}\prod_i\prod_j d\Re\{Z_{ij}\}d\Im\{Z_{ij}\}$$

for Z_{ij} complex variables

$$d_g\mathbf{A} \triangleq \prod_i\prod_j dA_{ij}d\overline{A}_{ij}$$

for A_{ij} and $\overline{A}_{ij}$ Grassmann variables

$$d\mu\left(\mathbf{Z}^{(a)},\mathbf{Z}^{(b)}\right) \triangleq \frac{1}{2\pi\imath}\prod_i\prod_j dZ_{ij}^{(a)} dZ_{ji}^{(b)}$$
for $Z_{ij}^{(a)}$ and $Z_{ij}^{(b)}$ complex variables.

The defining properties of a Grassmann variable are listed in Appendix A. With this notation as well as $\otimes$ the Kronecker product operator we specify the following identities, which are proven in [18].

- For $\mathbf{M}\in\mathbb{C}^{n\times n}$, $\mathbf{N}\in\mathbb{C}^{\nu\times\nu}$ positive definite, $\mathbf{O}\in\mathbb{C}^{\nu\times n}$ and $\mathbf{Z},\mathbf{P}\in\mathbb{C}^{n\times\nu}$, we have

$$\int \exp\left(-\frac{1}{2}\mathrm{Tr}(\mathbf{N}\mathbf{Z}^H\mathbf{M}\mathbf{X}+\mathbf{O}\mathbf{Z}-\mathbf{Z}^H\mathbf{P})\right)d_c\mathbf{Z} = (\det(\mathbf{N}\otimes\mathbf{M}))^{-1}\times\exp\left(-\frac{1}{2}\mathrm{Tr}(\mathbf{N}^{-1}\mathbf{O}\mathbf{M}^{-1}\mathbf{P})\right). \quad (8)$$

- For $\mathbf{M}\in\mathbb{C}^{n\times n}$, $\mathbf{N}\in\mathbb{C}^{\nu\times\nu}$ positive definite, and $\overline{\mathbf{O}},\overline{\mathbf{A}}$ and $\mathbf{A},\mathbf{P}$ $n\times\nu$ and $\nu\times n$ matrices, respectively, whose entries are Grassmann variables, we have

$$\int \exp(\mathrm{Tr}(\mathbf{N}\overline{\mathbf{A}}\mathbf{M}\mathbf{A}+\overline{\mathbf{O}}\mathbf{A}+\overline{\mathbf{A}}\mathbf{P}))d_g\mathbf{A} = \det(\mathbf{N}\otimes\mathbf{M})\exp(\mathrm{Tr}\,(\mathbf{N}^{-1}\overline{\mathbf{O}}\mathbf{M}^{-1}\mathbf{P})). \quad (9)$$

- For $\mathbf{Z}^{(a)},\mathbf{Z}^{(b)},\mathbf{O},\mathbf{P}\in\mathbb{C}^{\nu\times\nu}$, we have

$$\int \exp(\mathrm{Tr}(\mathbf{Z}^{(a)}\mathbf{Z}^{(b)}-\mathbf{Z}^{(a)}\mathbf{O}-\mathbf{P}\mathbf{Z}^{(b)}))d\mu(\mathbf{Z}^{(a)},\mathbf{Z}^{(b)}) = \exp(-\mathrm{Tr}(\mathbf{O}\mathbf{P})). \quad (10)$$

The application of these identities is known as the *replica trick*, which introduces multiple copies of the Gaussian integration that arises when computing the expectation of $\exp(-\nu I)$ over the elements of $\mathbf{H}_1$ and $\mathbf{H}_2$. We emphasize that the machinery of repeatedly applying the above identities in the evaluation of $g_I(\nu)$ (see Section V-A) requires ν to be a positive integer. In order to extract the (cumulant) moments of I from the respective generating function, we thus need to assume that $g_I(\nu)$ can be analytically continued at least in the positive vicinity of zero in the end. This assumption is applied without being proven anywhere in the literature yet. Nevertheless, all results obtained based on this assumption—including those derived below—show a good match with results obtained through computer simulations.

C. Saddle Point Integration

For the final evaluation of the moment generating function, we use the saddle point method. In its simplest form it is a useful tool to solve integrals of the form

$$\lim_{n\to\infty}\int e^{-n\cdot\Psi(x_1,\ldots,x_k)}\cdot dx_1\cdots dx_k,$$

where $\Psi(\cdot,\ldots,\cdot)$ is some function with well defined Hessian at its global minimum. For the sake of simplicity, we consider the univariate case in this section. In the actual proof of the Theorem we will then deal with integrals over multiple variables. Suppose we can rewrite the moment generating function of I in the form (as done in Section V-A)

$$g_I(\nu) = \int e^{-f(x,\nu,n)}dx$$

for some function f. By expanding f into a Taylor series in x around its global minimum at x_0 we can write

$$g_I(\nu) = e^{-f(x_0,\nu,n)} \times \int e^{-\frac{1}{2}f''(x_0,\nu,n)(x-x_0)^2-\frac{1}{6}f'''(x_0,\nu,n)(x-x_0)^3+\cdots}\cdot dx \quad (11)$$

where $(\,\cdot\,)'$ denotes derivative for x. From this expansion and the respective function $f(\cdot,\nu,n)$ it will be possible to show that (11) can be written as

$$g_I(\nu) = \exp(-\nu\xi_1(x_0)+\nu^2\cdot\xi_2(x_0)) + \sum_{k\geq 3}\nu^k\mathcal{O}(n^{-1})$$

with $\xi_1(\,\cdot\,)$ and $\xi_2(\,\cdot\,)$ functions that we determine in Section V-B. The fact that

$$\mathsf{E}[I^p] = (-1)^p\cdot\left(\frac{d^p}{d\nu^p}g_I(\nu)\right)\Big|_{\nu=0}$$

immediately reveals that the leading terms of mean and variance are given by ξ_1 and ξ_2, respectively. The $\mathcal{O}(n^{-1})$ scaling of the residual terms is proven in Section V-C. Comparing $\ln g_I(\nu)$ to the right hand side of (7) will reveal the higher cumulants to be $\mathcal{O}(n^{-1})$. Remember that we obtained (7) as a series expansion around $\nu=0$. We thus have implicitly assumed that the limit $n\to\infty$ and $\nu\to 0$ can be interchanged. This assumption is noncritical and made without proof in this paper.

In the subsequent sections, we apply this procedure in a multivariate framework. $f(\cdot,\nu,n)$ is then a function of multiple matrices (cf. next subsection), which appear inside trace and determinant operators. We make a symmetry assumption called the hypothesis of *replica symmetry*, namely that all these matrices are proportional to the identity matrix at the global minimum of $f(\cdot,\nu,n)$. This assumption is justified in [28].

We emphasize that it is this saddle point method that makes the following derivations a large n approximation. If we had another tool capable to solve the critical integral for finite n, the procedure could also be applied to obtain non-asymptotic results.

V. Proof

For the sake of clarity we structure the proof into three parts corresponding to the subsections below. In Section A, we repeatedly apply the integral identities stated in Section IV-B and thus bring the cumulant generating function into a form that allows for extracting mean and variance in Section B as well as all higher cumulant moments in Section C in the large n limit each.

A. Applying the Replica Trick

We introduce the auxiliary variables $\mathbf{Z}_1\in\mathbb{C}^{n_s\times\nu}$, $\mathbf{Z}_2\in\mathbb{C}^{n_d\times\nu}$, $\mathbf{Z}_3\in\mathbb{C}^{n_r\times\nu}$, $\mathbf{Z}_4\in\mathbb{C}^{n_r\times\nu}$, $\mathbf{Z}_5\in\mathbb{C}^{n_r\times\nu}$, and $\overline{\mathbf{A}}_1,\mathbf{A}_1$ ($\nu\times n_d$ and $n_d\times\nu$ Grassmann matrices), $\overline{\mathbf{A}}_2,\mathbf{A}_2$ ($\nu\times n_r$ and $n_r\times\nu$ Grassmann matrices) and evaluate the moment generating function of I by means of identities (8)–(10). More precisely, we start with applying the following identities (backwards each) in the following order with parameters as listed below.

- Identity (8) with $\mathbf{Z}=\mathbf{Z}_1$, $\mathbf{M}=\mathbf{I}_{n_d}+\frac{\alpha}{n_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H+\frac{\rho\alpha}{n_sn_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_2^H$, $\mathbf{N}=\mathbf{I}_\nu$, $\mathbf{O}=\mathbf{0}_{\nu\times n_d}$, and $\mathbf{P}=\mathbf{0}_{n_d\times\nu}$;
- Identity (9) with $\mathbf{A}=\mathbf{A}_1$, $\mathbf{M}=\mathbf{I}_{n_d}+\frac{\alpha}{n_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H$, $\mathbf{N}=\mathbf{I}_\nu$, $\mathbf{O}=\mathbf{0}_{\nu\times n_d}$, and $\mathbf{P}=\mathbf{0}_{n_d\times\nu}$;
- Identity (8) with $\mathbf{Z}=\mathbf{Z}_2$, $\mathbf{M}=\mathbf{I}_{n_s}$, $\mathbf{N}=\mathbf{I}_\nu$, $\mathbf{O}=\frac{\rho\alpha}{n_sn_r}\mathbf{Z}_1^H\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_2$, and $\mathbf{P}=\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_2^H\mathbf{Z}_1$;
- Identity (8) with $\mathbf{Z}=\mathbf{Z}_3$, $\mathbf{M}=\mathbf{I}_{n_r}$, $\mathbf{N}=\mathbf{I}_\nu$, $\mathbf{O}=\frac{\alpha}{n_r}\mathbf{Z}_1^H\tilde{\mathbf{H}}_2$, and $\mathbf{P}=\tilde{\mathbf{H}}_2^H\mathbf{Z}_1$;
- Identity (9) with $\mathbf{A}=\mathbf{A}_2$, $\mathbf{M}=\mathbf{I}_{n_r}$, $\mathbf{N}=\mathbf{I}_\nu$, $\overline{\mathbf{O}}=-\overline{\mathbf{A}}_1\tilde{\mathbf{H}}_2$, and $\mathbf{P}=\tilde{\mathbf{H}}_2^H\mathbf{A}_1$.

