Asymptotic Statistics of the Mutual Information for Spatially Correlated Rician Fading MIMO Channels With Interference

Erwin Riegler, *Member, IEEE*, and Giorgio Taricco, *Fellow, IEEE*

***Abstract*—The statistics of the mutual information of a separately correlated Rician fading multiple-input multiple-output (MIMO) channel in the presence of multiple-access interference are addressed in this paper. The approach followed is asymptotic in the number of transmit, receive, and interfering antennas, which are all assumed to grow asymptotically large while approaching finite ratios. In this asymptotic regime: i) the mean and the variance of the mutual information are calculated; and ii) the mutual information distribution is shown to converge to the Gaussian distribution, specified completely by the mean and variance, under some mild technical conditions. The asymptotic method adopted relies on two powerful tools developed in the context of theoretical physics: the *replica method* and *superanalysis*. The former has been already successfully applied in several research studies on MIMO systems. The application of these asymptotic results takes advantage of the fact that, in spite of being developed under the assumption of an asymptotically large number of antennas, they still represent a very accurate approximation even when the number of antennas is limited to a few units, as supported by the ample set of numerical results obtained by Monte Carlo simulations and reported in this paper.**

***Index Terms*—Multiple-input multiple-output (MIMO) system capacity, correlated Rician fading, replica method, superanalysis, interference channels.**

I. Introduction

MULTIPLE-INPUT multiple-output (MIMO) communication systems have drawn considerable attention for more than one decade because of the promise of increasing the capacity by a factor which is the minimum between the number of transmit antennas and the number of receive antennas employed at the link ends [16], [47], [52]. The promise was based on the early assumption that the channel gains were iid and Rayleigh distributed, which is the common ground of most early literature results.

This assumption, often referred to as *rich scattering assumption*, relied on the presence of a very large number of independent propagation paths from the transmitter to the receiver. The absence of line-of-sight (LOS) propagation implies that the path gains are Rayleigh-distributed and independence of the entries of the channel matrix refers to the spatial or antenna domain.

Subsequent studies showed that the rich scattering assumption was not always realistic and spatial correlation came often into play [20]. It has been noticed that antenna (or spatial) correlation changes drastically with the scattering environment, the distance between the transmitter and the receiver, the antenna configurations, and the Doppler spread [1], [39]. Although the growth rate of the asymptotic (in the number of antennas) capacity is affected by the presence of spatial correlation, it has been shown that capacity still scales linearly in the number of antennas under the assumption that the empirical eigenvalue distributions of the spatial correlation matrices converge in distribution to some limiting distributions as the number of transmit antennas tends to infinity [9]. This result holds in the case of perfect channel state information at the receiver (CSIR) and either perfect channel state information at the transmitter (CSIT) or no CSIT. In the absence of CSIT, the growth rate in the correlated case was shown to be smaller compared to the iid case. Oppositely, the authors showed that the situation changes in the low SNR regime where correlation helps to increase capacity in the case of full CSIT.

By using the replica method, Moustakas *et al.* [29] derived the mean and variance of the asymptotic capacity in the case of perfect CSIR and perfect channel distribution information at the transmitter (CDIT) with and without multiple-access interference and show that all higher moments tend to zero in the asymptotic limit. Five years later, the results from [29] have been rediscovered without using the replica method by Hachem *et al.* in [23] (these results are based on the Poincaré-Nash inequality and the application of Stieltjes-transform methods developed by Girko in [18]). This reflects the difficulty of the problem even for the central case, where the channel matrix doesn't have any LOS component, not to mention the case where multiple-access interference is present.

However, LOS propagation, which leads to consider a spatially correlated *Rician fading* model, has an important impact

Manuscript received November 17, 2008; revised December 11, 2009. Current version published March 17, 2010. The work of E. Riegler was supported by the STREP project No. IST-026905 (MASCOT) within the sixth Framework Programme of the European Commission. FTW Telecommunications Research Center Vienna is supported by the Austrian Government and by the City of Vienna within the competence center program COMET. The work of G. Taricco was supported in part by the STREP project No. IST-026905 (MASCOT) within the sixth Framework Programme of the European Commission. The material in this paper was presented in part at the IEEE GLOBECOM 2007, Communication Theory Symposium, Washington, DC, November 2007.

E. Riegler is with the Faculty of Electrical Engineering and Information Technology, Vienna University of Technology, A-1040 Vienna, Austria and also with the and the FTW Telecommunications Research Center Vienna, A-1220 Vienna, Austria (e-mail: erwin.riegler@nt.tuwien.ac.at).

G. Taricco is with the Dipartimento di Elettronica, Politecnico di Torino, 10129, Torino, Italy (e-mail: taricco@polito.it).

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

Color versions of Figures 1–14 in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2010.2040865

on the MIMO channel capacity. Many experimental and theoretical studies have shown that, in order to encompass correctly all the channel characteristics, correlated Rician fading models have to be considered [17], [19], [37]. This is due to the fact that the statistical model depends on the time scale of interest. For example, in the short term, the channel coefficients may have a nonzero mean. If the transmitter receives frequent updates of the channel statistics it can adapt to these time-varying updates of short-term channel statistics and capacity is increased relative to the transmission strategy associated with just the long-term channel statistics [19].

Nevertheless, until recently, the Rician fading channel model received more limited attention in the literature, as far as concerns its capacity, with some notable exceptions, which are reported in the following. The effect of imperfect feedback for the multiple-input single-output (MISO) case was considered by Visotsky and Madhow in [50]. The optimum input covariance in the MIMO case was found by Venkatesan *et al.* in [49] under the assumption of unit-rank average channel matrix. Later, Jafar and Goldsmith obtained in [26] necessary and sufficient conditions for the optimality of beamforming in the full CSIR case, assuming that the transmitter only knows either the average channel matrix (with arbitrary rank) or the channel covariance matrix. More recently, Hösli *et al.* [24] studied the dependence of channel capacity on the average channel matrix and showed that the mutual information increases monotonically (for a fixed covariance matrix, not a fixed SNR) with the Lowner partial order of the singular values of the average channel matrix. Assuming Rician fading and spatial correlation at the receiver only, Jayaweera and Poor [27] obtained the exact value of the capacity in the case of a transmitter which has no knowledge about the channel, and bounds on the capacity when the transmitter knows the Rician factor. Tulino *et al.* [48] derived the eigenvectors of the ergodic capacity achieving covariance matrix for several MIMO channel models of interests and suggested a numerical algorithm to obtain the eigenvalues. The asymptotic capacity for the Rician channel with perfect CSIR and CDIT has been obtained recently in [30]. Combining Rician fading and the Kronecker model, Taricco [43] derived the mean and variance of the asymptotic capacity in the case of perfect CSIR and CDIT and proved the asymptotic *Gaussianity* of the mutual information. Similar results for the ergodic capacity have been obtained in [12], [13], and [22] using the Stieltjes transform.

With the exception of [29], all previous results focused only on the single user case. A few papers considered the presence of multiple-access interference, corresponding to a more realistic multiuser MIMO scenario [3], [4], [6]–[8], [29]. Recently, Chiani *et al.* found a closed-form expression of the exact mean capacity for an uncorrelated Rayleigh fading MIMO channel with interference [7]. Notice that assimilating interference to additive noise *leads to incorrect capacity results*, as properly evidenced in [7]. The asymptotic (in the number of antennas) ergodic capacity with interference and ergodic sum-rate capacity of the separately correlated Rician fading MIMO channel have been obtained recently by the authors of these paper in [44], [45], and [35].

It is worth mentioning that the calculation of the *average* mutual information, in the presence of interference, can be reduced to the single user (interference-free) case using the linearity of the mean value. However, for *higher moments*, this is no longer the case. In this contribution we complement the results of [29] and [43] by extending the asymptotic analysis to the case of separately correlated Rician fading MIMO channels with multiple-access interference. We also show that, under mild technical conditions, all moments higher than the second tend to zero asymptotically (under the assumption that all number of antennas grow to infinity with their ratios approaching finite values). These results are applicable to the analysis of outage probability, where one is interested in finding a tradeoff between high average mutual information and low standard deviation (see [43, Remark IV.2]). While optimizing the asymptotic average mutual information in the presence of interference can be done by water-filling [45], a systematic way to optimize the asymptotic outage mutual information is still an open research question, even in the single user case.

The remainder of this paper is organized as follows. In Section II we summarize the notations in this paper. Section III describes the model of the separately correlated Rician fading MIMO channel in the presence of co-channel interference. This section also contains the key parameters, the normalizations, and the definition of the asymptotic setting. Section IV defines the cumulant moment generating function of a random mutual information, whose asymptotic derivation is among the goals of this paper. Section V summarizes the key concepts of two tools from theoretical physics upon which we built our results: the replica method and superanalysis. The main result of this paper is reported in Section VI, while the details of the asymptotic approximation of the cumulant generating function are collected into the Appendices. Section VII presents several numerical examples aimed not only to assess the accuracy of the approximation proposed but also to analyze a selection of communication scenarios of interest. Finally, Section VIII summarizes the results presented in this paper and provides our concluding remarks.

II. Notation

We denote column-vectors and matrices by lowercase and uppercase boldface characters, respectively. Grassmann variables are denoted by Greek letters while for complex numbers we use Latin symbols. The transpose of a matrix $\boldsymbol{A}$ is $\boldsymbol{A}^T$. The Hermitian transpose of $\boldsymbol{A}$ is $\boldsymbol{A}^H$. The trace of $\boldsymbol{A}$ is $\mathrm{tr}(\boldsymbol{A}) = \sum_a (\boldsymbol{A})_{aa}$. The exponential of the trace of $\boldsymbol{A}$ is denoted by $\mathrm{etr}(\boldsymbol{A}) = \exp(\mathrm{tr}(\boldsymbol{A}))$. The Frobenius norm of a matrix $\boldsymbol{A}$ is $\|\boldsymbol{A}\| = \sqrt{\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^H)}$. The spectral norm of $\boldsymbol{A}$ is defined as $\|\boldsymbol{A}\|_{sp} = \max(\sqrt{\lambda}|\lambda \text{ is an eigenvalue of } \boldsymbol{A}\boldsymbol{A}^H)$. $\boldsymbol{A} \oplus \boldsymbol{B}$ is the 2×2 block diagonal matrix with upper matrix block $\boldsymbol{A}$ and lower matrix block $\boldsymbol{B}$. $\boldsymbol{A} \otimes \boldsymbol{B}$ denotes the usual Kronecker product of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. Among the properties of the Kronecker product we recall the following: i) $\boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C}) = (\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C}$; ii) $\prod_k (\boldsymbol{A}_k \otimes \boldsymbol{B}_k) = (\prod_k \boldsymbol{A}_k) \otimes (\prod_k \boldsymbol{B}_k)$; iii) $(\boldsymbol{A} \otimes \boldsymbol{B})^H = \boldsymbol{A}^H \otimes \boldsymbol{B}^H$; (iv) $(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1}$ if $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible; v) $\mathrm{tr}(\boldsymbol{A} \otimes \boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A})\mathrm{tr}(\boldsymbol{B})$; vi) $\det(\boldsymbol{A} \otimes \boldsymbol{B}) = \det(\boldsymbol{A})^n \det(\boldsymbol{B})^m$ if $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ and $\boldsymbol{B} \in \mathbb{C}^{n \times n}$; (Vii) $(\boldsymbol{B}^T \otimes \boldsymbol{A})\mathrm{vec}\, \boldsymbol{C} = \mathrm{vec}\, \boldsymbol{A}\boldsymbol{C}\boldsymbol{B}$ where $\mathrm{vec}\, \boldsymbol{C}$ is the column vector obtained by stacking the columns of $\boldsymbol{C}$ on top of each other from left to right. $\boldsymbol{A}^{1/2}$ is the unique positive semidefinite matrix square-root of a positive semidefinite matrix $\boldsymbol{A}$. The notation $X \sim \mathcal{N}(\mu, \sigma^2)$ means that X is a complex Gaussian random variable with mean value $\mu = \mathrm{E}[X]$ and variance $\sigma^2 = \mathrm{E}[|X|^2] - |\mu|^2$. Given two functions $f(x)$ and $g(x)$, $f(x) = \Theta(g(x))$ if

$$0 < \liminf_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| \le \limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| < \infty$$

and $f(x) = O(g(x))$ if

$$\limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| < \infty.$$

III. MIMO Channel Model

We consider a narrowband block fading channel with n_R receive antennas, n_T transmit antennas from a transmitting user, and n_I transmit antennas from an interfering source (possibly representing several different users) characterized by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{H}_I\boldsymbol{x}_I + \boldsymbol{z}.$$