The first two steps allow to get rid of the determinants. Afterwards, we split the products $\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_2^H$ and $\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H$ at the expense of the introduced auxiliary matrices $\mathbf{Z}_2$ and $\mathbf{Z}_3$. The cumulant generating function can thus be rewritten as in (12) shown at the bottom of the next page.

In a next step, we also split the products $\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_1$ and $\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_2^H$ by applying (again in backward direction)

- Identity (8) with $\mathbf{Z}=\mathbf{Z}_4$, $\mathbf{M}=\mathbf{I}_{n_r}$, $\mathbf{N}=\mathbf{I}_\nu$, $\mathbf{O}=\mathbf{Z}_1^H\tilde{\mathbf{H}}_2$, and $\mathbf{P}=-\tilde{\mathbf{H}}_1\mathbf{Z}_2$;
- Identity (8) with $\mathbf{Z}=\mathbf{Z}_5$, $\mathbf{M}=\mathbf{I}_{n_r}$, $\mathbf{N}=\mathbf{I}_\nu$, $\mathbf{O}=\frac{\rho}{n_s}\mathbf{Z}_2^H\tilde{\mathbf{H}}_1^H$, and $\mathbf{P}=\frac{\alpha}{n_r}\tilde{\mathbf{H}}_2^H\mathbf{Z}_1$.

We thus obtain (13) on the bottom of the page.

We are now ready to perform the integration over $\mathbf{H}_1$ and $\mathbf{H}_2$. This can be done by applying (8) twice (in forward direction):

- Identity (8) with $\mathbf{Z}=\mathbf{H}_1$, $\mathbf{M}=\mathbf{I}_{n_r}$, $\mathbf{N}=\mathbf{I}_\nu$, $\mathbf{O}=\mathbf{Z}_2\mathbf{Z}_4^H$, and $\mathbf{P}=-\frac{\rho}{n_s}\mathbf{Z}_5\mathbf{Z}_2^H$;
- Identity (8) with $\mathbf{Z}=\mathbf{H}_2$, $\mathbf{M}=\mathbf{I}_{n_d}$, $\mathbf{N}=\mathbf{I}_\nu$, $\mathbf{O}=\mathbf{Z}_4\mathbf{Z}_1^H-\mathbf{Z}_3\mathbf{Z}_1^H-\mathbf{A}_2\overline{\mathbf{A}}_1$, and $\mathbf{P}=\frac{\alpha}{n_r}(\mathbf{Z}_1\mathbf{Z}_5^H-\mathbf{Z}_1\mathbf{Z}_3^H-\mathbf{A}_1\overline{\mathbf{A}}_2)$.

The cumulant moment generating function then simplifies to (14) at the bottom of the page.

Next, we split all quartic terms into quadratic terms by making use of (9) and (10) as follows.

- Identity (10) with $\mathbf{Z}^{(a)}=\mathbf{R}_1$, $\mathbf{Z}^{(b)}=\mathbf{Q}_1$, $\mathbf{O}=\frac{\rho}{n_s}\mathbf{Z}_2^H\tilde{\mathbf{T}}_s\mathbf{Z}_2$, and $\mathbf{P}=-\mathbf{Z}_4^H\tilde{\mathbf{T}}_r\mathbf{Z}_5$;
- Identity (10) with $\mathbf{Z}^{(a)}=\mathbf{R}_2$, $\mathbf{Z}^{(b)}=\mathbf{Q}_2$, $\mathbf{O}=\frac{\alpha}{n_r}\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1$, and $\mathbf{P}=\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{Z}_4$;
- Identity (10) $\mathbf{Z}^{(a)}=\mathbf{R}_3$, $\mathbf{Z}^{(b)}=\mathbf{Q}_3$, $\mathbf{O}=\frac{\alpha}{n_r}\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1$, and $\mathbf{P}=-\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{Z}_4$;
- Identity (9) with $\overline{\mathbf{A}}=\overline{\mathbf{R}}_4$, $\mathbf{A}=\mathbf{R}_4$, $\overline{\mathbf{O}}=-\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{Z}_4$, and $\mathbf{P}=-2\mathbf{Z}_1^H\mathbf{R}_d\mathbf{A}_1$;
- Identity (10) with $\mathbf{Z}^{(a)}=\mathbf{R}_5$, $\mathbf{Z}^{(b)}=\mathbf{Q}_5$, $\mathbf{O}=\frac{\alpha}{n_r}\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1$, and $\mathbf{P}=-\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{Z}_3$;
- Identity (10) with $\mathbf{Z}^{(a)}=\mathbf{R}_6$, $\mathbf{Z}^{(b)}=\mathbf{Q}_6$, $\mathbf{O}=\frac{\alpha}{n_r}\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1$, and $\mathbf{P}=\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{Z}_3$;
- Identity (9) with $\overline{\mathbf{A}}=-\overline{\mathbf{Q}}_4$, $\mathbf{A}=\mathbf{Q}_4$, $\overline{\mathbf{O}}=\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{Z}_3$, and $\mathbf{P}=-2\mathbf{Z}_1^H\mathbf{R}_d\mathbf{A}_1$;
- Identity (9) with $\overline{\mathbf{A}}=\overline{\mathbf{R}}_7$, $\mathbf{A}=\mathbf{R}_7$, $\overline{\mathbf{O}}=-\frac{\alpha}{n_r}\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{Z}_1$, and $\mathbf{P}=-2\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{A}_2$;
- Identity (9) with $\overline{\mathbf{A}}=-\overline{\mathbf{Q}}_7$, $\mathbf{A}=\mathbf{Q}_7$, $\overline{\mathbf{O}}=\frac{\alpha}{n_r}\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{Z}_1$, and $\mathbf{P}=-2\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{A}_2$;
- Identity (10) with $\mathbf{Z}^{(a)}=\mathbf{R}_8$, $\mathbf{Z}^{(b)}=\mathbf{Q}_8$, $\mathbf{O}=-2\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{A}_2$, and $\mathbf{P}=-\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{A}_1$.

Combining the integral measures with respect to the various $\mathbf{R}_i$'s an $\mathbf{Q}_i$'s into the single integral measure

$$d\lambda \triangleq d\mu(\mathbf{R}_1,\mathbf{Q}_1)d\mu(\mathbf{R}_2,\mathbf{Q}_2)d\mu(\mathbf{R}_3,\mathbf{Q}_3)d\mu(\mathbf{R}_5,\mathbf{Q}_5) \times d\mu(\mathbf{R}_6,\mathbf{Q}_6)d\mu(\mathbf{R}_8,\mathbf{Q}_8)\cdot d\mathbf{R}_4 d\mathbf{Q}_4 d\mathbf{R}_7 d\mathbf{Q}_7.$$

this results into

$$g_I(\nu)=\int\exp\left(-\frac{1}{2}\mathrm{Tr}\left(\mathbf{Z}_1^H\mathbf{Z}_1+\mathbf{Z}_2^H\mathbf{Z}_2+\mathbf{Z}_3^H\mathbf{Z}_3+\mathbf{Z}_4^H\mathbf{Z}_4 +\mathbf{Z}_5^H\mathbf{Z}_5-2\overline{\mathbf{A}}_1\mathbf{A}_1-2\overline{\mathbf{A}}_2\mathbf{A}_2\right)\right)$$

$$\begin{aligned}
g_I(\nu) &= \mathsf{E}[e^{-\nu I}]\\
&= \mathsf{E}\left[\left\{\frac{\det\left(\mathbf{I}_{n_d}+\frac{\alpha}{n_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H+\frac{\rho\alpha}{n_s n_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_2^H\right)}{\det\left(\mathbf{I}_{n_d}+\frac{\alpha}{n_r}\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_2^H\right)}\right\}^{-\nu}\right]\\
&= \int\int\exp\left(-\frac{1}{2}\mathrm{Tr}\left(\mathbf{Z}_1^H\mathbf{Z}_1+\mathbf{Z}_2^H\mathbf{Z}_2+\mathbf{Z}_3^H\mathbf{Z}_3\right)\right)\\
&\quad\times\exp\left(-\frac{1}{2}\mathrm{Tr}\left(\frac{\rho\alpha}{n_s n_r}\mathbf{Z}_2^H\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_2^H\mathbf{Z}_1-\mathbf{Z}_1^H\tilde{\mathbf{H}}_2\tilde{\mathbf{H}}_1\mathbf{Z}_2+\frac{\alpha}{n_r}\mathbf{Z}_3^H\tilde{\mathbf{H}}_2^H\mathbf{Z}_1-\mathbf{Z}_1^H\tilde{\mathbf{H}}_2\mathbf{Z}_3\right)\right)d_c\mathbf{Z}_1 d_c\mathbf{Z}_2 d_c\mathbf{Z}_3\\
&\quad\times\int\exp\left(\mathrm{Tr}\left(\overline{\mathbf{A}}_1\mathbf{A}_1+\overline{\mathbf{A}}_2\mathbf{A}_2+\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{H}}_2^H\mathbf{A}_1-\overline{\mathbf{A}}_1\tilde{\mathbf{H}}_2\mathbf{A}_2\right)\right)d_g\mathbf{A}_1 d_g\mathbf{A}_2\\
&\quad\times\exp\left(-\mathrm{Tr}\left(\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1+\tilde{\mathbf{H}}_2^H\tilde{\mathbf{H}}_2\right)\right)d\mathbf{H}_1^{(w)}d\mathbf{H}_2^{(w)} \qquad (12)\\
&= \int\int\int\exp\left(-\frac{1}{2}\mathrm{Tr}\left(\mathbf{Z}_1^H\mathbf{Z}_1+\mathbf{Z}_2^H\mathbf{Z}_2+\mathbf{Z}_3^H\mathbf{Z}_3+\mathbf{Z}_4^H\mathbf{Z}_4+\mathbf{Z}_5^H\mathbf{Z}_5\right)\right)\\
&\quad\times\exp\left(-\frac{1}{2}\mathrm{Tr}\left(\frac{\rho}{n_s}\mathbf{Z}_2^H\tilde{\mathbf{H}}_1^H\mathbf{Z}_5-\frac{\alpha}{n_r}\mathbf{Z}_5^H\tilde{\mathbf{H}}_2^H\mathbf{Z}_1+\mathbf{Z}_1^H\tilde{\mathbf{H}}_2\mathbf{Z}_4+\mathbf{Z}_4^H\tilde{\mathbf{H}}_1\mathbf{Z}_2\right)\right)d_c\mathbf{Z}_4 d_c\mathbf{Z}_5\\
&\quad\times\exp\left(-\frac{1}{2}\mathrm{Tr}\left(\frac{\alpha}{n_r}\mathbf{Z}_3^H\tilde{\mathbf{H}}_2^H\mathbf{Z}_1-\mathbf{Z}_1^H\tilde{\mathbf{H}}_2\mathbf{Z}_3\right)\right)d_c\mathbf{Z}_1 d_c\mathbf{Z}_2 d_c\mathbf{Z}_3\\
&\quad\times\int\exp\left(\mathrm{Tr}\left(\overline{\mathbf{A}}_1\mathbf{A}_1+\overline{\mathbf{A}}_2\mathbf{A}_2+\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{H}}_2^H\mathbf{A}_1-\overline{\mathbf{A}}_1\tilde{\mathbf{H}}_2\mathbf{A}_2\right)\right)d_g\mathbf{A}_1 d_g\mathbf{A}_2\\
&\quad\times\exp\left(-\mathrm{Tr}\left(\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1+\tilde{\mathbf{H}}_2^H\tilde{\mathbf{H}}_2\right)\right)d\mathbf{H}_1^{(w)}d\mathbf{H}_2^{(w)} \qquad (13)\\
&= \int\exp\left(-\frac{1}{2}\mathrm{Tr}\left(\mathbf{Z}_1^H\mathbf{Z}_1+\mathbf{Z}_2^H\mathbf{Z}_2+\mathbf{Z}_3^H\mathbf{Z}_3+\mathbf{Z}_4^H\mathbf{Z}_4+\mathbf{Z}_5^H\mathbf{Z}_5-2\overline{\mathbf{A}}_1\mathbf{A}_1-2\overline{\mathbf{A}}_2\mathbf{A}_2\right)\right)\\
&\quad\times\exp\left(-\frac{1}{4}\mathrm{Tr}\left(-\frac{\rho}{n_s}\mathbf{Z}_2^H\tilde{\mathbf{T}}_s\mathbf{Z}_2\mathbf{Z}_4^H\tilde{\mathbf{T}}_r\mathbf{Z}_5+\frac{\alpha}{n_r}\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{Z}_4-\frac{\alpha}{n_r}\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{Z}_4\right)\right)\\
&\quad\times\exp\left(-\frac{1}{4}\mathrm{Tr}\left(2\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{Z}_4\mathbf{Z}_1^H\mathbf{R}_d\mathbf{A}_1-\frac{\alpha}{n_r}\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{Z}_3+\frac{\alpha}{n_r}\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{Z}_3\right)\right)\\
&\quad\times\exp\left(-\frac{1}{4}\mathrm{Tr}\left(-2\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{Z}_3\mathbf{Z}_1^H\mathbf{R}_d\mathbf{A}_1+2\frac{\alpha}{n_r}\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{Z}_1\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{A}_2-2\frac{\alpha}{n_r}\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{Z}_1\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{A}_2\right.\right.\\
&\qquad\left.\left.-4\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{A}_2\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{A}_1\right)\right)d_c\mathbf{Z}_4 d_c\mathbf{Z}_5 d_c\mathbf{Z}_1 d_c\mathbf{Z}_2 d_c\mathbf{Z}_3 d_g\mathbf{A}_1 d_g\mathbf{A}_2 \qquad (14)
\end{aligned}$$

$$
\begin{aligned}
&\times\ \exp(\mathrm{Tr}(\mathbf{R}_1\mathbf{Q}_1+\mathbf{R}_2\mathbf{Q}_2+\mathbf{R}_3\mathbf{Q}_3+\overline{\mathbf{R}}_4\mathbf{R}_4\\
&\qquad+\mathbf{R}_5\mathbf{Q}_5+\mathbf{R}_6\mathbf{Q}_6-\overline{\mathbf{Q}}_4\mathbf{Q}_4\\
&\qquad+\overline{\mathbf{R}}_7\mathbf{R}_7-\overline{\mathbf{Q}}_7\mathbf{Q}_7-\mathbf{R}_8\mathbf{Q}_8))\\
&\times\ \exp\left(-\frac{1}{2}\mathrm{Tr}\left(\frac{\rho}{n_s}\mathbf{R}_1\mathbf{Z}_2^H\tilde{\mathbf{T}}_s\mathbf{Z}_2\right.\right.\\
&\left.\left.\quad-\mathbf{Q}_1\mathbf{Z}_4^H\tilde{\mathbf{T}}_r\mathbf{Z}_5+\frac{\alpha}{n_r}\mathbf{R}_2\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1+\mathbf{Q}_2\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{Z}_4\right)\right)\\
&\times\ \exp\left(-\frac{1}{2}\mathrm{Tr}\left(\frac{\alpha}{n_r}\mathbf{R}_3\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1\right.\right.\\
&\left.\left.-\mathbf{Q}_3\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{Z}_4+\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{Z}_4\mathbf{R}_4+2\overline{\mathbf{R}}_4\mathbf{Z}_1^H\mathbf{R}_d\mathbf{A}_1\right)\right)\\
&\times\ \exp\left(-\frac{1}{2}\mathrm{Tr}\left(\frac{\alpha}{n_r}\mathbf{R}_5\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1\right.\right.\\
&\left.\left.-\mathbf{Q}_5\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{Z}_3+\frac{\alpha}{n_r}\mathbf{R}_6\mathbf{Z}_1^H\mathbf{R}_d\mathbf{Z}_1+\mathbf{Q}_6\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{Z}_3\right)\right)\\
&\times\ \exp\left(-\frac{1}{2}\mathrm{Tr}\left(-\frac{\alpha}{n_r}\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{Z}_3\mathbf{Q}_4-2\overline{\mathbf{Q}}_4\mathbf{Z}_1^H\mathbf{R}_d\mathbf{A}_1\right.\right.\\
&\left.\left.+\frac{\alpha}{n_r}\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{Z}_1\mathbf{R}_7+2\overline{\mathbf{R}}_7\mathbf{Z}_5^H\tilde{\mathbf{T}}_r\mathbf{A}_2\right)\right)\\
&\times\ \exp\left(-\frac{1}{2}\mathrm{Tr}\left(-\frac{\alpha}{n_r}\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{Z}_1\mathbf{Q}_7-2\overline{\mathbf{Q}}_7\mathbf{Z}_3^H\tilde{\mathbf{T}}_r\mathbf{A}_2\right.\right.\\
&\left.\left.\qquad-2\frac{\alpha}{n_r}\mathbf{R}_8\overline{\mathbf{A}}_2\tilde{\mathbf{T}}_r\mathbf{A}_2-2\mathbf{Q}_8\overline{\mathbf{A}}_1\mathbf{R}_d\mathbf{A}_1\right)\right)\\
&\times\ d_c\mathbf{Z}_4d_c\mathbf{Z}_5d_c\mathbf{Z}_1d_c\mathbf{Z}_2d_c\mathbf{Z}_3d_g\mathbf{A}_1d_g\mathbf{A}_2d\lambda
\end{aligned}\tag{15}
$$