Here, $\boldsymbol{y} \in \mathbb{C}^{n_R}$ is the received signal vector, $\boldsymbol{x} \in \mathbb{C}^{n_T}$ is the transmitted complex, Gaussian distributed signal vector with zero mean and covariance $\boldsymbol{Q} = \mathrm{E}[\boldsymbol{x}\boldsymbol{x}^H]$, $\boldsymbol{x}_I \in \mathbb{C}^{n_I}$ is the interfering complex, Gaussian distributed signal vector with zero mean and covariance $\boldsymbol{Q}_I = \mathrm{E}[\boldsymbol{x}_I\boldsymbol{x}_I^H]$, and $\boldsymbol{z} \in \mathbb{C}^{n_R}$ is additive zero-mean white noise with iid entries $\boldsymbol{z}_a \sim \mathcal{N}_c(0,1)$. The channel matrices $\boldsymbol{H} \in \mathbb{C}^{n_R\times n_T}$ and $\boldsymbol{H}_I \in \mathbb{C}^{n_R\times n_I}$ model separately (or Kronecker) correlated Rician fading. Thus, they can be written as

$$\begin{cases} \boldsymbol{H} = \bar{\boldsymbol{H}} + \boldsymbol{R}^{1/2}\boldsymbol{H}_w\boldsymbol{T}^{1/2} \\ \boldsymbol{H}_I = \bar{\boldsymbol{H}}_I + \boldsymbol{R}_I^{1/2}\boldsymbol{H}_{w,I}\boldsymbol{T}_I^{1/2} \end{cases} \tag{1}$$

where $\bar{\boldsymbol{H}}$ and $\bar{\boldsymbol{H}}_I$ represent the mean values of $\boldsymbol{H}$ and $\boldsymbol{H}_I$, respectively, and are related to the presence of LOS components, $(\boldsymbol{H}_w)_{ab}, (\boldsymbol{H}_{w,I})_{ab} \sim \mathcal{N}_c(\boldsymbol{0},\boldsymbol{1})$, and the positive semidefinite matrices $\boldsymbol{T}(\boldsymbol{T}_I)$ and $\boldsymbol{R}(\boldsymbol{R}_I)$ are the transmit signal (interference) and receive signal (interference) spatial correlation matrices, respectively. The covariance between the different entries of $\boldsymbol{H}$ and $\boldsymbol{H}_I$ satisfies the identities

$$\begin{cases} \mathrm{cov}((\boldsymbol{H})_{ij}(\boldsymbol{H})_{i'j'}) = (\boldsymbol{R})_{ii'}(\boldsymbol{T})^*_{jj'} \\ \mathrm{cov}((\boldsymbol{H}_I)_{ij}(\boldsymbol{H}_I)_{i'j'}) = (\boldsymbol{R}_I)_{ii'}(\boldsymbol{T}_I)^*_{jj'}. \end{cases}$$

Remark III.1: (Extension to multiple interfering transmitter) (3) can be applied to model multiple interfering transmitter under the assumption of a *common receive correlation matrix for each interfering source*. Indeed, assume that we have N_I interfering users, so that the channel of interfering user i $(i = 1,\ldots,N_I)$ is of the form $\boldsymbol{H}_I^i = \bar{\boldsymbol{H}}^i + \boldsymbol{R}^{1/2}\boldsymbol{W}^i{\boldsymbol{T}_I^i}^{1/2}$ and $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \sum_{i=1}^{N_I}\boldsymbol{H}_I^i\boldsymbol{x}_I^i + \boldsymbol{z}$.

Then set $\boldsymbol{x} \triangleq [{\boldsymbol{x}_I^1}^T,\ldots,{\boldsymbol{x}_I^{N_I}}^T]^T$, $\boldsymbol{T} \triangleq \bigoplus_{i=1}^{N_I}\boldsymbol{T}_I^i$, $\boldsymbol{W} \triangleq [\boldsymbol{W}^1,\ldots,\boldsymbol{W}^{N_I}]$, and $\bar{\boldsymbol{H}} \triangleq [\bar{\boldsymbol{H}}^1,\ldots,\bar{\boldsymbol{H}}^{N_I}]$. Different receive correlation matrices for the interfering transmitter can be modeled be introducing virtual delays in combination with a wideband channel model, as proposed in [36] and [46].

A. Normalizations

Following the approach in [43], we define

$$\begin{cases} \tilde{\boldsymbol{H}} \triangleq \bar{\boldsymbol{H}}\boldsymbol{Q}^{1/2} & \tilde{\boldsymbol{T}} = \boldsymbol{T}^{1/2}\boldsymbol{Q}\boldsymbol{T}^{1/2} \\ \tilde{\boldsymbol{H}}_I \triangleq \bar{\boldsymbol{H}}_I\boldsymbol{Q}_I^{1/2} & \tilde{\boldsymbol{T}}_I = \boldsymbol{T}_I^{1/2}\boldsymbol{Q}_I\boldsymbol{T}_I^{1/2}. \end{cases} \tag{2}$$

Then, the transmitted signal and interference covariance matrices are implicitly accounted for into $\tilde{\boldsymbol{H}}$, $\tilde{\boldsymbol{H}}_I$, $\tilde{\boldsymbol{T}}$, and $\tilde{\boldsymbol{T}}_I$. According to these definitions, the total received power is given by

$$\mathrm{E}[\|\boldsymbol{y}\|^2] = \underbrace{\|\tilde{\boldsymbol{H}}\|^2 + \mathrm{tr}(\boldsymbol{R})\mathrm{tr}(\tilde{\boldsymbol{T}})}_{\text{signal}} + \underbrace{\|\tilde{\boldsymbol{H}}_I\|^2 + \mathrm{tr}(\boldsymbol{R}_I)\mathrm{tr}(\tilde{\boldsymbol{T}}_I)}_{\text{interference}} + n_R.$$

Splitting the total received power components into direct and diffuse parts we obtain the Rician factors

$$K = \frac{\|\tilde{\boldsymbol{H}}\|^2}{\mathrm{tr}(\boldsymbol{R})\mathrm{tr}(\tilde{\boldsymbol{T}})}\; K_I = \frac{\|\tilde{\boldsymbol{H}}_I\|^2}{\mathrm{tr}(\boldsymbol{R}_I)\mathrm{tr}(\tilde{\boldsymbol{T}}_I)}. \tag{3}$$

Then, we can define the signal-to-noise and the interference-to-noise power ratios as

$$\begin{cases} \mathsf{SNR} = \dfrac{(K+1)\mathrm{tr}(\boldsymbol{R})\mathrm{tr}(\tilde{\boldsymbol{T}})}{n_R} \\ \mathsf{INR} = \dfrac{(K_I+1)\mathrm{tr}(\boldsymbol{R}_I)\mathrm{tr}(\tilde{\boldsymbol{T}}_I)}{n_R} \end{cases} \tag{4}$$

with signal-to-interference power ratio $\mathsf{SIR} = \mathsf{SNR}/\mathsf{INR}$. It can be noticed that the previous definitions of the Rician factors and of the SNRs in (3) and (4) depend on the full transmit covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{Q}_I$, and not only on their traces, unless they are proportional to the identity matrix. This may be an issue when the channel capacity is evaluated against the SNR since $\boldsymbol{Q}$ *is not* proportional to the identity matrix if it is optimized to achieve the ergodic capacity (see [43, Remark II.1]).

In what follows we further assume that:

1) The spatial correlation matrices are normalized by

$$\mathrm{tr}(\boldsymbol{R}) = \mathrm{tr}(\boldsymbol{R}_I) = n_R$$
$$\mathrm{tr}(\boldsymbol{T}) = n_T$$
$$\mathrm{tr}(\boldsymbol{T}_I) = n_I.$$

2) The matrices defined in (2) are normalized according to the identities

$$\begin{cases} \|\tilde{\boldsymbol{H}}\|^2 = \dfrac{n_R K\mathsf{SNR}}{K+1} & \mathrm{tr}(\tilde{\boldsymbol{T}}) = \dfrac{\mathsf{SNR}}{K+1} \\ \|\tilde{\boldsymbol{H}}_I\|^2 = \dfrac{n_R K_I\mathsf{INR}}{K_I+1} & \mathrm{tr}(\tilde{\boldsymbol{T}}_I) = \dfrac{\mathsf{INR}}{K_I+1}. \end{cases} \tag{5}$$

B. Asymptotic Setup

We define the following asymptotic setting:

1) We assume that $n_T, n_R, n_I \to \infty$ while the ratios $n_T/n_R \to \kappa \in (0,\infty)$ and $n_I/n_R \to \kappa_I \in (0,\infty)$.
2) The Rician factors K, K_I and the signal-to-noise and interference-to-noise ratios SNR and INR are kept fixed as $n_T, n_R, n_I \to \infty$.

3) The spatial correlation matrices are bounded by

$$\|\boldsymbol{R}\|_{sp} = \Theta(1) \quad \|\boldsymbol{R}_I\|_{sp} = \Theta(1)$$
$$\|\tilde{\boldsymbol{T}}\|_{sp} = \Theta(1/n_R) \quad \|\tilde{\boldsymbol{T}}_I\|_{sp} = \Theta(1/n_R).$$

4) The LOS matrices are bounded by

$$\|\tilde{\boldsymbol{H}}\|_{sp} = O(1) \quad \|\tilde{\boldsymbol{H}}_I\|_{sp} = O(1).$$

IV. Mutual Information and Cumulant Generating Function

We assume that the receiver knows *exactly* the realizations of the channel matrices $\boldsymbol{H}$ and $\boldsymbol{H}_I$. The random mutual information is given by [7], [29]

$$\mathcal{I} \triangleq \ln\left(\frac{\det(\boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H)}{\det(\boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H)}\right) \tag{6}$$

nat/complex dimension. Here, the channel matrices $\boldsymbol{H}$ and $\boldsymbol{H}_I$ are distributed according to (1). Their randomness determines the statistic distribution of $\mathcal{I}$, which is specified at the second order by the first two cumulant moments, i.e., the mean and the variance, given by

$$\mu_{\mathcal{I}} \triangleq \mathrm{E}_{\boldsymbol{H}_w,\boldsymbol{H}_{w,I}}[\mathcal{I}] \quad \text{and} \quad \sigma_{\mathcal{I}}^2 \triangleq \mathrm{E}_{\boldsymbol{H}_w,\boldsymbol{H}_{w,I}}[\mathcal{I}^2] - \mu_{\mathcal{I}}^2$$

respectively.

In our numerical simulations in Section VII we will compare the asymptotic statistics of $\mathcal{I}$ with the asymptotic statistics of

$$\mathcal{I}_G \triangleq \ln\left(\frac{\det(\boldsymbol{I}_{n_R} + \mathrm{E}_{\boldsymbol{H}_I}[\boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H] + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H)}{\det(\boldsymbol{I}_{n_R} + \mathrm{E}_{\boldsymbol{H}_I}[\boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H])}\right) \tag{7}$$

nat/complex dimension. The mean and variance of $\mathcal{I}_G$ read as

$$\mu_{\mathcal{I}_G} \triangleq \mathrm{E}_{\boldsymbol{H}_w}[\mathcal{I}_G] \quad \text{and} \quad \sigma_{\mathcal{I}_G}^2 \triangleq \mathrm{E}_{\boldsymbol{H}_w}\left[\mathcal{I}_G^2\right] - \mu_{\mathcal{I}_G}^2$$

respectively. The advantage of the random quantity $\mathcal{I}_G$ is that the computation of its mean and its variance can be accomplished with methods obtained for the single user case [43]. Note that $\mu_{\mathcal{I}_G} \leq \mu_{\mathcal{I}}$ due to Jensen's inequality [10, p. 27]. Thus, ergodic capacity (we assume CDIT) is always lower bounded by ergodic capacity when interference is treated as Gaussian noise, as is stated in [7, Section V].

Splitting the $\ln$ in (6) in two parts, we can see that the mean can be obtained by the same asymptotic methods which are used in the interference-free case [7], [44], [45]. However, these methods cannot be used to derive the variance $\sigma_{\mathcal{I}}^2$ or higher moments of $\mathcal{I}$ (because the variance and higher moments of $\mathcal{I}$ are not linear in $\mathcal{I}$), but *superanalysis* has to be applied. Thus, in order to calculate the asymptotic cumulant moments of $\mathcal{I}$, we introduce the *moment generating function*

$$\begin{aligned} G(a) &\triangleq \mathrm{E}[\exp(-a\mathcal{I})] \\ &= \mathrm{E}\left[\frac{\det\left(\boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\right)^{-a}}{\det\left(\boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H\right)^{-a}}\right] \end{aligned} \tag{8}$$

and the *cumulant generating function (CGF)* as $g(a) \triangleq \ln(G(a))$. This definition allows us to obtain the mean and the variance of $\mathcal{I}$ as

$$\mu_{\mathcal{I}} = -g'(0^+) \quad \text{and} \quad \sigma_{\mathcal{I}}^2 = g''(0^+).$$

V. Replica Method and Superanalysis

To derive an asymptotic expression for $g(a)$ we resort to the following tools from theoretical physics: the *replica method* and *superanalysis*. Here we briefly explain the main ideas behind these concepts.

A. Replica Method

Originally, the replica method was introduced in [14] to develop a mathematical model of magnetism, which is characterized by randomness in interactions between magnetic moments. Since then, the replica method has been applied to the physical model of spin glasses [15] and found later application in a variety of research areas, such as digital communications (see [33] for a full account). In the following we briefly summarize the concept in view of its application to our analysis.

The replica method is based on the so called *replica trick*, which allows to evaluate the expectation of the logarithm of a random variable Z by using its moment generating function $g(Z)$. The starting point is the following identity:

$$\mathrm{E}[\ln(Z)] = -\lim_{a\to 0}\frac{d}{da}\mathrm{E}[Z^{-a}].$$

The expectation $\mathrm{E}[Z^{-a}]$ is obtained for positive integer values of a and the result is *extended* to a real interval near 0^+. Information theoretical applications assume that $\ln(Z)$ is a random instance of a mutual information, such as (6), and the expectation is taken with respect to the channel gain (corresponding to the matrices $\boldsymbol{H}$ and $\boldsymbol{H}_I$ in our case). When the evaluation of $\mathrm{E}[Z^{-a}]$ is too hard, asymptotic approximations (in our case in the number of antennas) are called for, typically based on the method of saddlepoint approximation [5].

To sum up, the key assumptions for the replica method are listed as follows:

1) (*Extension from positive integers*) The function $G(a)$ can be extended from the positive integers to a right neighborhood of $a = 0$.
2) (*Interchange of limits*) The limits $n_R \to \infty$ and $a \to 0+$ can be interchanged.
3) (*Replica symmetry*) The saddlepoint approximation admits a real and rotation invariant saddlepoint in the replica space. This assumption is quite plausible because (i) we want to obtain real moments and (ii) there should not be a preferred direction in the replica space.

In our case, the replica space is the space spanned by the auxiliary variables which will be introduced in Section VI-C, (30)–(32). The function F in (36) depends only on these auxiliary variables. To perform a saddlepoint approximation of F we will make an ansatz for the saddlepoint $\mathcal{S}$ in (38). In (38), all matrices evaluated at the saddlepoint are assumed to be scalar multiples of the identity matrix. These matrices are real and rotationally invariant.

For a discussion of these assumptions in a more general context see, for example, [41] and the references therein.

B. Superanalysis

The presence of interference in our framework leads to a determinant with positive exponent $(a > 0)$ in the expression of the moment-generating function (8), which can not be handled as a convergent multi-dimensional Gaussian integral[1]. A similar problem was addressed in the Rayleigh case by [29], who suggested the introduction of (anticommuting) Grassmann variables.

In that contest, (commuting) complex and (anticommuting) Grassmann variables decouple nicely[2] and need not to be considered simultaneously. On the contrary, in our Rician channel contest, complex and Grassmann variables mix together and this is why we have to resort to *superanalysis*.

The origins of superanalysis go back to a paper from 1953 by Schwinger [38], who presented a theoretical analysis based on commuting and anticommuting variables without rigorous mathematical proofs. A detailed explanation of superanalysis is provided by the excellent book of Berezin [2] and a short introduction to Grassmann variables can be found in Appendix II of [29].

As far as our present application is concerned, we have to deal with *superfunctions*, which are functions from a vector space $\mathbb{C}^M$ (or more general domain) into a *Grassmann algebra* $\mathcal{A}$, which is an algebra over the complex numbers $\mathbb{C}$ with a set of anticommuting generators $\mathbb{G} \triangleq \{\theta_1, \ldots, \theta_N\}$. The anticommuting property implies that $\theta_i\theta_j = -\theta_j\theta_i \;\forall\, i, j = 1, \ldots, N$. These generators are called *Grassmann variables*. Though the theory applies even in the case of infinitely many generators, in our case it is sufficient to assume that N is finite. Thus, a superfunction is defined as a power series

$$F(z_1, \ldots, z_M, \theta_1, \ldots, \theta_N) \triangleq \sum_{i_1,\ldots,i_N} f_{i_1,\ldots,i_N}(z_1, \ldots, z_M)\theta_1^{i_1} \cdots \theta_N^{i_N} \quad (9)$$

where $f_{i_1,\ldots,i_N}(z_1, \ldots, z_M)$ are ordinary complex valued functions on $\mathbb{C}^M$. Convergence of superfunctions is defined elementwise for each component $f_{i_1,\ldots,i_N}(z_1, \ldots, z_M)$. We say that a superfunction $F(x, \theta_1, \ldots, \theta_N) = O(g(x))$ if $f_{i_1,\ldots,i_N}(x) = O(g(x))$ for all indices $i_1, \ldots, i_N$ appearing in the power series (9). Integrals of superfunctions [2, p.76] with respect to Grassmann variables are defined by the following rule:

$$\int d_g\theta_1 \ldots d_g\theta_n f(\theta_1, \ldots, \theta_n) = \frac{\partial}{\partial\theta_1} \cdots \frac{\partial}{\partial\theta_n} f(\theta_1, \ldots, \theta_n).$$

Moreover, $d_g\theta_i d_g\theta_j = -d_g\theta_j d_g\theta_i$, and $d_g\theta_i\theta_j = -\theta_j d_g\theta_i$. Essentially, integrals over Grassmann variables act as derivatives which is the reason why Grassmann variables can be used to rewrite positive powers of determinants[3].

[1]See Identity A.1.

[2]More precisely, the matrices defined in (34) become block diagonal.

[3]See Identity B.1.

VI. Main Result

This section contains the main contribution of this paper, summarized in Theorem 1, which characterizes the asymptotic statistics of the mutual information $\mathcal{I}$ defined in (6) as a Gaussian random variable. More precisely, we derive an asymptotic approximation for the CGF of $\mathcal{I}$ under the asymptotic settings stated in Section III-B. Theorem 1 allows to derive the asymptotic mean and variance of $\mathcal{I}$ and to show that all its higher moments tend to zero as $n_R \to \infty$.

Theorem 1: The CGF of the random mutual information $\mathcal{I}$ defined in (6) has the following asymptotic expansion:

$$\begin{aligned} g(a) \sim -a(&\ln\det(\boldsymbol{K_0}) - \ln\det(\boldsymbol{M_0}) \\ &- w_1 z_1 - w_2 z_2 + w_3 z_3) \\ &+ \frac{a^2}{2}\ln\left(\frac{\det(\boldsymbol{G_2})^2}{-\det(\boldsymbol{M_2})\det(\boldsymbol{K_2})}\right). \end{aligned} \quad (10)$$

In this expression, we defined the matrices

$$\boldsymbol{K_0} \triangleq \begin{pmatrix} \hat{\boldsymbol{R}} & \tilde{\boldsymbol{H}} & \tilde{\boldsymbol{H}}_I \\ -\tilde{\boldsymbol{H}}^H & \hat{\boldsymbol{T}} & \boldsymbol{0} \\ -\tilde{\boldsymbol{H}}_I^H & \boldsymbol{0} & \hat{\boldsymbol{I}} \end{pmatrix} \quad (11)$$

$$\boldsymbol{M_0} \triangleq \begin{pmatrix} \check{\boldsymbol{R}} & -\tilde{\boldsymbol{H}}_I \\ \tilde{\boldsymbol{H}}_I^H & \check{\boldsymbol{I}} \end{pmatrix} \quad (12)$$

and those reported in (13), shown at the bottom of the page. The above matrices are based on the following two sets of auxiliary matrices:

$$\begin{cases} \boldsymbol{A_M} \triangleq (\check{\boldsymbol{R}} + \tilde{\boldsymbol{H}}_I\check{\boldsymbol{I}}^{-1}\tilde{\boldsymbol{H}}_I^H)^{-1} \\ \boldsymbol{B_M} \triangleq \check{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I\boldsymbol{D_M} \\ \boldsymbol{C_M} \triangleq -\check{\boldsymbol{I}}^{-1}\tilde{\boldsymbol{H}}_I^H\boldsymbol{A_M} \\ \boldsymbol{D_M} \triangleq (\check{\boldsymbol{I}} + \tilde{\boldsymbol{H}}_I^H\check{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I)^{-1} \end{cases} \quad (14)$$

and

$$\boldsymbol{M_2} \triangleq \begin{pmatrix} \operatorname{tr}((\boldsymbol{A_M R}_I)^2) & 1 + \operatorname{tr}((\boldsymbol{C_M R}_I)(\boldsymbol{B_M}\tilde{\boldsymbol{T}}_I)) \\ 1 + \operatorname{tr}((\boldsymbol{B_M}\tilde{\boldsymbol{T}}_I)(\boldsymbol{C_M R}_I)) & \operatorname{tr}((\boldsymbol{D_M}\tilde{\boldsymbol{T}}_I)^2) \end{pmatrix}$$

$$-\boldsymbol{K_2} \triangleq \begin{pmatrix} \operatorname{tr}((\boldsymbol{A_K R})^2) & \operatorname{tr}(\boldsymbol{D_K R B_K}\tilde{\boldsymbol{T}}) + 1 & \operatorname{tr}(\boldsymbol{A_K R A_K R}_I) & \operatorname{tr}(\boldsymbol{G_K R C_K}\tilde{\boldsymbol{T}}_I) \\ \operatorname{tr}(\boldsymbol{B_K}\tilde{\boldsymbol{T}}\boldsymbol{D_K R}) + 1 & \operatorname{tr}((\boldsymbol{E_K}\tilde{\boldsymbol{T}})^2) & \operatorname{tr}(\boldsymbol{B_K}\tilde{\boldsymbol{T}}\boldsymbol{D_K R}_I) & \operatorname{tr}(\boldsymbol{H_K}\tilde{\boldsymbol{T}}\boldsymbol{F_K}\tilde{\boldsymbol{T}}_I) \\ \operatorname{tr}(\boldsymbol{A_K R}_I\boldsymbol{A_K R}) & \operatorname{tr}(\boldsymbol{D_K R}_I\boldsymbol{B_K}\tilde{\boldsymbol{T}}) & \operatorname{tr}((\boldsymbol{A_K R}_I)^2) & \operatorname{tr}(\boldsymbol{G_K R}_I\boldsymbol{C_K}\tilde{\boldsymbol{T}}_I) + 1 \\ \operatorname{tr}(\boldsymbol{C_K}\tilde{\boldsymbol{T}}_I\boldsymbol{G_K R}) & \operatorname{tr}(\boldsymbol{F_K}\tilde{\boldsymbol{T}}_I\boldsymbol{H_K}\tilde{\boldsymbol{T}}) & \operatorname{tr}(\boldsymbol{C_K}\tilde{\boldsymbol{T}}_I\boldsymbol{G_K R}_I) + 1 & \operatorname{tr}((\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I)^2) \end{pmatrix}$$

$$-\boldsymbol{G_2} \triangleq \begin{pmatrix} \operatorname{tr}(\boldsymbol{C_K}\tilde{\boldsymbol{T}}_I\boldsymbol{C_M R}_I) - 1 & \operatorname{tr}(\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I\boldsymbol{D_M}\tilde{\boldsymbol{T}}_I) \\ \operatorname{tr}(\boldsymbol{A_K R}_I\boldsymbol{A_M R}_I) & \operatorname{tr}(\boldsymbol{G_K R}_I\boldsymbol{B_M}\tilde{\boldsymbol{T}}_I) - 1 \end{pmatrix} \quad (13)$$

$$\begin{cases} \boldsymbol{A_K} \triangleq \left(\hat{\boldsymbol{R}} + \tilde{\boldsymbol{H}}\hat{\boldsymbol{T}}^{-1}\tilde{\boldsymbol{H}}^H + \tilde{\boldsymbol{H}}_I\hat{\boldsymbol{I}}^{-1}\tilde{\boldsymbol{H}}_I^H\right)^{-1} \\ \boldsymbol{B_K} \triangleq -\hat{\boldsymbol{R}}^{-1}(\tilde{\boldsymbol{H}}\boldsymbol{E_K} + \tilde{\boldsymbol{H}}_I\boldsymbol{H_K}) \\ \boldsymbol{C_K} \triangleq -\hat{\boldsymbol{R}}^{-1}(\tilde{\boldsymbol{H}}\boldsymbol{F_K} + \tilde{\boldsymbol{H}}_I\boldsymbol{I_K}) \\ \boldsymbol{D_K} \triangleq \hat{\boldsymbol{T}}^{-1}\tilde{\boldsymbol{H}}^H\boldsymbol{A_K} \\ \boldsymbol{E_{K_s}} \triangleq \Big(\hat{\boldsymbol{T}} + \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}} - \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I \\ \qquad \cdot(\hat{\boldsymbol{I}} + \tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I)^{-1}\tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}\Big)^{-1} \\ \boldsymbol{F_K} \triangleq -(\hat{\boldsymbol{T}} + \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}})^{-1}\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I\boldsymbol{I_K} \\ \boldsymbol{G_K} \triangleq \hat{\boldsymbol{I}}^{-1}\tilde{\boldsymbol{H}}_I^H\boldsymbol{A_K} \\ \boldsymbol{H_K} \triangleq -\left(\hat{\boldsymbol{I}} + \tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I\right)^{-1}\tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}\boldsymbol{E_K} \\ \boldsymbol{I_K} \triangleq \Big(\hat{\boldsymbol{I}} + \tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I - \tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}} \\ \qquad \cdot(\hat{\boldsymbol{T}} + \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}})^{-1}\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I\Big)^{-1}. \end{cases} \tag{15}$$

Above, we used the following short-hand notation:

$$\begin{cases} \hat{\boldsymbol{R}} \triangleq \boldsymbol{I}_{n_R} + z_1\boldsymbol{R} + z_2\boldsymbol{R}_I \\ \hat{\boldsymbol{T}} \triangleq \boldsymbol{I}_t + w_1\tilde{\boldsymbol{T}} \\ \hat{\boldsymbol{I}} \triangleq \boldsymbol{I}_{n_I} + w_2\tilde{\boldsymbol{T}}_I \\ \check{\boldsymbol{R}} \triangleq \boldsymbol{I}_{n_R} + z_3\boldsymbol{R}_I \\ \check{\boldsymbol{I}} \triangleq \boldsymbol{I}_{n_I} + w_3\tilde{\boldsymbol{T}}_I. \end{cases} \tag{16}$$

The positive real numbers $\{w_1, w_2, z_1, z_2\}$ and $\{w_3, z_3\}$ are the solutions of the following two systems of coupled fixed-point equations:

$$\begin{cases} w_1 = \mathrm{tr}(\boldsymbol{A_K}\boldsymbol{R}) & z_1 = \mathrm{tr}(\boldsymbol{E_K}\tilde{\boldsymbol{T}}) \\ w_2 = \mathrm{tr}(\boldsymbol{A_K}\boldsymbol{R}_I) & z_2 = \mathrm{tr}(\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I) \\ w_3 = \mathrm{tr}(\boldsymbol{A_M}\boldsymbol{R}_I) & z_3 = \mathrm{tr}(\boldsymbol{D_M}\tilde{\boldsymbol{T}}_I). \end{cases} \tag{17}$$

From the CGF expansion (10), we obtain directly the asymptotic mean and variance of $\mathcal{I}$ as

$$\mu_{\mathcal{I}} \sim \ln\det(\boldsymbol{K_0}) - \ln\det(\boldsymbol{M_0}) - w_1z_1 - w_2z_2 + w_3z_3$$

and

$$\sigma_{\mathcal{I}}^2 \sim \ln\left(\frac{\det(\boldsymbol{G_2})^2}{-\det(\boldsymbol{M_2})\det(\boldsymbol{K_2})}\right).$$

Higher-order moments tend to zero as $n_R \to \infty$, so that $\mathcal{I}$ converges in distribution to the Gaussian random variable $\mathcal{N}(\mu_I, \sigma_I^2)$.