Finally, we perform the integrations with respect to $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{Z}_3$, $\mathbf{Z}_5$, $\mathbf{Z}_6$, $\mathbf{Z}_8$ $\mathbf{A}_1$ and $\mathbf{A}_2$ analogously to those with respect to $\mathbf{H}_1$ and $\mathbf{H}_2$ by means of identity (8) for the complex integrals and identity (9) for the Grassmanian ones. This is done in four steps. We start out with integrating with respect to $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{Z}_6$, $\mathbf{Z}_8$. Next, we perform the integration with respect to $\mathbf{Z}_5$ and $\mathbf{A}_1$. Then, we continue with $\mathbf{Z}_3$, before integrating with respect to $\mathbf{A}_2$ finally. This yields

$$g_I(\nu)=\int\exp(-\mathcal{S})\cdot d\lambda \tag{16}$$

with

$$
\begin{aligned}
\mathcal{S}=&-\mathrm{Tr}(\mathbf{R}_1\mathbf{Q}_1+\mathbf{R}_2\mathbf{Q}_2+\mathbf{R}_3\mathbf{Q}_3\\
&+\overline{\mathbf{R}}_4\mathbf{R}_4+\mathbf{R}_5\mathbf{Q}_5+\mathbf{R}_6\mathbf{Q}_6\\
&-\overline{\mathbf{Q}}_4\mathbf{Q}_4+\overline{\mathbf{R}}_7\mathbf{R}_7-\overline{\mathbf{Q}}_7\mathbf{Q}_7-\mathbf{R}_8\mathbf{Q}_8)\\
&+\ln\det\left(\mathbf{I}_{n_s\nu}+\frac{\rho}{n_s}\mathbf{R}_1^*\right)\\
&+\ln\det\left(\mathbf{I}_{n_r\nu}+\frac{\alpha}{n_r}(\mathbf{R}_2^*+\mathbf{R}_3^*+\mathbf{R}_5^*+\mathbf{R}_6^*)\right)\\
&+\ln\det\left(\mathbf{I}_{n_r\nu}+\mathbf{Q}_1^*\mathbf{Q}_2^*-\mathbf{Q}_5^*\mathbf{Q}_1^*\mathbf{Q}_3^*+\mathbf{Q}_6^*+\mathbf{Q}_1^*\mathbf{Q}_2^*\mathbf{Q}_6^*\right)\\
&-\ln\det\left(\mathbf{I}_{n_r\nu}+\frac{\alpha}{n_r}\mathbf{R}_8^*\right.\\
&\qquad-\frac{\alpha}{n_r}(\mathbf{I}_{n_r\nu}+\mathbf{Q}_1^*\mathbf{Q}_2^*)^{-1}\overline{\mathbf{R}}_7^*\mathbf{Q}_1^*\mathbf{R}_4^*\\
&\quad+\frac{\alpha}{n_r}\left[\mathbf{I}_{n_r\nu}-(\mathbf{I}_{n_r\nu}+\mathbf{Q}_1^*\mathbf{Q}_2^*)^{-1}\mathbf{Q}_5^*\mathbf{Q}_1^*\mathbf{Q}_3^*+\mathbf{Q}_6^*\right]^{-1}\\
&\qquad\times\left[\overline{\mathbf{R}}_7^*\mathbf{Q}_1^*\mathbf{Q}_3^*\mathbf{Q}_4^*(\mathbf{I}_{n_r\nu}+\mathbf{Q}_1^*\mathbf{Q}_2^*)^{-1}\right.\\
&\qquad-\overline{\mathbf{R}}_7^*\mathbf{Q}_1^*\mathbf{Q}_3^*(\mathbf{I}_{n_r\nu}+\mathbf{Q}_1^*\mathbf{Q}_2^*)^{-1}\\
&\qquad\times\mathbf{Q}_5^*\mathbf{Q}_1^*\mathbf{R}_4^*(\mathbf{I}_{n_r\nu}+\mathbf{Q}_1^*\mathbf{Q}_2^*)^{-1}\\
&\qquad\left.\left.+\overline{\mathbf{Q}}_7^*(\mathbf{I}_{n_r\nu}+\mathbf{Q}_1^*\mathbf{Q}_2^*)^{-1}\mathbf{Q}_5^*\mathbf{Q}_1^*\mathbf{R}_4^*-\overline{\mathbf{Q}}_7^*\mathbf{Q}_4^*\right]\right)\\
&-\ln\det\left(\mathbf{I}_{n_r\nu}+\mathbf{Q}_8^*+\frac{\alpha}{n_r}\left[\mathbf{I}_{n_r\nu}\right.\right.\\
&\qquad\left.+\frac{\alpha}{n_r}(\mathbf{R}_2^*+\mathbf{R}_3^*+\mathbf{R}_5^*+\mathbf{R}_6^*)\right]^{-1}\\
&\qquad\left.\left[\overline{\mathbf{R}}_4^*\mathbf{Q}_7^*-\overline{\mathbf{Q}}_4^*\mathbf{Q}_7^*-\overline{\mathbf{R}}_4^*\mathbf{R}_7^*+\overline{\mathbf{Q}}_4^*\mathbf{R}_7^*\right]\right)
\end{aligned}\tag{17}
$$

as well as $\mathbf{R}_1^*=\mathbf{R}_1\otimes\tilde{\mathbf{T}}_s$, $\mathbf{R}_2^*=\mathbf{R}_2\otimes\mathbf{R}_d$, $\mathbf{R}_3^*=\mathbf{R}_2\otimes\mathbf{R}_d$, $\mathbf{R}_4^*=\mathbf{R}_4\otimes\mathbf{R}_d$, $\overline{\mathbf{R}}_4^*=\overline{\mathbf{R}}_4\otimes\mathbf{R}_d$, $\mathbf{R}_5^*=\mathbf{R}_2\otimes\mathbf{R}_d$, $\mathbf{R}_6^*=\mathbf{R}_2\otimes\mathbf{R}_d$, $\mathbf{R}_7^*=\mathbf{R}_7\otimes\mathbf{R}_d$, $\overline{\mathbf{R}}_7^*=\overline{\mathbf{R}}_7\otimes\mathbf{R}_d$, $\mathbf{R}_8^*=\mathbf{R}_8\otimes\tilde{\mathbf{T}}_r$, $\mathbf{Q}_1^*=\mathbf{Q}_1\otimes\mathbf{R}_r$, $\mathbf{Q}_2^*=\mathbf{Q}_2\otimes\tilde{\mathbf{T}}_r$, $\mathbf{Q}_3^*=\mathbf{Q}_3\otimes\tilde{\mathbf{T}}_r$, $\mathbf{Q}_4^*=\mathbf{Q}_4\otimes\mathbf{R}_d$, $\overline{\mathbf{Q}}_4^*=\overline{\mathbf{Q}}_4\otimes\mathbf{R}_d$, $\mathbf{Q}_5^*=\mathbf{Q}_5\otimes\tilde{\mathbf{T}}_r$, $\mathbf{Q}_6^*=\mathbf{Q}_6\otimes\tilde{\mathbf{T}}_r$, $\mathbf{Q}_7^*=\mathbf{Q}_7\otimes\mathbf{R}_d$, $\overline{\mathbf{Q}}_7^*=\overline{\mathbf{Q}}_7\otimes\mathbf{R}_d$, and $\mathbf{Q}_8^*=\mathbf{Q}_8\otimes\mathbf{R}_d$.

At this point we have shaped the problem into the form of (11), where the role of x is played by the introduced $\nu\times\nu$ auxiliary matrices. Note that there appears no matrix with one of its dimension equal to n_s, n_r or n_d in $\mathcal{S}$ anymore.

B. Evaluating Mean and Variance

In order to evaluate the remaining integral in (16) by means of saddle point integration, we need to expand $\mathcal{S}$ into a Taylor series in $\boldsymbol{\delta}\mathbf{R}_1,\boldsymbol{\delta}\mathbf{Q}_1,\ldots,\boldsymbol{\delta}\mathbf{R}_8,\boldsymbol{\delta}\mathbf{Q}_8$ around its minimum. This expansion corresponds to the expansion in x in Section IV.C. With $\mathcal{S}_p$ denoting the pth order term in the series, the expansion looks as follows

$$\mathcal{S}=\mathcal{S}_0+\mathcal{S}_2+\mathcal{S}_3+\ldots \tag{18}$$

By symmetry all complex matrices are assumed to be proportional to the identity matrix at the minimum of $\mathcal{S}$ (*replica symmetry*). The Grassmann matrices have to vanish in order to obtain a real solution (by definition real numbers cannot be Grassmann numbers, since they commute). Thus, to develop the Taylor series (18) in this point, we write

$$
\begin{array}{ll}
\mathbf{R}_1=r_1n_s\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{R}_1 & \mathbf{Q}_1=q_1\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{Q}_1\\
\mathbf{R}_2=r_2n_r\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{R}_2 & \mathbf{Q}_2=q_2\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{Q}_2\\
\mathbf{R}_3=r_3n_r\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{R}_3 & \mathbf{Q}_3=q_3\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{Q}_3\\
\mathbf{R}_4=\boldsymbol{\delta}\mathbf{R}_4 & \mathbf{Q}_4=\boldsymbol{\delta}\mathbf{Q}_4\\
\overline{\mathbf{R}}_4=\boldsymbol{\delta}\overline{\mathbf{R}}_4 & \overline{\mathbf{Q}}_4=\boldsymbol{\delta}\overline{\mathbf{Q}}_4\\
\mathbf{R}_5=r_5n_r\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{R}_5 & \mathbf{Q}_5=q_5\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{Q}_5\\
\mathbf{R}_6=r_6n_r\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{R}_6 & \mathbf{Q}_6=q_6\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{Q}_6\\
\overline{\mathbf{R}}_7=\boldsymbol{\delta}\mathbf{R}_7 & \overline{\mathbf{Q}}_7=\boldsymbol{\delta}\mathbf{Q}_7\\
\overline{\mathbf{R}}_7=\boldsymbol{\delta}\overline{\mathbf{R}}_7 & \overline{\mathbf{Q}}_7=\boldsymbol{\delta}\overline{\mathbf{Q}}_7\\
\mathbf{R}_8=r_8n_r\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{R}_8 & \mathbf{Q}_8=q_8\mathbf{I}_\nu+\boldsymbol{\delta}\mathbf{Q}_8.
\end{array}
$$

By definition, $\mathcal{S}_0$ is given by (17) evaluated at the minimum of $\mathcal{S}$, i.e.,

$$
\begin{aligned}
\mathcal{S}_0=\nu\cdot&\left[\ln\det\left(\mathbf{I}_{n_s}+\rho r_1\tilde{\mathbf{T}}_s\right)\right.\\
&+\ln\det\left(\mathbf{I}_{n_r}+\alpha(r_2+r_3+r_5+r_6)\mathbf{R}_d\right)\\
&+\ln\det\left(\mathbf{I}_{n_r}+q_1\mathbf{R}_rq_2\tilde{\mathbf{T}}_r-q_5\tilde{\mathbf{T}}_rq_1\mathbf{R}_rq_3\tilde{\mathbf{T}}_r\right.\\
&\left.+q_6\tilde{\mathbf{T}}_r+q_1\mathbf{R}_rq_2\tilde{\mathbf{T}}_rq_6\tilde{\mathbf{T}}_r\right)\\
&-\ln\det\left(\mathbf{I}_{n_r}+\alpha r_8\tilde{\mathbf{T}}_r\right)-\ln\det\left(\mathbf{I}_{n_r}+q_8\mathbf{R}_d\right)\\
&-(n_sr_1q_1+n_rr_2q_2+n_rr_3q_3\\
&\left.+n_rr_5q_5+n_rr_6q_6-n_rr_8q_8)\right].
\end{aligned}\tag{19}
$$

The respective coefficients r_i and q_i have to ensure that $\mathcal{S}_1=0$. They are found by differentiating (19) for each of them and setting the resulting expressions to zero. The derivatives for the r_i's (note that we can summarize $r_2+r_3+r_5+r_6+r_8\triangleq\tilde{r}_2$ by symmetry) yield

$$
\begin{aligned}
0&=q_1-\frac{1}{n_s}\mathrm{Tr}\left\{\rho\tilde{\mathbf{T}}_s\left[\mathbf{I}_{n_s}+\rho r_1\tilde{\mathbf{T}}_s\right]^{-1}\right\}\\
0&=q_2-\frac{1}{n_r}\mathrm{Tr}\left\{\alpha\mathbf{R}_d\left[\mathbf{I}_{n_r}+\alpha\tilde{r}_2\mathbf{R}_d\right]^{-1}\right\}
\end{aligned}
$$

$$0 = q_3 - \frac{1}{n_r}\mathrm{Tr}\left\{\alpha \mathbf{R}_d \left[\mathbf{I}_{n_r} + \alpha \tilde{r}_2 \mathbf{R}_d\right]^{-1}\right\}$$
$$0 = q_5 - \frac{1}{n_r}\mathrm{Tr}\left\{\alpha \mathbf{R}_d \left[\mathbf{I}_{n_r} + \alpha \tilde{r}_2 \mathbf{R}_d\right]^{-1}\right\}$$
$$0 = q_6 - \frac{1}{n_r}\mathrm{Tr}\left\{\alpha \mathbf{R}_d \left[\mathbf{I}_{n_r} + \alpha \tilde{r}_2 \mathbf{R}_d\right]^{-1}\right\}$$
$$0 = q_8 - \frac{1}{n_r}\mathrm{Tr}\left\{\alpha \tilde{\mathbf{T}}_r \left[\mathbf{I}_{n_r} + \alpha r_3 \tilde{\mathbf{T}}_r\right]^{-1}\right\}.$$

We see that $q_2 = q_3 = q_5 = q_6$. Taking this into account the derivatives for the q_i's yield

$$0 = r_1 - \frac{1}{n_s}\mathrm{Tr}\left\{q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r \times \left[\mathbf{I}_{n_r} + q_1 q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r + q_2 \tilde{\mathbf{T}}_r\right]^{-1}\right\}$$
$$0 = \tilde{r}_2 - \frac{1}{n_r}\mathrm{Tr}\left\{(\tilde{\mathbf{T}}_r + q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r) \times \left[\mathbf{I}_{n_r} + q_1 q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r + q_2 \tilde{\mathbf{T}}_r\right]^{-1}\right\}$$
$$0 = r_8 - \frac{1}{n_r}\mathrm{Tr}\left\{\mathbf{R}_d \left[\mathbf{I}_{n_d} + q_8 \mathbf{R}_d\right]^{-1}\right\}.$$

The leading term thus simplifies to

$$\begin{aligned}\mathcal{S}_0 = \nu \cdot \Big[&\ln\det\left(\mathbf{I}_{n_s} + \rho r_1 \tilde{\mathbf{T}}_s\right) \\ &+ \ln\det\left(\mathbf{I}_{n_r} + \alpha \tilde{r}_2 \mathbf{R}_d\right) \\ &+ \ln\det\left(\mathbf{I}_{n_r} + q_1 q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r + q_2 \tilde{\mathbf{T}}_r\right) \\ &- \ln\det\left(\mathbf{I}_{n_r} + \alpha r_8 \tilde{\mathbf{T}}_r\right) - \ln\det\left(\mathbf{I}_{n_r} + q_8 \mathbf{R}_d\right) \\ &-(n_s r_1 q_1 + n_r \tilde{r}_2 q_2 - n_r r_8 q_8)\Big] \triangleq \nu \cdot \xi_1. \end{aligned} \tag{20}$$

with

$$q_1 = \frac{1}{n_s}\mathrm{Tr}\left\{\rho \tilde{\mathbf{T}}_s \left[\mathbf{I}_{n_s} + \rho r_1 \tilde{\mathbf{T}}_s\right]^{-1}\right\}$$
$$q_2 = \frac{1}{n_r}\mathrm{Tr}\left\{\alpha \mathbf{R}_d \left[\mathbf{I}_{n_r} + \alpha \tilde{r}_2 \mathbf{R}_d\right]^{-1}\right\}$$
$$q_8 = \frac{1}{n_r}\mathrm{Tr}\left\{\alpha \tilde{\mathbf{T}}_r \left[\mathbf{I}_{n_r} + \alpha r_3 \tilde{\mathbf{T}}_r\right]^{-1}\right\}$$
$$r_1 = \frac{1}{n_s}\mathrm{Tr}\left\{q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r \times \left[\mathbf{I}_{n_r} + q_1 q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r + q_2 \tilde{\mathbf{T}}_r\right]^{-1}\right\}$$
$$\tilde{r}_2 = \frac{1}{n_r}\mathrm{Tr}\left\{(\tilde{\mathbf{T}}_r + q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r) \times \left[\mathbf{I}_{n_r} + q_1 q_2 \mathbf{R}_r \tilde{\mathbf{T}}_r + q_2 \tilde{\mathbf{T}}_r\right]^{-1}\right\}$$
$$r_8 = \frac{1}{n_r}\mathrm{Tr}\left\{\mathbf{R}_d \left[\mathbf{I}_{n_d} + q_8 \mathbf{R}_d\right]^{-1}\right\}.$$

We note that $\xi_1(\cdot)$ in (20) is the multivariate version of the function mentioned in Section IV-C. We see that $\xi_1(\cdot)$ is $\Theta(n)$. This quantity will turn out to correspond to the mean of I in the large n limit.

At this point, we make use of the variable transformations $\mathbf{R}_x \to \boldsymbol{\delta}\mathbf{R}_x$ and $\mathbf{Q}_x \to \boldsymbol{\delta}\mathbf{Q}_x$ for $x = 1 \ldots 8$, which preserve the integral measures. Furthermore, we define

$$\mathbf{x}_{ab}^{(1)} \triangleq [\delta R_{1,ab}, \delta Q_{1,ab}, \delta R_{2,ab}, \delta Q_{2,ab}, \delta R_{3,ab}, \delta Q_{3,ab}, \delta R_{5,ab}, \delta Q_{5,ab}, \delta R_{6,ab}, \delta Q_{6,ab}]^T$$
$$\mathbf{x}_{ab}^{(2)} \triangleq [\delta R_{8,ab}, \delta Q_{8,ab}]^T$$
$$\mathbf{x}_{ab}^{(3)} \triangleq [\delta \overline{R}_{4,ab}, \delta R_{4,ab}, \delta Q_{4,ab}, \delta \overline{Q}_{4,ab}, \delta R_{7,ab}, \delta \overline{R}_{7,ab}, \delta Q_{7,ab}, \delta \overline{Q}_{7,ab}]^T.$$