Proof: The Theorem is proved in Sections VI-A to VI-D.

A. Rewriting Determinants

In what follows we use repeatedly Identities A.1 and B.1 in order to represent the determinants appearing in (8) in integral form. The numerator of (8) can be written as

$$\begin{aligned} &\det\left(\boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\right)^{-a} \\ &= \int D_c\boldsymbol{U}\int D_c\boldsymbol{V}\int D_c\boldsymbol{W} \\ &\quad \cdot \mathrm{etr}(-\pi(\boldsymbol{U}^H\boldsymbol{U} + \boldsymbol{V}^H\boldsymbol{V} + \boldsymbol{W}^H\boldsymbol{W})) \\ &\quad \cdot \tilde{f}(\boldsymbol{U},\boldsymbol{V})f_w(\boldsymbol{U},\boldsymbol{V})\tilde{f}_I(\boldsymbol{U},\boldsymbol{W})f_{w,I}(\boldsymbol{U},\boldsymbol{W}) \end{aligned} \tag{18}$$

where integration is carried out over the domains $\boldsymbol{U} \in \mathbb{C}^{n_R\times a}$, $\boldsymbol{V} \in \mathbb{C}^{n_T\times a}$, $\boldsymbol{W} \in \mathbb{C}^{n_I\times a}$, and we set

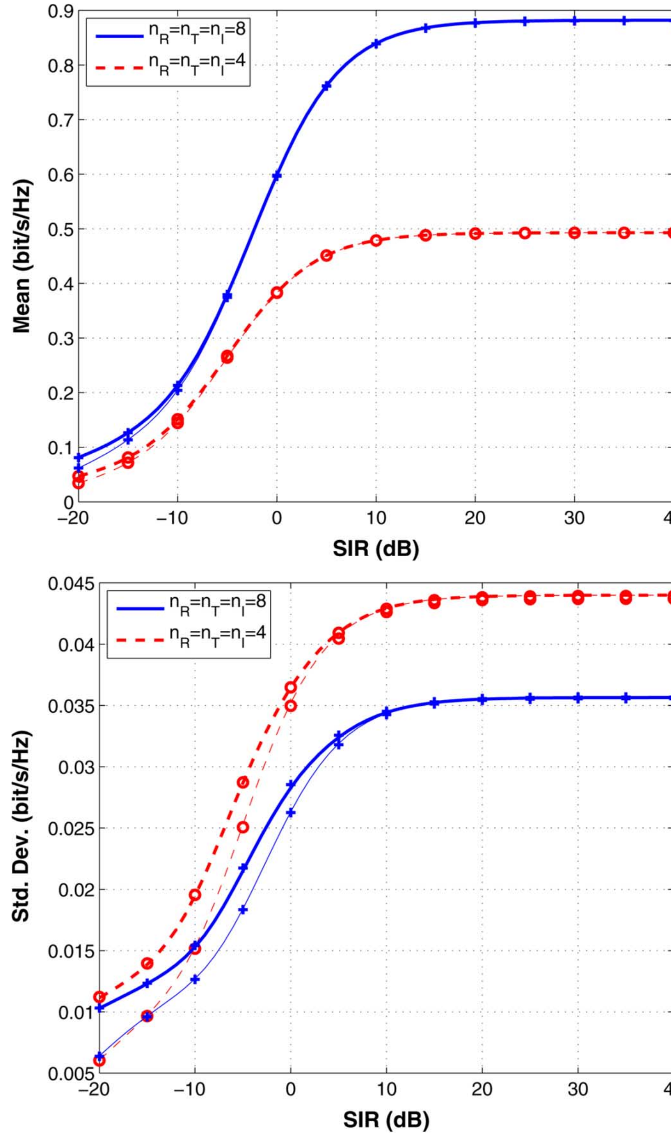

Fig. 1. Mean and standard deviation of mutual information $\mathcal{I}$ versus SIR with SNR $= -10$ dB.

$$\tilde{f}(\boldsymbol{U},\boldsymbol{V}) \triangleq \mathrm{etr}(-\pi i(\boldsymbol{U}^H\tilde{\boldsymbol{H}}\boldsymbol{V} + \boldsymbol{V}^H\tilde{\boldsymbol{H}}^H\boldsymbol{U})) \tag{19}$$

$$\tilde{f}_I(\boldsymbol{U},\boldsymbol{W}) \triangleq \mathrm{etr}(-\pi i(\boldsymbol{U}^H\tilde{\boldsymbol{H}}_I\boldsymbol{W} + \boldsymbol{W}^H\tilde{\boldsymbol{H}}_I^H\boldsymbol{U})) \tag{20}$$

$$\begin{aligned} f_w(\boldsymbol{U},\boldsymbol{V}) \triangleq \mathrm{etr}\Big(&-\pi i(\boldsymbol{T}^{1/2}\boldsymbol{Q}^{1/2}\boldsymbol{V}\boldsymbol{U}^H\boldsymbol{R}^{1/2}\boldsymbol{H}_w \\ &+\boldsymbol{H}_w^H\boldsymbol{R}^{1/2}\boldsymbol{U}\boldsymbol{V}^H\boldsymbol{Q}^{1/2}\boldsymbol{T}^{1/2})\Big) \end{aligned} \tag{21}$$

$$\begin{aligned} f_{w,I}(\boldsymbol{U},\boldsymbol{W}) \triangleq \mathrm{etr}\Big(&-\pi i\Big(\boldsymbol{T}_I^{1/2}\boldsymbol{Q}_I^{1/2}\boldsymbol{W}\boldsymbol{U}^H\boldsymbol{R}_I^{1/2}\boldsymbol{H}_{w,I} \\ &+\boldsymbol{H}_{w,I}^H\boldsymbol{R}_I^{1/2}\boldsymbol{U}\boldsymbol{W}^H\boldsymbol{Q}_I^{1/2}\boldsymbol{T}_I^{1/2}\Big)\Big) \end{aligned} \tag{22}$$

where $i \triangleq \sqrt{-1}$. The derivation of (18) is described in detail in Appendix C. Similarly, we can write the inverse of the denominator in the argument of the logarithm from (8) as

$$\begin{aligned} &\det\left(\boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H\right)^a \\ &= \int D_g(\boldsymbol{\Psi},\bar{\boldsymbol{\Psi}})\int D_g(\boldsymbol{\Omega},\bar{\boldsymbol{\Omega}})\mathrm{etr}(\bar{\boldsymbol{\Psi}}\boldsymbol{\Psi} + \bar{\boldsymbol{\Omega}}\boldsymbol{\Omega}) \\ &\quad \cdot \tilde{h}_I(\boldsymbol{\Psi},\bar{\boldsymbol{\Psi}},\boldsymbol{\Omega},\bar{\boldsymbol{\Omega}})h_{w,I}(\boldsymbol{\Psi},\bar{\boldsymbol{\Psi}},\boldsymbol{\Omega},\bar{\boldsymbol{\Omega}}) \end{aligned} \tag{23}$$

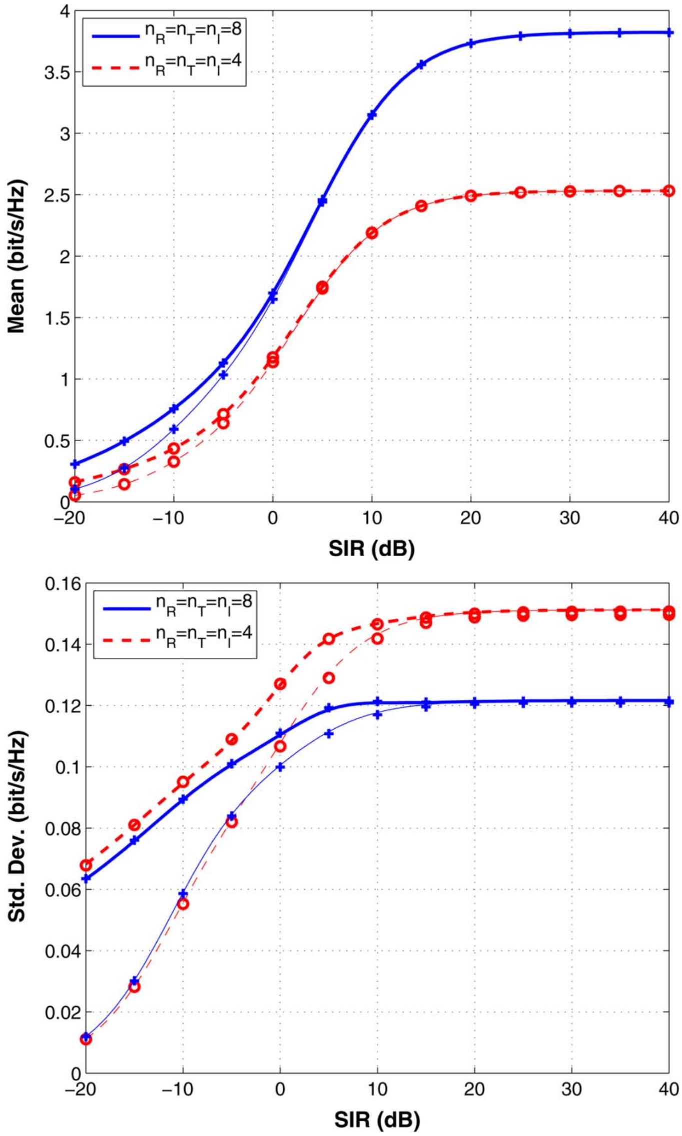

Fig. 2. Same as Fig. 1 but $\mathsf{SNR} = 0$ dB.

where integration is carried out over the domains $\boldsymbol{\Psi} \in \mathbb{G}^{n_R \times a}$, $\bar{\boldsymbol{\Psi}} \in \mathbb{G}^{a \times n_R}$, $\boldsymbol{\Omega} \in \mathbb{G}^{n_I \times a}$, $\bar{\boldsymbol{\Omega}} \in \mathbb{G}^{a \times n_I}$, and

$$\tilde{h}_I(\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}) \triangleq \operatorname{etr}\left(\bar{\boldsymbol{\Omega}}\tilde{\boldsymbol{H}}_I^H \boldsymbol{\Psi} - \bar{\boldsymbol{\Psi}}\tilde{\boldsymbol{H}}_I \boldsymbol{\Omega}\right) \tag{24}$$

$$h_{w,I}(\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}) \triangleq \operatorname{etr}\Big(\boldsymbol{T}_I^{1/2}\boldsymbol{Q}_I^{1/2}\boldsymbol{\Omega}\bar{\boldsymbol{\Psi}}\boldsymbol{R}_I^{1/2}\boldsymbol{H}_{w,I} - \boldsymbol{H}_{w,I}^H \boldsymbol{R}_I^{1/2}\boldsymbol{\Psi}\bar{\boldsymbol{\Omega}}\boldsymbol{Q}_I^{1/2}\boldsymbol{T}_I^{1/2}\Big). \tag{25}$$

The derivation of (23), based on Grassmann integrals, is reported in Appendix D. In the remainder of this proof we adopted the following notation:

1) Symbols with a tilde, i.e., $\tilde{f}$, $\tilde{f}_I$, and $\tilde{h}_I$, depend only on LOS components and do not affect expectations. The index "I" stands for interference, meaning dependence on $\bar{\boldsymbol{H}}_I$ and not on $\bar{\boldsymbol{H}}$.
2) Quantities with an index "w," i.e., f_w, $f_{w,I}$, and $h_{w,I}$, refer to expectations. The index "I" has the same meaning as in point 1).

B. Calculating Expectations

The expectation in (8) can now be calculated. Using Identity (A.1) we find that

$$\begin{aligned} &\mathrm{E}_{\boldsymbol{H}_w, \boldsymbol{H}_{w,I}}[f_w(\boldsymbol{U},\boldsymbol{V}) f_{w,I}(\boldsymbol{U},\boldsymbol{W}) h_{w,I}(\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}})] \\ &\quad = \mathrm{E}_{\boldsymbol{H}_w}[f_w(\boldsymbol{U},\boldsymbol{V})] \\ &\qquad \cdot \mathrm{E}_{\boldsymbol{H}_{w,I}}[f_{w,I}(\boldsymbol{U},\boldsymbol{W}) h_{w,I}(\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}})] \\ &\quad = h_I(\boldsymbol{U},\boldsymbol{W}; \boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}) f(\boldsymbol{U},\boldsymbol{V}) f_I(\boldsymbol{U},\boldsymbol{W}) \end{aligned}$$

with

$$f(\boldsymbol{U},\boldsymbol{V}) \triangleq \operatorname{etr}(-\pi^2 \boldsymbol{V}^H \tilde{\boldsymbol{T}} \boldsymbol{V} \boldsymbol{U}^H \boldsymbol{R} \boldsymbol{U}) \tag{26}$$

$$f_I(\boldsymbol{U},\boldsymbol{W}) \triangleq \operatorname{etr}(-\pi^2 \boldsymbol{W}^H \tilde{\boldsymbol{T}}_I \boldsymbol{W} \boldsymbol{U}^H \boldsymbol{R}_I \boldsymbol{U}) \tag{27}$$

and

$$\begin{aligned} &h_I(\boldsymbol{U},\boldsymbol{W}; \boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}) \\ &\quad \triangleq \operatorname{etr}(\bar{\boldsymbol{\Omega}}\tilde{\boldsymbol{T}}_I \boldsymbol{\Omega}\bar{\boldsymbol{\Psi}}\boldsymbol{R}_I\boldsymbol{\Psi}) \\ &\qquad \cdot \operatorname{etr}(-\pi(\bar{\boldsymbol{\Psi}}\boldsymbol{R}_I\boldsymbol{U}\boldsymbol{W}^H\tilde{\boldsymbol{T}}_I\boldsymbol{\Omega} \\ &\qquad + \bar{\boldsymbol{\Omega}}\tilde{\boldsymbol{T}}_I\boldsymbol{W}\boldsymbol{U}^H\boldsymbol{R}_I\boldsymbol{\Psi})). \end{aligned} \tag{28}$$

To sum up, we end up with

$$\begin{aligned} G(a) = &\int D_c\boldsymbol{U} \int D_c\boldsymbol{V} \int D_c\boldsymbol{W} \\ &\cdot \operatorname{etr}(-\pi(\boldsymbol{U}^H\boldsymbol{U} + \boldsymbol{V}^H\boldsymbol{V} + \boldsymbol{W}^H\boldsymbol{W})) \\ &\cdot \tilde{f}(\boldsymbol{U},\boldsymbol{V})\tilde{f}_I(\boldsymbol{U},\boldsymbol{W}) f(\boldsymbol{U},\boldsymbol{V}) f_I(\boldsymbol{U},\boldsymbol{W}) \\ &\cdot \int D_g(\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}) \int D_g(\boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}) \operatorname{etr}(\bar{\boldsymbol{\Psi}}\boldsymbol{\Psi} + \bar{\boldsymbol{\Omega}}\boldsymbol{\Omega}) \\ &\cdot \tilde{h}_I(\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}) h_I(\boldsymbol{U},\boldsymbol{W}; \boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}). \end{aligned} \tag{29}$$

Here, $\tilde{f}$, $\tilde{f}_I$, and $\tilde{h}_I$ depend on the LOS components of the channel matrices and are defined in (19), (20), and (24), respectively. The functions f, f_I, and h_I are from (26), (27), and (28), respectively.