With this notation we can write the moment generating function in terms of the Hessians of (17), $\mathbf{V}_1$, $\mathbf{V}_2$ and $\mathbf{V}_3$, as defined in (5), as

$$\begin{aligned} g_I(\nu) &= e^{-\mathcal{S}_0} \int \exp(-\mathcal{S}_2 - \mathcal{S}_3 - \mathcal{S}_4 - \ldots) \cdot d\lambda \\ &= e^{-\mathcal{S}_0} \int \exp(-\mathcal{S}_2)\Big\{1 - [\mathcal{S}_3 + \mathcal{S}_4 + \ldots] \\ &\qquad + \frac{1}{2}[\mathcal{S}_3 + \mathcal{S}_4 + \ldots]^2 - \ldots\Big\} \cdot d\lambda \qquad (21) \\ &= e^{-\mathcal{S}_0} \int \exp\left(-\frac{1}{2}\sum_{i=1}^{3}\sum_{a,b=1}^{\nu} \mathbf{x}_{ab}^{(i)T} \mathbf{V}_i \mathbf{x}_{ab}^{(i)}\right) \\ &\qquad \times \Big\{1 - [\mathcal{S}_3 + \mathcal{S}_4 + \ldots] \\ &\qquad + \frac{1}{2}[\mathcal{S}_3 + \mathcal{S}_4 + \ldots]^2 - \ldots\Big\} \cdot d\lambda \qquad (22) \\ &= e^{-\mathcal{S}_0}\Bigg[\left|\frac{\det \mathbf{V}_1 \det \mathbf{V}_2}{(\det \mathbf{V}_3)^2}\right|^{-\frac{\nu^2}{2}} \\ &\qquad + \int \exp\left(-\frac{1}{2}\sum_{i=1}^{3}\sum_{a,b=1}^{\nu} \mathbf{x}_{ab}^{(i)T} \mathbf{V}_i \mathbf{x}_{ab}^{(i)}\right) \\ &\qquad \times \Big\{-[\mathcal{S}_3 + \mathcal{S}_4 + \ldots] \\ &\qquad + \frac{1}{2}[\mathcal{S}_3 + \mathcal{S}_4 + \ldots]^2 - \ldots\Big\} \cdot d\lambda\Bigg]. \qquad (23) \end{aligned}$$

In (21) we expanded $\exp(-\mathcal{S}_3 - \mathcal{S}_4 - \cdots)$ into a series. The evaluation of the integral over the first term in (22) is provided in [18]. We note that $\xi_2(\cdot) \triangleq -\ln|\det \mathbf{V}_1| - \ln|\det \mathbf{V}_2| + 2\ln|\det \mathbf{V}_3|$, which will turn out to correspond to the variance of I in the large n limit, is $\mathcal{O}(1)$. Again, $\xi_2(\cdot)$ is the multivariate version of the function mentioned in Section IV-C.

C. Proving Gaussianity

We next show, that the remaining integral expression

$$\int \exp\left(-\frac{1}{2}\sum_{i=1}^{3}\sum_{a,b=1}^{\nu} \mathbf{x}_{ab}^{(i)T} \mathbf{V}_i \mathbf{x}_{ab}^{(i)}\right) \times \left\{-[\mathcal{S}_3 + \mathcal{S}_4 + \ldots + \frac{1}{2}[\mathcal{S}_3 + \mathcal{S}_4 + \ldots]^2 - \ldots\right\} \cdot d\lambda \tag{24}$$

is $\mathcal{O}(n^{-1})$. To see this, we need to consider the various Taylor coefficients of the $\mathcal{S}_p$ for $p > 2$ first. By inspecting (17), we note that

1a) a differentiation for either $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}_3$, $\mathbf{R}_5$, $\mathbf{R}_6$ or $\mathbf{R}_8$ yields a multiplication by a term that is $\mathcal{O}(n^{-1})$,

2a) a differentiation for either $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{Q}_3$, $\mathbf{Q}_5$, $\mathbf{Q}_6$ or $\mathbf{Q}_8$ does not change the order with respect to n,

3a) *two* differentiations for Grassmann variables (note that odd numbers of differentiations yield zero Taylor coefficients) yield a multiplication by a term that is $\mathcal{O}(n^{-1})$.

Therefore, a Taylor coefficient resulting from i, j and k differentiations of the first, second and third type, respectively, will be $\mathcal{O}(n^{1-i-k/2})$. Moreover, a product of t Taylor coefficients resulting from $i_1,\, j_1,\, k_1,\, i_2,\, j_2,\, k_2, \ldots, i_t,\, j_t,\, k_t$ differentiations of the first, second and third type, each, will be $\mathcal{O}(n^{t-\sum_l (i_l+k_l/2)})$.

Next, consider integrals of the form

$$\int \exp\left(-\frac{1}{2}\sum_{i=1}^{3}\sum_{a,b=1}^{\nu}\mathbf{x}_{ab}^{(i)T}\mathbf{V}_i\mathbf{x}_{ab}^{(i)}\right)\prod_{i,i\neq 4,7}\boldsymbol{\delta}\mathbf{R}_i \times \prod_{j,j\neq 4,7}\boldsymbol{\delta}\mathbf{Q}_j \prod_{k_1,k_2,k_3,k_4=4,7}\boldsymbol{\delta}\overline{\mathbf{R}}_{k_1}\boldsymbol{\delta}\mathbf{Q}_{k_2}\boldsymbol{\delta}\overline{\mathbf{Q}}_{k_3}\boldsymbol{\delta}\mathbf{R}_{k_4}d\lambda.$$

For the complex matrices Wick's theorem allows us to split the integral into sums of products of integrals involving only quadratic correlations. Furthermore, it states that for odd numbers of multipliers the integral evaluates to zero. Ignoring the Grassmann matrices for the moment we can extract the order of these correlations in the following. We define $\mathbf{V}$ as the joint Hessian

$$\mathbf{V} \triangleq \begin{pmatrix}\mathbf{V}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{V}_2\end{pmatrix}$$

and note that $\det(\mathbf{V})$ is $\mathcal{O}(1)$. Also, we define $\mathbf{x} \triangleq [\mathbf{x}_{ab}^{(1),T}, \mathbf{x}_{ab}^{(2),T}]^T$ and denote the integral measure $d\lambda$ without all Grassmann contributions by $d\tilde{\lambda}$. With this notation, we can extract the orders of the three kinds of arising quadratic correlations by applying the second part of Wick's theorem:

1b)

$$\exp\left(-\frac{1}{2}\sum_{a,b=1}^{\nu}\mathbf{x}^T\mathbf{V}\mathbf{x}\right)\cdot\delta R_{i,ab}\cdot\delta R_{j,cd}\cdot d\tilde{\lambda} = \delta_{ad}\delta_{bc}|\det(\mathbf{V})|^{-\frac{\nu^2}{2}}\cdot\frac{\det(\mathbf{V}^{(2i-1,2j-1)})}{\det(\mathbf{V})} = \Theta(n),$$

2b)

$$\int \exp\left(-\frac{1}{2}\sum_{a,b=1}^{\nu}\mathbf{x}^T\mathbf{V}\mathbf{x}\right)\cdot\delta Q_{i,ab}\cdot\delta Q_{j,cd}\cdot d\tilde{\lambda} = \delta_{ad}\delta_{bc}|\det(\mathbf{V})|^{-\frac{\nu^2}{2}}\cdot\frac{\det(\mathbf{V}^{(2i,2j)})}{\det(\mathbf{V})} = \mathcal{O}(n^{-1}),$$

3b)

$$\int \exp\left(-\frac{1}{2}\sum_{a,b=1}^{\nu}\mathbf{x}^T\mathbf{V}\mathbf{x}\right)\cdot\delta R_{i,ab}\cdot\delta Q_{j,cd}\cdot d\tilde{\lambda} = -\delta_{ad}\delta_{bc}|\det(\mathbf{V})|^{-\frac{\nu^2}{2}}\cdot\frac{\det(\mathbf{V}^{(2i-1,j)})}{\det(\mathbf{V})} = \mathcal{O}(1).$$

By $\det(\mathbf{V}^{(a,b)})$ we denote the sub-determinant when the ath row and the bth column in the matrix is deleted, δ_{xy} denotes the Kronecker delta function. The orders follow, since deleting odd lines/columns in $\mathbf{V}$ amounts to a multiplication of the respective determinant by a multiplier which is $\Theta(n)$, while deleting even lines/columns in $\mathbf{V}$ amounts to a multiplication of the respective determinant by a multiplier which is $\mathcal{O}(n^{-1})$. The Grassmannian integrations are easily verified to yield $\mathcal{O}(n^0)$ multipliers, since also the elements of $\mathbf{V}_3$ are $\mathcal{O}(n^0)$.

Combining 1a) and 1b), 2a) and 2b) as well as 3a) and 3b), we can finally summarize, that terms resulting from the evaluation of (24) are

$$\mathcal{O}\left(n^{t-\sum_{x=1}^{t}\frac{i_x+j_x+k_x}{2}}\right), \text{if } \sum_{x=1}^{t} i_x+j_x+k_x \text{ is even}$$

or zero otherwise. Here, t denotes the number of involved Taylor coefficients, i_x, j_x, k_x the number of derivatives of kind 1, 2, and 3. Note that i_x, j_x and k_x also correspond to the number of multipliers arising with the Taylor coefficient in the correlation. Since $\sum_{x=1}^{t}\frac{i_x+j_x+k_x}{2} > t$ for $p > 2$, we conclude that all appearing terms in the integral are $\mathcal{O}(n^{-1})$ or smaller.