C. Calculating Integrals

In order to integrate out the complex matrix variables $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{W}$ and the Grassmann matrix variables $\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}$, we have to disentangle the quartic products in the exponents of f, f_I, and h_I. Applying Identity A.2 to f in (26) and f_I in (27) yields

$$\begin{aligned} f(\boldsymbol{U},\boldsymbol{V}) = &\int d\mu(\boldsymbol{W_1}, \boldsymbol{Z_1}) \\ &\cdot \operatorname{etr}(\boldsymbol{Z_1}\boldsymbol{W_1} - \pi\boldsymbol{V}^H\tilde{\boldsymbol{T}}\boldsymbol{V}\boldsymbol{W_1} - \pi\boldsymbol{Z_1}\boldsymbol{U}^H\boldsymbol{R}\boldsymbol{U}) \end{aligned} \tag{30}$$

$$\begin{aligned} f_I(\boldsymbol{U},\boldsymbol{W}) = &\int d\mu(\boldsymbol{W_2}, \boldsymbol{Z_2}) \\ &\cdot \operatorname{etr}(\boldsymbol{Z_2}\boldsymbol{W_2} - \pi\boldsymbol{W}^H\tilde{\boldsymbol{T}}_I\boldsymbol{W}\boldsymbol{W_2} - \pi\boldsymbol{Z_2}\boldsymbol{U}^H\boldsymbol{R}_I\boldsymbol{U}). \end{aligned} \tag{31}$$

Applying Identities A.2 and B.1 (with $\boldsymbol{A} = \boldsymbol{I}_a$ and $\boldsymbol{B} = \boldsymbol{I}_a$) to h_I in (28) leads to (32).

$$\begin{aligned} h_I(\boldsymbol{U},\boldsymbol{W}; \boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}) = &\int d\mu(\boldsymbol{W_3}, \boldsymbol{Z_3}) \\ \times &\operatorname{etr}(\boldsymbol{Z_3}\boldsymbol{W_3} + \bar{\boldsymbol{\Omega}}\tilde{\boldsymbol{T}}_I\boldsymbol{\Omega}\boldsymbol{W_3} - \boldsymbol{Z_3}\bar{\boldsymbol{\Psi}}\boldsymbol{R}_I\boldsymbol{\Psi}) \\ \cdot &\int D_g(\boldsymbol{\Theta_1}, \bar{\boldsymbol{\Theta}}_\mathbf{1}) \operatorname{etr}(\bar{\boldsymbol{\Theta}}_\mathbf{1}\boldsymbol{\Theta_1} + \sqrt{\pi}\bar{\boldsymbol{\Psi}}\boldsymbol{R}_I\boldsymbol{U}\boldsymbol{\Theta_1} \\ &+ \sqrt{\pi}\bar{\boldsymbol{\Theta}}_\mathbf{1}\boldsymbol{W}^H\tilde{\boldsymbol{T}}_I\boldsymbol{\Omega}) \end{aligned}$$

$$\cdot \int D_g(\boldsymbol{\Theta_2}, \bar{\boldsymbol{\Theta}}_2) \mathrm{etr}(\bar{\boldsymbol{\Theta}}_2 \boldsymbol{\Theta_2} + \sqrt{\pi} \bar{\boldsymbol{\Omega}} \tilde{\boldsymbol{T}}_I \boldsymbol{W} \boldsymbol{\Theta_2} + \sqrt{\pi} \bar{\boldsymbol{\Theta}}_2 \boldsymbol{U}^H \boldsymbol{R}_I \boldsymbol{\Psi}) \quad (32)$$

Inserting (30), (31), and (32) into the expression of $G(a)$ in (29) allows us to perform the integration over matrices $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{W}$ and $\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}, \boldsymbol{\Omega}, \bar{\boldsymbol{\Omega}}$. The generating function (8) can be written as in (33).

$$\begin{aligned} G(a) = &\int d\mu(\boldsymbol{W_1}, \boldsymbol{Z_1}) \int d\mu(\boldsymbol{W_2}, \boldsymbol{Z_2}) \int d\mu(\boldsymbol{W_3}, \boldsymbol{Z_3}) \\ &\times \int D_g(\boldsymbol{\Theta_1}, \bar{\boldsymbol{\Theta}}_1) \int D_g(\boldsymbol{\Theta_2}, \bar{\boldsymbol{\Theta}}_2) \\ &\cdot \mathrm{etr}(\boldsymbol{Z_1}\boldsymbol{W_1} + \boldsymbol{Z_2}\boldsymbol{W_2} + \boldsymbol{Z_3}\boldsymbol{W_3}) \\ &\times \mathrm{etr}(\bar{\boldsymbol{\Theta}}_1 \boldsymbol{\Theta_1}) \mathrm{etr}(\bar{\boldsymbol{\Theta}}_2 \boldsymbol{\Theta_2}) \\ &\times \det(\boldsymbol{M}) \det(\boldsymbol{K} + \bar{\boldsymbol{\Gamma}} \boldsymbol{M}^{-1} \boldsymbol{\Gamma})^{-1} \end{aligned} \quad (33)$$

In this equation, we used the definitions given by (34), shown at the bottom of the page. The derivation of (33) is based on Identities A.1 and B.1 and is postponed to Appendix E.

D. Saddlepoint Approximation

Note that we have to perform a *simultaneous* saddlepoint approximation in the Grassmann and complex integration variables appearing in (33). To this end, we define the *supermatrix* $\boldsymbol{\mathfrak{X}}$ [2, p. 82] by

$$\boldsymbol{\mathfrak{X}} = \begin{pmatrix} \boldsymbol{M} & -\boldsymbol{\Gamma} \\ \bar{\boldsymbol{\Gamma}} & \boldsymbol{K} \end{pmatrix}$$

with superdeterminant $\mathrm{sdet}(\boldsymbol{\mathfrak{X}})$ and supertrace $\mathrm{str}(\boldsymbol{\mathfrak{X}})$

$$\begin{aligned} \mathrm{sdet}(\boldsymbol{\mathfrak{X}}) &\triangleq \det(\boldsymbol{M}) \det(\boldsymbol{K} + \bar{\boldsymbol{\Gamma}} \boldsymbol{M}^{-1} \boldsymbol{\Gamma})^{-1} \\ \mathrm{str}(\boldsymbol{\mathfrak{X}}) &\triangleq \mathrm{tr}(\boldsymbol{M}) - \mathrm{tr}(\boldsymbol{K}). \end{aligned}$$

Equation (33) can then be rewritten as

$$\begin{aligned} G(a) = &\int d\mu(\boldsymbol{W_1}, \boldsymbol{Z_1}) \int d\mu(\boldsymbol{W_2}, \boldsymbol{Z_2}) \int d\mu(\boldsymbol{W_3}, \boldsymbol{Z_3}) \\ &\cdot \int D_g(\boldsymbol{\Theta_1}, \bar{\boldsymbol{\Theta}}_1) \int D_g(\boldsymbol{\Theta_2}, \bar{\boldsymbol{\Theta}}_2) \exp(-F) \end{aligned} \quad (35)$$

with

$$\begin{aligned} F \triangleq\ & F(\boldsymbol{W_1}, \boldsymbol{Z_1}, \boldsymbol{W_2}, \boldsymbol{Z_2}, \boldsymbol{W_3}, \boldsymbol{Z_3}, \boldsymbol{\Theta_1}, \bar{\boldsymbol{\Theta}}_1, \boldsymbol{\Theta_2}, \bar{\boldsymbol{\Theta}}_2) \\ =& -\ln \mathrm{sdet}(\boldsymbol{\mathfrak{X}}) - \mathrm{tr}(\boldsymbol{Z_1}\boldsymbol{W_1} + \boldsymbol{Z_2}\boldsymbol{W_2} + \boldsymbol{Z_3}\boldsymbol{W_3} \\ &+ \bar{\boldsymbol{\Theta}}_1 \boldsymbol{\Theta_1} + \bar{\boldsymbol{\Theta}}_2 \boldsymbol{\Theta_2}). \end{aligned} \quad (36)$$

Using the fact that $\ln \mathrm{sdet}(\boldsymbol{\mathfrak{X}}) = \mathrm{str}(\ln \boldsymbol{\mathfrak{X}})$ [2, p. 112] and the multiplicative property of superdeterminants [2, p. 101], we get the following expansion:

$$\ln \mathrm{sdet}(\boldsymbol{\mathfrak{X}}) = \ln \mathrm{sdet}(\boldsymbol{\mathfrak{X}}|_{\mathcal{S}}) + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \mathrm{str}((\boldsymbol{\mathfrak{X}}^{-1}|_{\mathcal{S}} \delta \boldsymbol{\mathfrak{X}})^k) \quad (37)$$

where $\mathcal{S}$ is assumed to be a real and replica-symmetric saddlepoint

$$\begin{aligned} \boldsymbol{W_1} &= w_1 \boldsymbol{I}_a + \delta \boldsymbol{W_1} \\ \boldsymbol{Z_1} &= z_1 \boldsymbol{I}_a + \delta \boldsymbol{Z_1} \\ \boldsymbol{W_2} &= w_2 \boldsymbol{I}_a + \delta \boldsymbol{W_2} \\ \boldsymbol{Z_2} &= z_2 \boldsymbol{I}_a + \delta \boldsymbol{Z_2} \\ \boldsymbol{W_3} &= w_3 \boldsymbol{I}_a + \delta \boldsymbol{W_3} \\ \boldsymbol{Z_3} &= -z_3 \boldsymbol{I}_a - \delta \boldsymbol{Z_3} \\ \boldsymbol{\Theta_1} &= \delta \boldsymbol{\Theta_1} \\ \bar{\boldsymbol{\Theta}}_1 &= \delta \bar{\boldsymbol{\Theta}}_1 \\ \boldsymbol{\Theta_2} &= \delta \boldsymbol{\Theta_2} \\ \bar{\boldsymbol{\Theta}}_2 &= \delta \bar{\boldsymbol{\Theta}}_2 \end{aligned} \quad (38)$$

of F in (36) and $|_{\mathcal{S}}$ denotes evaluation at the saddlepoint.[4] Inserting (37) and the ansatz for the saddlepoint (38) into (36) leads to an expansion $F = \sum_{k=0}^{\infty} F_k$ with

$$\begin{aligned} F_0 &\triangleq -\ln \mathrm{sdet}(\boldsymbol{\mathfrak{X}})|_{\mathcal{S}} - a(w_1 z_1 + w_2 z_2 - w_3 z_3) \\ F_1 &\triangleq -\mathrm{str}(\boldsymbol{\mathfrak{X}}^{-1}|_{\mathcal{S}} \delta \boldsymbol{\mathfrak{X}}) \\ &\quad - \mathrm{tr}(z_1 \delta \boldsymbol{W_1} + w_1 \delta \boldsymbol{Z_1} + z_2 \delta \boldsymbol{W_2} \\ &\qquad + w_2 \delta \boldsymbol{Z_2} - z_3 \delta \boldsymbol{W_3} - w_3 \delta \boldsymbol{Z_3}) \\ F_2 &\triangleq \frac{1}{2} \mathrm{str}((\boldsymbol{\mathfrak{X}}^{-1}|_{\mathcal{S}} \delta \boldsymbol{\mathfrak{X}})^2) \\ &\quad - \mathrm{tr}(\delta \boldsymbol{Z_1} \delta \boldsymbol{W_1} + \delta \boldsymbol{Z_2} \delta \boldsymbol{W_2} \\ &\qquad - \delta \boldsymbol{Z_3} \delta \boldsymbol{W_3} + \delta \bar{\boldsymbol{\Theta}}_1 \delta \boldsymbol{\Theta_1} + \delta \bar{\boldsymbol{\Theta}}_2 \delta \boldsymbol{\Theta_2}) \\ F_k &\triangleq \frac{(-1)^k}{k} \mathrm{str}((\boldsymbol{\mathfrak{X}}^{-1}|_{\mathcal{S}} \delta \boldsymbol{\mathfrak{X}})^k) \quad \forall\, k > 2. \end{aligned} \quad (39)$$

Before starting to evaluate the different F_k's, we need to elaborate

$$\boldsymbol{\mathfrak{X}}^{-1}|_{\mathcal{S}} \delta \boldsymbol{\mathfrak{X}} = \begin{pmatrix} \boldsymbol{M}^{-1}|_{\mathcal{S}} \delta \boldsymbol{M} & -\boldsymbol{M}^{-1}|_{\mathcal{S}} \delta \boldsymbol{\Gamma} \\ \boldsymbol{K}^{-1}|_{\mathcal{S}} \delta \bar{\boldsymbol{\Gamma}} & \boldsymbol{K}^{-1}|_{\mathcal{S}} \delta \boldsymbol{K} \end{pmatrix}.$$

[4] The symbol δ in (38) denotes deviations around the saddlepoint $\mathcal{S}$. For example, $\boldsymbol{W_1} = \boldsymbol{W_1}|_{\mathcal{S}} + \delta \boldsymbol{W_1} = w_1 \boldsymbol{I}_a + \delta \boldsymbol{W_1}$, which defines the deviation $\delta \boldsymbol{W_1} \in \mathbb{C}^{a \times a}$ in a unique way. See also [29, eq. (18)] and [43, p. 13].

$$\begin{cases} \boldsymbol{M} \triangleq \begin{pmatrix} \boldsymbol{I}_{n_R a} - \boldsymbol{Z_3}^T \otimes \boldsymbol{R}_I & -\boldsymbol{I}_a \otimes \tilde{\boldsymbol{H}}_I \\ \boldsymbol{I}_a \otimes \tilde{\boldsymbol{H}}_I^H & \boldsymbol{I}_{n_I a} + \boldsymbol{W_3}^T \otimes \tilde{\boldsymbol{T}}_I \end{pmatrix} \\ \boldsymbol{\Gamma} \triangleq \begin{pmatrix} \boldsymbol{\Theta_1}^T \otimes \boldsymbol{R}_I & 0 & 0 \\ 0 & 0 & \boldsymbol{\Theta_2}^T \otimes \tilde{\boldsymbol{T}}_I \end{pmatrix} \\ \bar{\boldsymbol{\Gamma}} \triangleq \begin{pmatrix} \bar{\boldsymbol{\Theta}}_2^T \otimes \boldsymbol{R}_I & 0 \\ 0 & 0 \\ 0 & \bar{\boldsymbol{\Theta}}_1^T \otimes \tilde{\boldsymbol{T}}_I \end{pmatrix} \\ \boldsymbol{K} \triangleq \begin{pmatrix} \boldsymbol{I}_{n_R a} + \boldsymbol{Z_1}^T \otimes \boldsymbol{R} + \boldsymbol{Z_2}^T \otimes \boldsymbol{R}_I & \boldsymbol{I}_a \otimes \tilde{\boldsymbol{H}} & \boldsymbol{I}_a \otimes \tilde{\boldsymbol{H}}_I \\ -\boldsymbol{I}_a \otimes \tilde{\boldsymbol{H}}^H & \boldsymbol{I}_{n_T a} + \boldsymbol{W_1}^T \otimes \tilde{\boldsymbol{T}} & 0 \\ -\boldsymbol{I}_a \otimes \tilde{\boldsymbol{H}}_I^H & 0 & \boldsymbol{I}_{n_I a} + \boldsymbol{W_2}^T \otimes \tilde{\boldsymbol{T}}_I \end{pmatrix}. \end{cases} \quad (34)$$

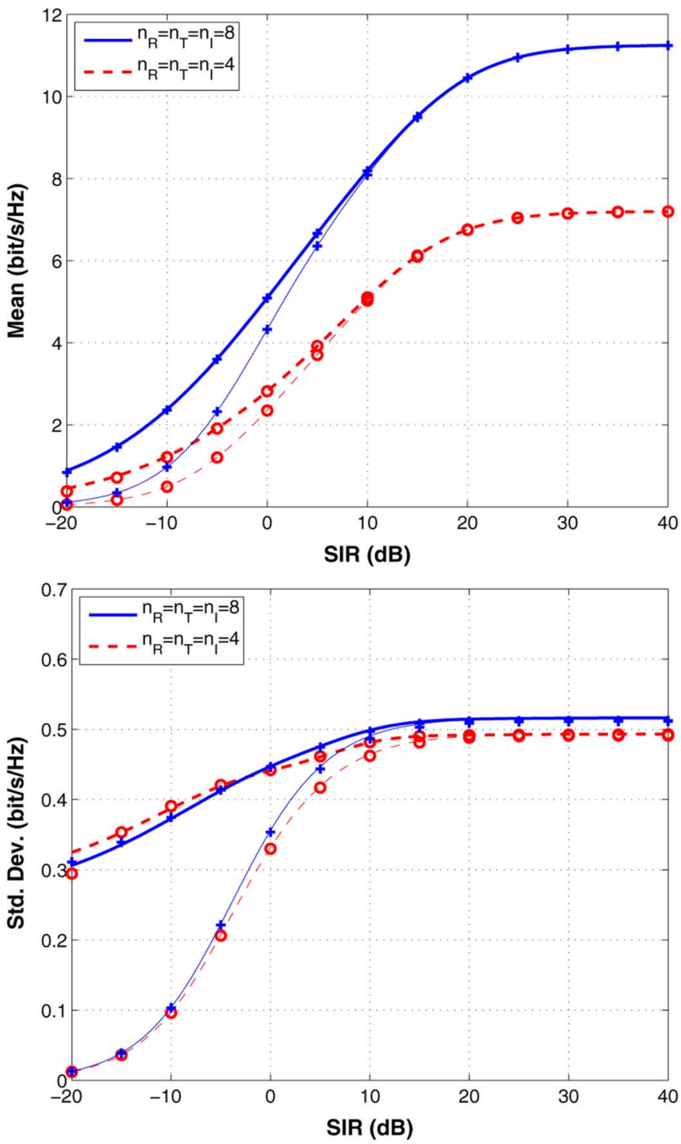

Fig. 3. Same as Fig. 1 but $\mathsf{SNR} = 10$ dB.

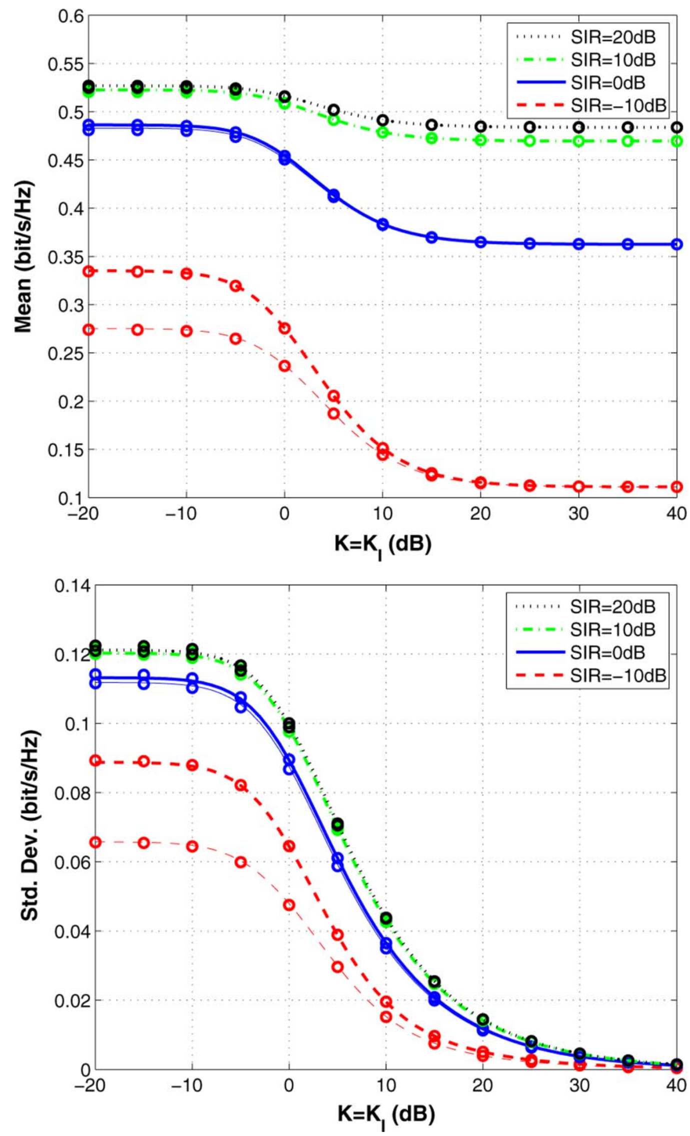

Fig. 4. Mean and standard deviation of $\mathcal{I}$ as a function of $K = K_I$ with $\mathsf{SNR} = -10$ dB.

To this end, we have to calculate the inverse of the 2×2 and 3×3 block matrices $\boldsymbol{M}|_{\mathcal{S}}$ and $\boldsymbol{K}|_{\mathcal{S}}$, respectively. The block inverse of $\boldsymbol{M}|_{\mathcal{S}}$ is obtained by using standard results from linear algebra [25, p. 18]. We get

$$\boldsymbol{M}^{-1}|_{\mathcal{S}}\delta\boldsymbol{M} = \begin{pmatrix} \delta\boldsymbol{Z}_{\boldsymbol{3}}^{T} \otimes \boldsymbol{A}_{\boldsymbol{M}}\boldsymbol{R}_{I} & \delta\boldsymbol{W}_{\boldsymbol{3}}^{T} \otimes \boldsymbol{B}_{\boldsymbol{M}}\tilde{\boldsymbol{T}}_{I} \\ \delta\boldsymbol{Z}_{\boldsymbol{3}}^{T} \otimes \boldsymbol{C}_{\boldsymbol{M}}\boldsymbol{R}_{I} & \delta\boldsymbol{W}_{\boldsymbol{3}}^{T} \otimes \boldsymbol{D}_{\boldsymbol{M}}\tilde{\boldsymbol{T}}_{I} \end{pmatrix}$$

and

$$\boldsymbol{M}^{-1}|_{\mathcal{S}}\delta\boldsymbol{\Gamma} = \begin{pmatrix} \delta\boldsymbol{\Theta}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{A}_{\boldsymbol{M}}\boldsymbol{R}_{I} & 0 & \delta\boldsymbol{\Theta}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{B}_{\boldsymbol{M}}\tilde{\boldsymbol{T}}_{I} \\ \delta\boldsymbol{\Theta}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{C}_{\boldsymbol{M}}\boldsymbol{R}_{I} & 0 & \delta\boldsymbol{\Theta}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{D}_{\boldsymbol{M}}\tilde{\boldsymbol{T}}_{I} \end{pmatrix}$$

with the matrices $\boldsymbol{A}_{\boldsymbol{M}}, \boldsymbol{B}_{\boldsymbol{M}}, \boldsymbol{C}_{\boldsymbol{M}}, \boldsymbol{D}_{\boldsymbol{M}}$ defined in (14). Next, we define the matrices $\hat{\boldsymbol{R}}$, $\hat{\boldsymbol{T}}$, and $\hat{\boldsymbol{I}}$ according to (16) and calculate the block inverse of $\boldsymbol{K}|_{\mathcal{S}}$. This can be done by iterating the 2×2 block matrix inversion formula [25, p. 18], which eventually leads to (40), shown at the bottom of the page, and to

$$\boldsymbol{K}^{-1}|_{\mathcal{S}}\delta\bar{\boldsymbol{\Gamma}} = \begin{pmatrix} \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{A}_{\boldsymbol{K}}\boldsymbol{R}_{I} & \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{C}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}_{I} \\ \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{D}_{\boldsymbol{K}}\boldsymbol{R}_{I} & \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{F}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}_{I} \\ \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{G}_{\boldsymbol{K}}\boldsymbol{R}_{I} & \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{I}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}_{I} \end{pmatrix}$$

where the block matrices $\boldsymbol{A}_{\boldsymbol{K}}, \ldots, \boldsymbol{I}_{\boldsymbol{K}}$ are defined in (15).