We can thus rewrite (23) as

$$g_I(\nu) = e^{-\mathcal{S}_0}\cdot\left[\left|\frac{\det\mathbf{V}_1\det\mathbf{V}_2}{(\det\mathbf{V}_3)^2}\right|^{-\frac{\nu^2}{2}} + \mathcal{O}(n^{-1})\right].$$

After multiplying out the determinant, the cumulant generating function is given by

$$\begin{aligned}\ln g_I(\nu) &= \ln\left\{e^{-\mathcal{S}_0}\left|\frac{\det\mathbf{V}_1\det\mathbf{V}_2}{(\det\mathbf{V}_3)^2}\right|^{-\frac{\nu^2}{2}} \times\left(1+\left|\frac{\det\mathbf{V}_1\det\mathbf{V}_2}{(\det\mathbf{V}_3)^2}\right|^{\frac{\nu^2}{2}}\cdot\mathcal{O}(n^{-1})\right)\right\}\\ &= -\nu\cdot\xi_1 - \frac{\nu^2}{2}(\ln|\det(\mathbf{V}_1)| + \ln|\det(\mathbf{V}_2)| - 2\ln|\det(\mathbf{V}_3)|) + \ln(1+\mathcal{O}(n^{-1}))\\ &= -\nu\cdot\xi_1 + \frac{\nu^2}{2}\cdot\xi_2 + \mathcal{O}(n^{-1}).\end{aligned}$$

A comparison of coefficients with (7) immediately reveals

$$\mathsf{E}[I] = \xi_1 + \mathcal{O}(n^{-1})$$

and

$$\mathsf{Var}[I] = \xi_2 + \mathcal{O}(n^{-1}).$$

Also, the $\mathcal{C}_p$ for $p > 2$ are $\mathcal{O}(n^{-1})$ and thus vanish for $n \to \infty$. This implies that I is Gaussian distributed in this limit [29, Th. 1]. Note that indeed the residual term of the variance can be shown to be $\mathcal{O}(n^{-2})$ in the same way as it is done in [18]. The reason behind this is that no $\mathcal{O}(n^{-1})$ term proportional to ν^2 is generated in (23). We skip this (in the present case very tedious) derivation for reasons of brevity.

VI. Comparison With Simulation Results

We verify the results stated in the theorem by means of computer experiments. For the mean, this is done through Monte Carlo simulations. The respective plots are shown in Figs. 2 and 3. In Fig. 2, we present the ergodic mutual information versus the SNR for $n = n_s = n_r = n_d = 2, 4$ and 8. We observe that even for only two antennas the approximation is reasonable, for four antennas the match is close to perfect, while for eight antennas no difference between analytic approximation and numeric evaluation can be seen anymore. In Fig. 3, we test a configuration with different antenna array sizes at source, relay and destination and correlated fading. Defining the Toeplitz matrix

$$\boldsymbol{\Omega}(\gamma,n) \triangleq \begin{pmatrix} 1 & \gamma & \gamma^2 & \cdots & \gamma^n\\ \gamma & 1 & \gamma & \ddots & \\ \gamma^2 & \gamma & 1 & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \gamma\\ \gamma^n & & & \gamma & 1\end{pmatrix}$$

we choose $\tilde{\mathbf{T}}_s = \boldsymbol{\Omega}(1/2, n_s)$, $\mathbf{R}_r = \boldsymbol{\Omega}(3/8, 3n_s)$, $\tilde{\mathbf{T}}_r = \boldsymbol{\Omega}(1/4, 3n_s)$ and $\mathbf{R}_d = \boldsymbol{\Omega}(1/8, 2n_s)$ and obtain very good matches between analytic solutions and simulation results. The reason why

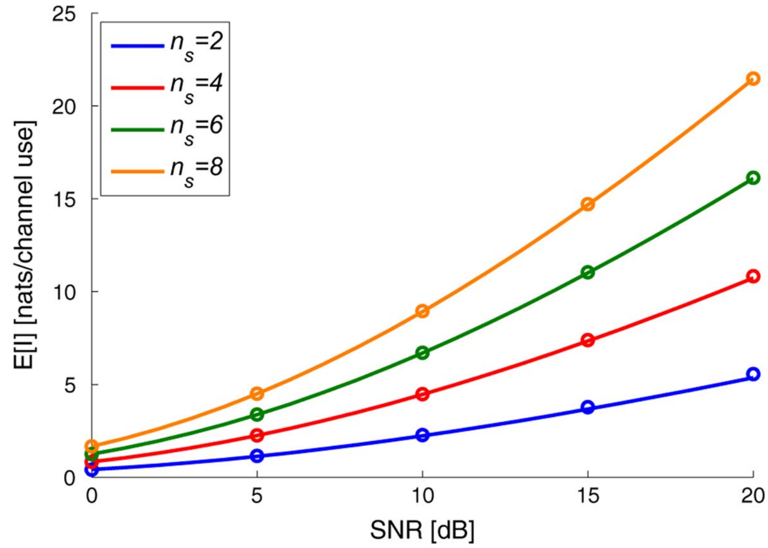

Fig. 2. Mean mutual information versus SNR for $n_s = n_r = n_d$ and i.i.d. channel matrix entries—solid lines are analytical approximations, circles mark true mutual information as obtained through Monte Carlo simulations.

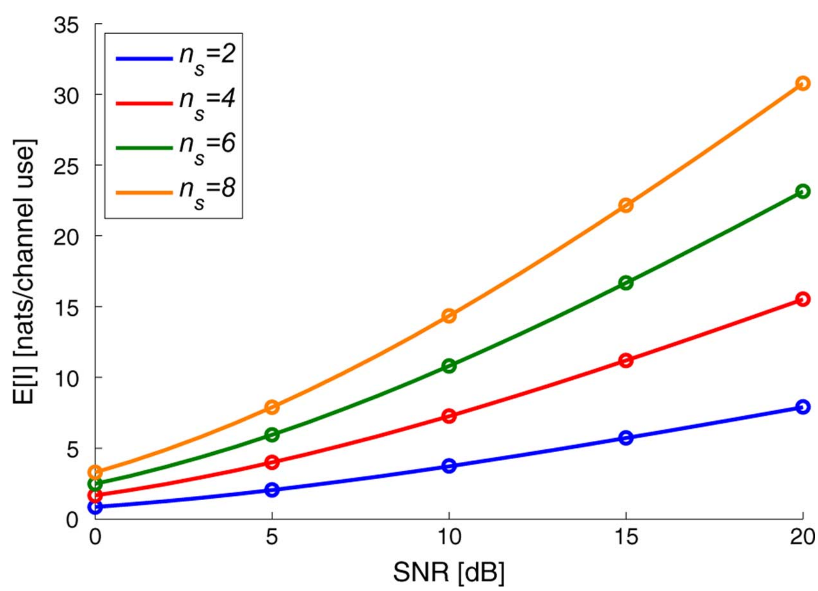

Fig. 3. Mean mutual information versus SNR for $n_r = 3n_s$,$n_d = 2n_s$ and correlation matrices as given in Section VI—solid lines are analytical approximations, circles mark true mutual information as obtained through Monte Carlo simulations.

the results look even tighter in this case lies in the larger number of antennas involved in total.

In order to also verify our results for the higher cumulant moments we compare the empirical cumulative distribution function (CDF) of the mutual information to a Gaussian CDF with mean and variance given in the theorem. The respective plots are shown in Figs. 4 and 5. We observe that the analytic approximation becomes tight indeed as $n = n_s = n_r = n_d$ increases (see Fig. 4). For $n = 8$ even the tails of the distribution are reasonably approximated, which is an important issue for the characterization of the outage capacity. Again, we also present the curves for $\tilde{\mathbf{T}}_s = \mathbf{\Omega}(1/2, n_s)$, $\mathbf{R}_r = \mathbf{\Omega}(3/8, 3n_s)$, $\tilde{\mathbf{T}}_r = \mathbf{\Omega}(1/4, 3n_s)$ and $\mathbf{R}_d = \mathbf{\Omega}(1/8, 2n_s)$ in Fig. 5. Our simulation results thus also demonstrate that the replica method—despite its deficiency of not being mathematically rigorous yet—indeed reveals the correct solution to our problem.

VII. Conclusion

Having used the framework developed in [21] and [18], we have evaluated the cumulant moments of the mutual information for amplify-and-forward MIMO relay channels with Gaussian input in the asymptotic regime of large antenna numbers. Similarly to the case of ordinary point-to-point MIMO channels, we observe that all cumulant moments higher than the variance vanish as the antenna array sizes grow large and conclude that the respective mutual information is Gaussian distributed. For mean and variance, we

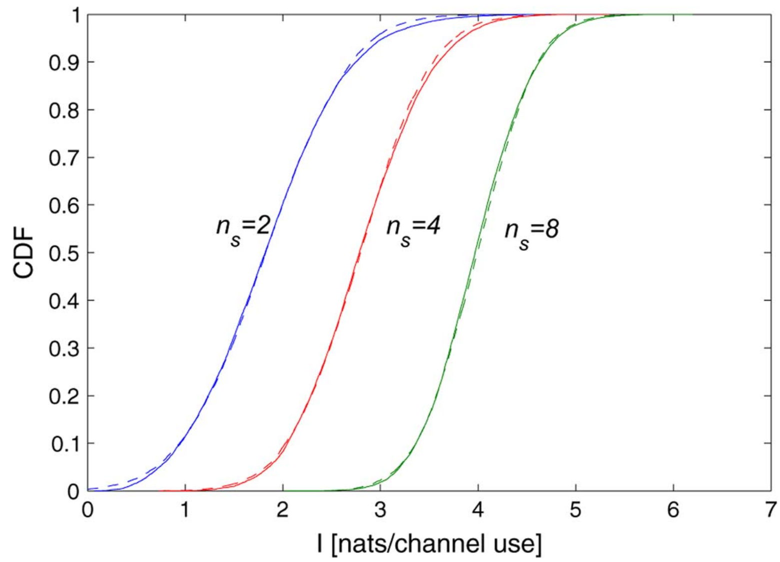

Fig. 4. Cumulative distribution function of mutual information for $n = n_s = n_r = n_d$ and i.i.d. channel matrix entries. Dashed lines represent Gaussian distributions with analytically computed mean and variance. The solid lines are the empirical distributions obtained through simulations.