Now, we can evaluate the terms F_k defined in (39) as follows:

$$\boldsymbol{K}^{-1}|_{\mathcal{S}}\delta\boldsymbol{K} = \begin{pmatrix} \delta\boldsymbol{Z}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{A}_{\boldsymbol{K}}\boldsymbol{R} + \delta\boldsymbol{Z}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{A}_{\boldsymbol{K}}\boldsymbol{R}_{I} & \delta\boldsymbol{W}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{B}_{\boldsymbol{K}}\tilde{\boldsymbol{T}} & \delta\boldsymbol{W}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{C}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}_{I} \\ \delta\boldsymbol{Z}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{D}_{\boldsymbol{K}}\boldsymbol{R} + \delta\boldsymbol{Z}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{D}_{\boldsymbol{K}}\boldsymbol{R}_{I} & \delta\boldsymbol{W}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{E}_{\boldsymbol{K}}\tilde{\boldsymbol{T}} & \delta\boldsymbol{W}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{F}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}_{I} \\ \delta\boldsymbol{Z}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{G}_{\boldsymbol{K}}\boldsymbol{R} + \delta\boldsymbol{Z}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{G}_{\boldsymbol{K}}\boldsymbol{R}_{I} & \delta\boldsymbol{W}_{\boldsymbol{1}}^{T} \otimes \boldsymbol{H}_{\boldsymbol{K}}\tilde{\boldsymbol{T}} & \delta\boldsymbol{W}_{\boldsymbol{2}}^{T} \otimes \boldsymbol{I}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}_{I} \end{pmatrix} \tag{40}$$

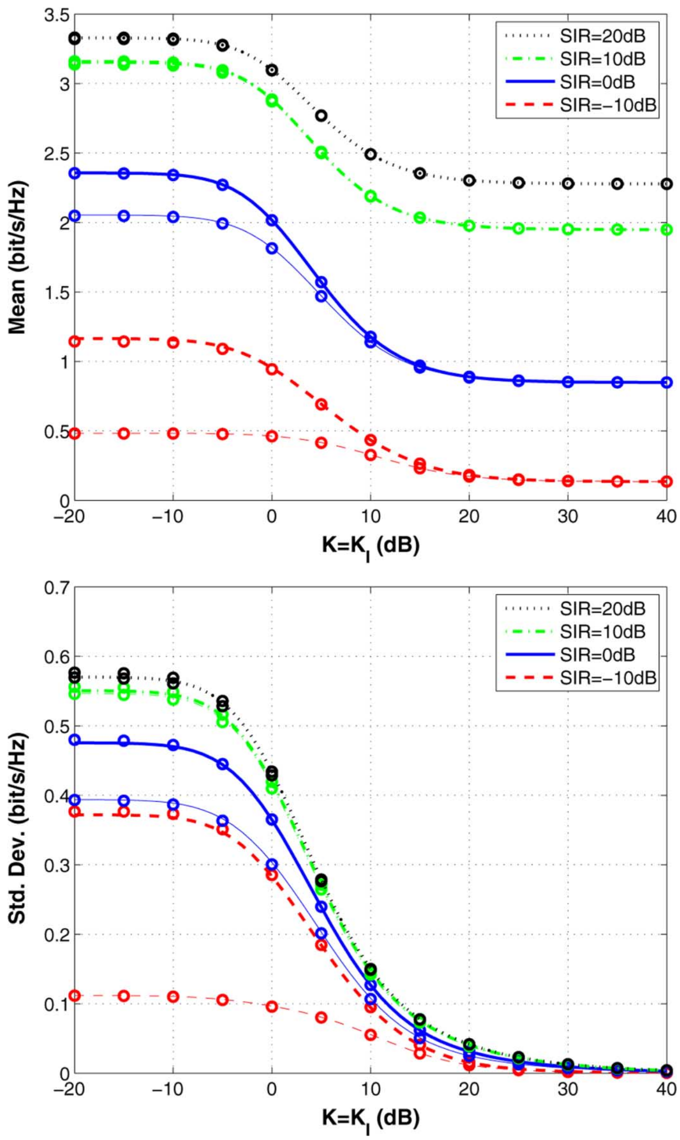

Fig. 5. Same as Fig. 4 but $\mathsf{SNR} = 0$ dB.

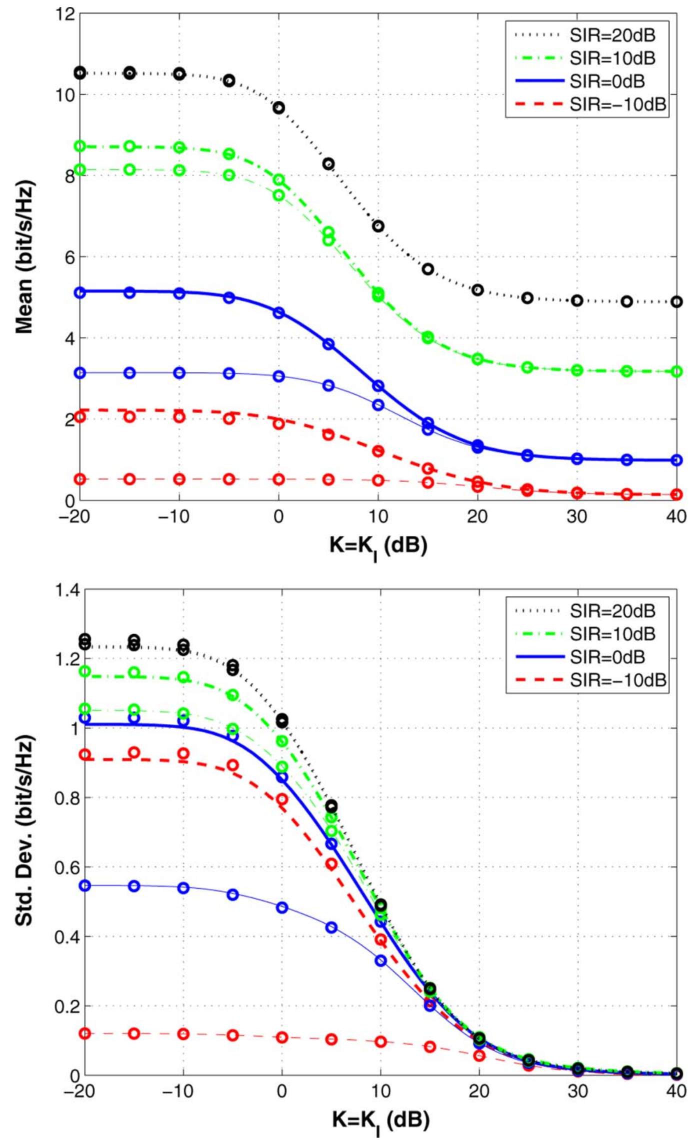

Fig. 6. Same as Fig. 4 but $\mathsf{SNR} = 10$ dB.

1) The constant term F_0 in the saddlepoint approximation is given by

$$\begin{aligned}F_0 &= -\ln \mathrm{sdet}(\mathfrak{X})|_{\mathcal{S}} - a(w_1z_1 + w_2z_2 - w_3z_3)\\ &= -\ln\det(\boldsymbol{M})|_{\mathcal{S}} + \ln\det(\boldsymbol{K})|_{\mathcal{S}}\\ &\quad - a(w_1z_1 + w_2z_2 - w_3z_3)\\ &= a(\ln\det(\boldsymbol{K_0}) - w_1z_1 - w_2z_2)\\ &\quad - a(\ln\det(\boldsymbol{M_0}) - w_3z_3))\end{aligned} \tag{41}$$

where the matrices $\boldsymbol{M_0}$ and $\boldsymbol{K_0}$ are defined in (11) and (12).

2) The first-order term F_1 in the saddlepoint approximation is given by (42).

$$\begin{aligned}F_1 &= -\mathrm{tr}(\boldsymbol{M}^{-1}|_{\mathcal{S}}\delta\boldsymbol{M}) + \mathrm{tr}(\boldsymbol{K}^{-1}|_{\mathcal{S}}\delta\boldsymbol{K})\\ &\quad - z_1\delta\boldsymbol{W_1} - w_1\delta\boldsymbol{Z_1} - z_2\delta\boldsymbol{W_2}\\ &\quad - w_2\delta\boldsymbol{Z_2} + z_3\delta\boldsymbol{W_3} + w_3\delta\boldsymbol{Z_3}\\ &= (\mathrm{tr}(\boldsymbol{E_K}\tilde{\boldsymbol{T}}) - z_1)\delta\boldsymbol{W_1} + (\mathrm{tr}(\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I) - z_2)\\ &\quad \times \delta\boldsymbol{W_2} - (\mathrm{tr}(\boldsymbol{D_M}\tilde{\boldsymbol{T}}_I) - z_3)\delta\boldsymbol{W_3}\\ &\quad + (\mathrm{tr}(\boldsymbol{A_K}\boldsymbol{R}) - w_1)\delta\boldsymbol{Z_1} + (\mathrm{tr}(\boldsymbol{A_K}\boldsymbol{R}_I) - w_2)\\ &\quad \times \delta\boldsymbol{Z_2} - (\mathrm{tr}(\boldsymbol{A_M}\boldsymbol{R}_I) - w_3)\delta\boldsymbol{Z_3}.\end{aligned} \tag{42}$$

Setting $F_1 = 0$ gives the two systems of coupled fixed-point equations in (17).

3) The second-order term F_2 in the saddlepoint approximation is given by

$$\begin{aligned}F_2 &= \frac{1}{2}\mathrm{tr}((\boldsymbol{M}^{-1}|_{\mathcal{S}}\delta\boldsymbol{M})^2) - \frac{1}{2}\mathrm{tr}((\boldsymbol{K}^{-1}|_{\mathcal{S}}\delta\boldsymbol{K})^2)\\ &\quad + \mathrm{tr}((\boldsymbol{K}^{-1}|_{\mathcal{S}}\delta\bar{\boldsymbol{\Gamma}})(\boldsymbol{M}^{-1}|_{\mathcal{S}}\delta\boldsymbol{\Gamma}))\\ &\quad - \mathrm{tr}(\delta\boldsymbol{Z_1}\delta\boldsymbol{W_1} + \delta\boldsymbol{Z_2}\delta\boldsymbol{W_2} - \delta\boldsymbol{Z_3}\delta\boldsymbol{W_3}\\ &\qquad + \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{1}}\delta\boldsymbol{\Theta_1} + \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{2}}\delta\boldsymbol{\Theta_2}).\end{aligned} \tag{43}$$

Defining the matrices $\boldsymbol{M_2}$, $\boldsymbol{K_2}$, and $\boldsymbol{G_2}$ as in (13) and the vectors

$$\begin{aligned}\boldsymbol{x} &\triangleq (\delta\boldsymbol{Z_3}, \delta\boldsymbol{W_3})^T\\ \boldsymbol{y} &\triangleq (\delta\boldsymbol{Z_1}, \delta\boldsymbol{W_1}, \delta\boldsymbol{Z_2}, \delta\boldsymbol{W_2})^T\\ \boldsymbol{\theta} &\triangleq (\delta\boldsymbol{\Theta_1}, \delta\boldsymbol{\Theta_2})^T\\ \bar{\boldsymbol{\theta}} &\triangleq (\delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{1}}, \delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{2}})^T\end{aligned}$$

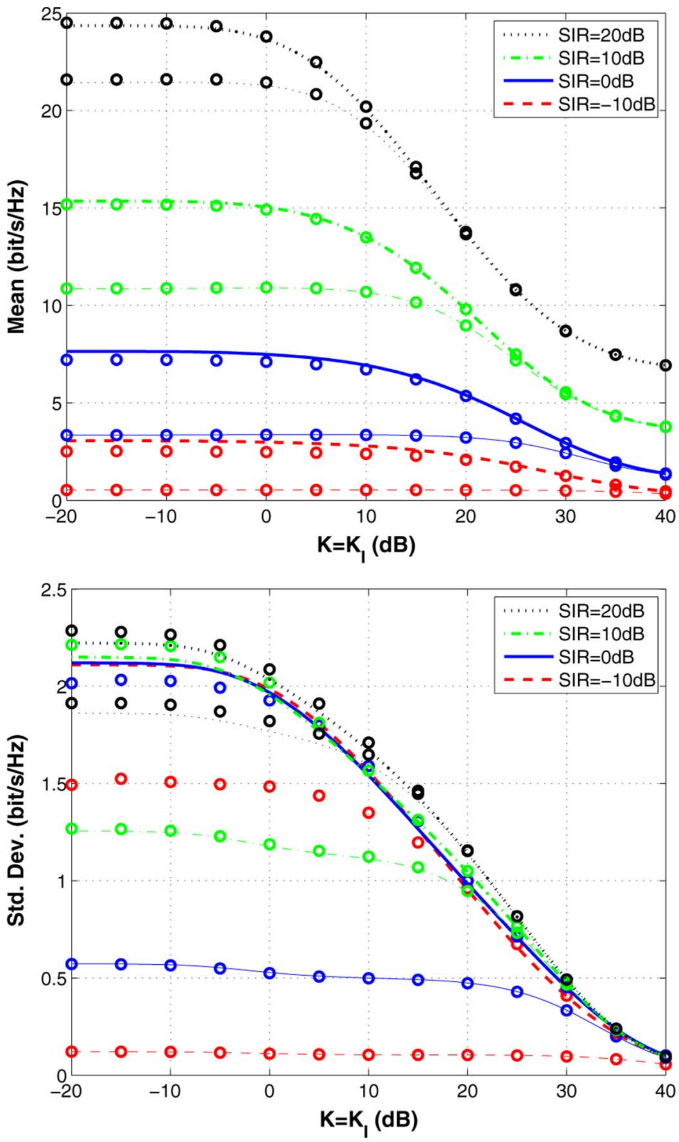

Fig. 7. Same as Fig. 4 but $\mathsf{SNR} = 30$ dB.

equation (43) can be written as

$$F_2 = \frac{1}{2}\mathrm{tr}(\boldsymbol{x}^T(\boldsymbol{M_2} \otimes \boldsymbol{I}_a)\boldsymbol{x}) + \frac{1}{2}\mathrm{tr}(\boldsymbol{y}^T(\boldsymbol{K_2} \otimes \boldsymbol{I}_a)\boldsymbol{y}) - \mathrm{tr}(\bar{\boldsymbol{\theta}}^T(\boldsymbol{G_2} \otimes \boldsymbol{I}_a)\boldsymbol{\theta}). \tag{44}$$

4) Higher-order terms (F_k with $k > 2$) in the saddlepoint approximation tend to zero as $n_R \to \infty$. This result is proved in Appendix F by using the asymptotic setup defined in Section III-B.