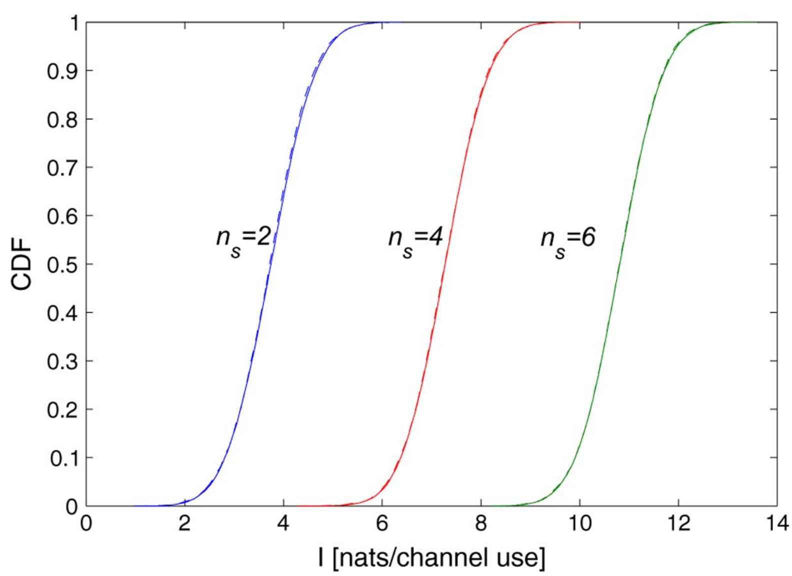

Fig. 5. Cumulative distribution function of mutual information for $n_r = 3n_s, n_d = 2n_d$ and correlation matrices as given in Section VI. Dashed lines represent Gaussian distributions with analytically computed mean and variance. The solid lines are the empirical distributions obtained through simulations.

obtain expressions that allow for an analytic evaluation. Computer experiments show, that the derived expressions serve as excellent approximations even for channels with only very few antennas. The results confirm the linear scaling of the ergodic mutual information in the antenna array size and reveal that the respective variance is upper bounded by a constant.

Appendix A
Preliminaries of Grassmann Variables

Grassmann algebra is a concept from mathematical physics. A Grassmann variable (also called an anticommuting number) is a quantity that anticommutes with other Grassmann numbers but commutes with (ordinary) complex numbers. With θ_1, θ_2 Grassmann variables and λ a complex number the defining properties are

$$\lambda\theta_1 = \theta_1\lambda$$
$$\theta_1\theta_2 = -\theta_2\theta_1.$$

With θ_3 another Grassmann variable further properties are

$$\theta_1(\theta_2\theta_3) = \theta_3(\theta_1\theta_2)$$
$$\theta_1^2 = 0$$
$$\exp(\theta_1\theta_2) = 1 + \theta_1\theta_2.$$

Integration over Grassmann variables is defined by the following two properties

$$\int d\theta = 0$$
$$\int \theta d\theta = 1.$$

Note that also the differentials are anticommuting, i.e., $d\theta_1 d\theta_2 = -d\theta_2 d\theta_1$. Further details about integrals over Grassmann variables such as variable transformation can be found in the Appendix of [18].

APPENDIX B
WICK'S THEOREM

With $\mathbf{V} \in \mathbb{C}^{N \times N}$, $\mathbf{x} \in \mathbb{C}^{N \times 1}$ and an integral measure $d\alpha(\mathbf{x}) = 1/\sqrt{2\pi} dx_1, \ldots, dx_N$ we have

$$\begin{aligned}(\det \mathbf{V})^{\frac{1}{2}} \int \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{V}\mathbf{x}\right) \cdot \prod_{k=1}^{M} x_k \cdot d\alpha(\mathbf{x}) \\ = \sum_{\text{pairs}} (\det \mathbf{V})^{\frac{1}{2}} \int \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{V}\mathbf{x}\right) \cdot x_{i,1}x_{i,2} \cdot d\alpha(\mathbf{x}) \\ \times \cdots \\ \times (\det \mathbf{V})^{\frac{1}{2}} \int \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{V}\mathbf{x}\right) \cdot x_{i_{M-1}} x_{i_M} \cdot d\alpha(\mathbf{x}) \quad (25)\end{aligned}$$

if M is even. For odd M the expression evaluates to zero. The sum in (25) is over all possible rearrangements of the orderings of the indices such that different indexes are paired with each other (with each distinct pairing being counted once).

Furthermore, we have that

$$(\det \mathbf{V})^{\frac{1}{2}} \int \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{V}\mathbf{x}\right) x_i x_j \cdot d\alpha(\mathbf{x}) = [V^{-1}]_{i,j}$$

with $[\mathbf{V}^{-1}]_{i,j}$ the element in the ith row and jth column of $\mathbf{V}^{-1}$. We will also need that

$$[\mathbf{V}^{-1}]_{i,j} = \det \mathbf{V}^{(i,j)} / \det \mathbf{V}$$

with $\mathbf{V}^{(i,j)}$ an $N-1 \times N-1$ matrix, where the ith row and the jth column of $\mathbf{V}^{-1}$ are deleted.

ACKNOWLEDGMENT

The authors would like to thank Aris Moustakas for various very valuable advices.

REFERENCES

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

[2] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I & Part II," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1948, Nov. 2003.

[4] A. Wittneben and B. Rankov, "Impact of cooperative relays on the capacity of rank-deficient MIMO channels," in *Proc. Mobile and Wireless Communications Summit (IST)*, Aveiro, Portugal, Jun. 2003.

[5] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[6] M. Gastpar, G. Kramer, and M. Vetterli, "The multiple-relay channel: Coding and antenna-clustering capacity," in *Proc. IEEE Int. Symp. Inf. Theory*, Lausanne, Switzerland, Jun. 2002.

[7] A. Høst-Madsen, "On the capacity of wireless relaying," in *Proc. 56th IEEE Vehicular Technology Conf.*, Seoul, Korea, Sep. 2002, vol. 3, pp. 1333–1337.

[8] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[9] M. Khojastepour, B. Aazhang, and A. Sabharwal, "On the capacity of 'cheap' relay networks," in *Proc. Conf. Information Science and Systems (CISS)*, Princeton, NJ, Apr. 2003.

[10] V. I. Morgenshtern and H. Bölcskei, "Crystallization in large wireless networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3319–3349, Oct. 2007.

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

[12] P. B. Rapajic and D. Popescu, "Information capacity of a random signature multiple-input multiple-output channel," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1245–1248, Sep.. 2000.

[13] X. Mestre, J. R. Fonollosa, and A. Pags-Zamora, "Capacity of MIMO channels: Asymptotic evaluation under correlated fading," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 637–650, Jun. 2003.

[14] M. Debbah and R. R. Müller, "MIMO channel modeling and the principle of maximum entropy," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1667–1690, May 2005.

[15] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple receive antenna architectures," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.

[16] B. Hochwald, T. L. Marzetta, and T. V. , "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[17] M. A. Kamath and B. L. Hughes, "The asymptotic capacity of multiple-antenna Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4325–4333, Dec. 2005.

[18] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large n analysis," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, p. 2545, Oct. 2003.

[19] S. F. Edwards and P. W. Anderson, "Theory of spin glasses," *J. Phys. F: Metal Phys.*, vol. 5, no. 52, pp. 965–974, May 1975.

[20] 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.

[21] A. M. Sengupta and P. P. Mitra, Capacity of multivariate channels with multiplicative noise: I. Random matrix techniques and large-n expansions for full transfer matrices Oct. 2000, LANL arXiv:Physics/0010081.

[22] R. Müller, "A random matrix model of communication via antenna arrays," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2495–2506, Sep. 2002.

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

[24] A. J. Paulraj, R. U. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, May 2003.

[25] H. Orcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of 'Kronecker' MIMO radio channel model," *Electron. Lett.*, vol. 39, no. 16, pp. 1209–1210, Aug. 2003.

[26] W. Weichselberger, M. Herdin, H. Orcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[27] Z. Bai and J. Silverstein, "CLT for linear spectral statistics of large-dimensional sample covariance matrices," *Ann. Probab.*, vol. 32, no. 1A, pp. 553–605, 2004.

[28] A. M. Sengupta and P. P. Mitra, "Distributions of singular values for some random matrices," *Phys. Rev. E*, vol. 60, no. 3, pp. 3389–3392, 1999.

[29] S. Janson, "Normal convergence by higher semiinvariants with applications to sums of dependent random variables and random graphs," *Ann. Probab.*, vol. 16, no. 1, pp. 305–312, 1988.