We can summarize the results obtained as

$$\begin{aligned} G(a) &\sim \exp(-F_0) \int D_g(\boldsymbol{\Theta}, \bar{\boldsymbol{\Theta}}) \exp(\mathrm{tr}(\bar{\boldsymbol{\theta}}^T(\boldsymbol{G_2} \otimes \boldsymbol{I}_a)\boldsymbol{\theta})) \\ &\quad \cdot \int d\mu(\delta\boldsymbol{W_1}, \delta\boldsymbol{Z_1}) \int d\mu(\delta\boldsymbol{W_2}, \delta\boldsymbol{Z_2}) \int d\mu(\delta\boldsymbol{W_3}, \delta\boldsymbol{Z_3}) \\ &\quad \cdot \exp\left(-\frac{1}{2}\mathrm{tr}(\boldsymbol{x}^T(\boldsymbol{M_2} \otimes \boldsymbol{I}_a)\boldsymbol{x} + \boldsymbol{y}^T(\boldsymbol{K_2} \otimes \boldsymbol{I}_a)\boldsymbol{y})\right) \\ &= \exp(-F_0)\det(\boldsymbol{G_2})^{2a} \int d\mu(\delta\boldsymbol{W_1}, \delta\boldsymbol{Z_1}) \\ &\quad \cdot \int d\mu(\delta\boldsymbol{W_2}, \delta\boldsymbol{Z_2})\mathrm{etr}\left(-\frac{1}{2}\boldsymbol{y}^T(\boldsymbol{K_2} \otimes \boldsymbol{I}_a)\boldsymbol{y}\right) \\ &\quad \cdot \int d\mu(\delta\boldsymbol{W_3}, \delta\boldsymbol{Z_3})\mathrm{etr}\left(-\frac{1}{2}\boldsymbol{x}^T(\boldsymbol{M_2} \otimes \boldsymbol{I}_a)\boldsymbol{x}\right) \end{aligned}$$

where we used Identity (B.1) with $\boldsymbol{A} = \boldsymbol{I}_{2a}$ and $\boldsymbol{B} = \boldsymbol{G_2} \otimes I_a$. The integrals can be evaluated as

$$\begin{aligned} G(a) &\sim \exp\{-a[(\ln\det(\boldsymbol{K_0}) - w_1 z_1 - w_2 z_2) \\ &\quad -(\ln\det(\boldsymbol{M_0}) - w_3 z_3)]\} \left\{\frac{\det(\boldsymbol{G_2})^2}{-\det(\boldsymbol{M_2})\det(\boldsymbol{K_2})}\right\}^{\frac{a^2}{2}}. \end{aligned} \tag{45}$$

Finally, the logarithm of (45) gives the result reported in (10).

VII. Numerical Results

In this section, we present some numerical examples to assess the accuracy of the asymptotic analysis provided above. Moreover, the numerical results are aimed at describing the influence of the channel parameters on the mutual information statistics.

Following the model proposed in [28], the transmit and receive correlation matrices are assumed to be *exponential*, i.e., their elements are obtained by raising a common base to the power corresponding to the absolute value of the difference of their indexes in the matrix. The key advantage of this assumption is simplicity: the whole matrix depends on a single parameter. Thus, the correlation matrices are defined as

$$\begin{aligned} (\boldsymbol{R})_{ij} &= \alpha_R^{|i-j|} & (\boldsymbol{R}_I)_{ij} &= \alpha_{R_I}^{|i-j|} \\ (\boldsymbol{T})_{ij} &= \alpha_T^{|i-j|} & (\boldsymbol{T}_I)_{ij} &= \alpha_{T_I}^{|i-j|}. \end{aligned}$$

The average channel matrices $\bar{\boldsymbol{H}}$ and $\bar{\boldsymbol{H}}_I$ are assumed to have all constant entries.

Following these assumptions, we define the following set of parameters:

1) The number of antennas is equal to $n_T = n_R = n_I = 4$.
2) The Rician factors are $K = K_I = 10$ dB.
3) $\alpha_R = \alpha_T = \alpha_{R_I} = \alpha_{T_I} = 0$, i.e., there is no spatial correlation.
4) The covariance matrix of the (interfering) signal is proportional to the identity matrix.

A certain parameter in this set will be reintroduced in a simulation if it differs from this definition.

In all figures, thick lines refer to the mutual information defined in (6), accounting for the real statistics of multiple-access interference, and thin lines refer to the mutual information defined in (7), where interference is considered as additive Gaussian noise. The markers '+' and '∘" report the corresponding Monte Carlo simulation results with the goal of assessing the accuracy of the asymptotic approach when the number of antennas is small.

A. Impact of Interference

Figs. 1 to 3 show the mean and the standard deviation of the mutual information for $\mathsf{SNR} = -10, 0,$ and 10 dB, respectively, for $n_T = n_R = n_I \in \{4, 8\}$ as the SIR ranges from -20 to 40 dB. The figures show that the Gaussian interference assumption corresponds to lower values of mutual information, as already noticed, e.g., in [7] and [8]. This effect is emphasized as the SIR gets lower, whereas it has more limited impact for larger

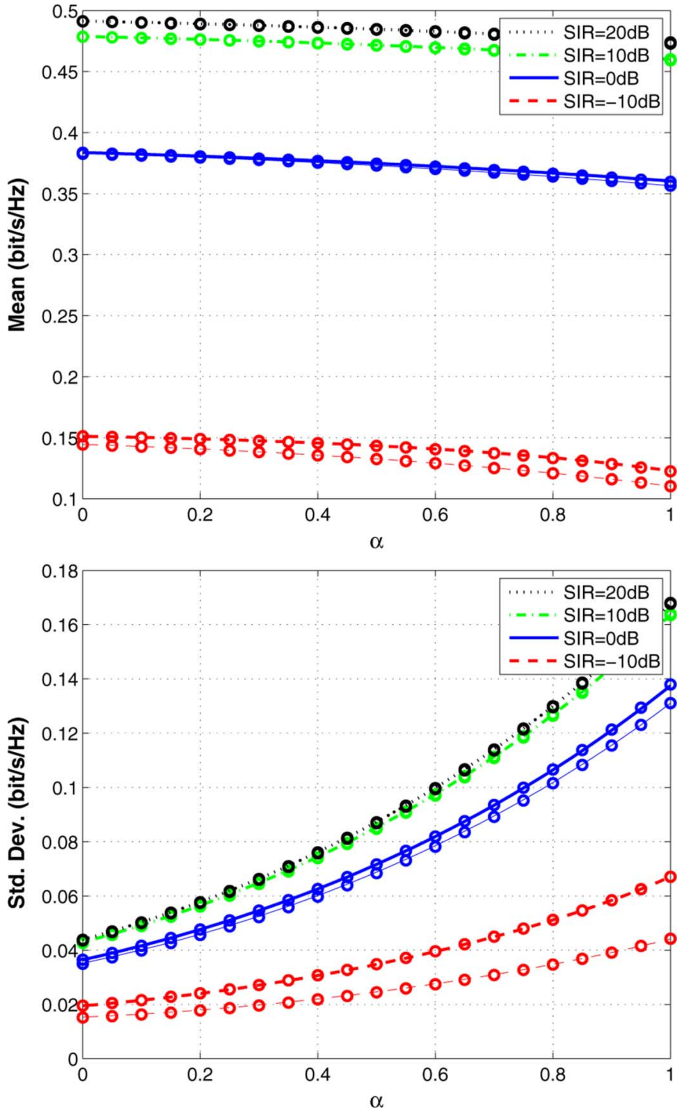

Fig. 8. Mean and standard deviation of random mutual information as a function of α with SNR $= -10$ dB.

SIR. Moreover, the effect is more visible at higher SNR for a fixed SIR, as expected. Similar effects are noticed for the mutual information standard deviation.

B. Impact of Rician Factor

Figs. 4 to 7 report the mean and the standard deviation of the mutual information for SNR $= -10, 0,$ and 10 dB, respectively, for SIR $\in \{-10, 0, 10, 20\}$ dB. Here, the Rician factors range from $K = K_I = -20$ to $K = K_I = 40$ dB. It can be noticed that all the mean (or ergodic) mutual information plots decrease monotonically versus the Rician factor. The same effect is visible for the standard deviation plots.

The effect of considering interference as additive Gaussian noise is more noticeable when the Rician factor and the SIR are low and when the SNR is high. In other words, proper interference cancelation is more effective when there is a weak LOS component, when the useful signal power is high (high SNR), and when the interfering signal power is still higher (low SIR). Finally, it can be noticed that for very high SNR (30 dB) and low SIR (-10 dB) the accuracy of standard deviation is limited, as can be seen in Fig. 7. Similar effects have also been reported in [45]. This can be explained by noting that, in this low-SIR regime, the mutual information is obtained as the difference between two large numerical values. Thus, we believe that the results are affected by significant cancelation errors that amplify the difference between the asymptotic and the simulation results.

C. Impact of Spatial Correlation

Figs. 8 to 10 show the mean and the standard deviation of the mutual information for SNR $= -10, 0,$ and 10 dB, respectively, for SIR $\in \{-10, 0, 10, 20\}$ dB. Spatial correlation is accounted for by the common exponential basis of all the spatial correlation matrices, namely, $\alpha_R = \alpha_T = \alpha_{R_I} = \alpha_{T_I} = \alpha \in (0, 1)$. We can see that the effect of increasing spatial correlation decreases the mean (ergodic) mutual information. The reduction is more sensible at higher SNR. Notice that we assumed iid power allocation, i.e., $\boldsymbol{Q}$ is proportional to the identity matrix. In [45] it is shown that optimizing $\boldsymbol{Q}$ (in order to achieve the ergodic capacity) leads to different results at low SIR ratio, which confirms the fact that spatial correlation can sometimes increase capacity. As far as concerns the standard deviation, it can be noticed that there are an SNR and a SIR threshold after which the standard deviation starts to decrease. Indeed, at SNR $= -10$ dB the standard deviation increases with α for all values of SIR. For SNR $= 0$ dB the curve for SIR $= -10$ dB starts decreasing at $\alpha = 0.8$. For SNR $= 10$ dB this effect is still more pronounced.

D. Impact of Antenna Ratio

Figs. 11 to 13 report the mean and the standard deviation of the mutual information for SNR $= -10, 0,$ and 10 dB, respectively, for SIR $\in \{-10, 0, 10, 20\}$ dB. In this case, the number of receive and interfering antennas are kept fixed to $n_R = n_I = 2$, while $n_T \in \{n \in \mathbb{N} : 1 \le n \le 10\}$ ($\mathbb{N}$ denotes the set of natural numbers). For low SNR, the mean (ergodic) mutual information is almost constant while the standard deviation decreases sharply with n_T at all values of the SIR. As the SNR increases, the mean (ergodic) mutual information exhibits a slightly higher rate of increase but remains limited by the low n_R assumption. Finally, it can be noticed that, when the number of antennas is low ($n_T < 4$ and $n_R = n_I = 2$), the accuracy of the asymptotic approximation decreases slightly at high SNR and low SIR.

To sum up, the numerical results we presented show that the asymptotic approximation provides very accurate results in broad range of channel settings even with a small number of antennas. As expected, treating co-channel interference as Gaussian noise reduces the mutual information. This effect is most visible at low SIR and Rician factor and, at fixed SIR, for high SNR.

E. Asymptotic Gaussianity of $\mathcal{I}$

Fig. 14 shows the cumulative density function (CDF) of the mutual information $\mathcal{I}$ given by (6) (thick lines) and $\mathcal{I}_G$ given by (7) (thin lines) for $n_T = n_R = n_I \in \{4, 8\}$ and SNR $= 10$ dB. The figure below corresponds to SIR $= -10$ dB and the figure above to SIR $= 0$ dB. Dashed blue curves are obtained

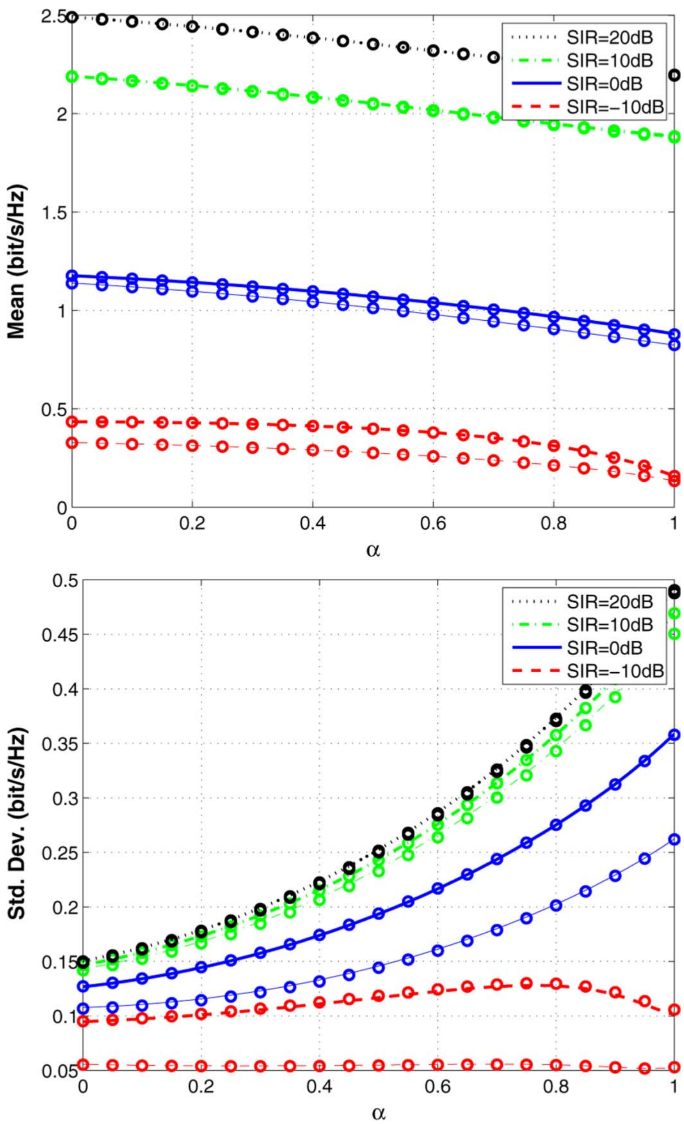

Fig. 9. Same as Fig. 8 but $\mathsf{SNR} = 0$ dB.

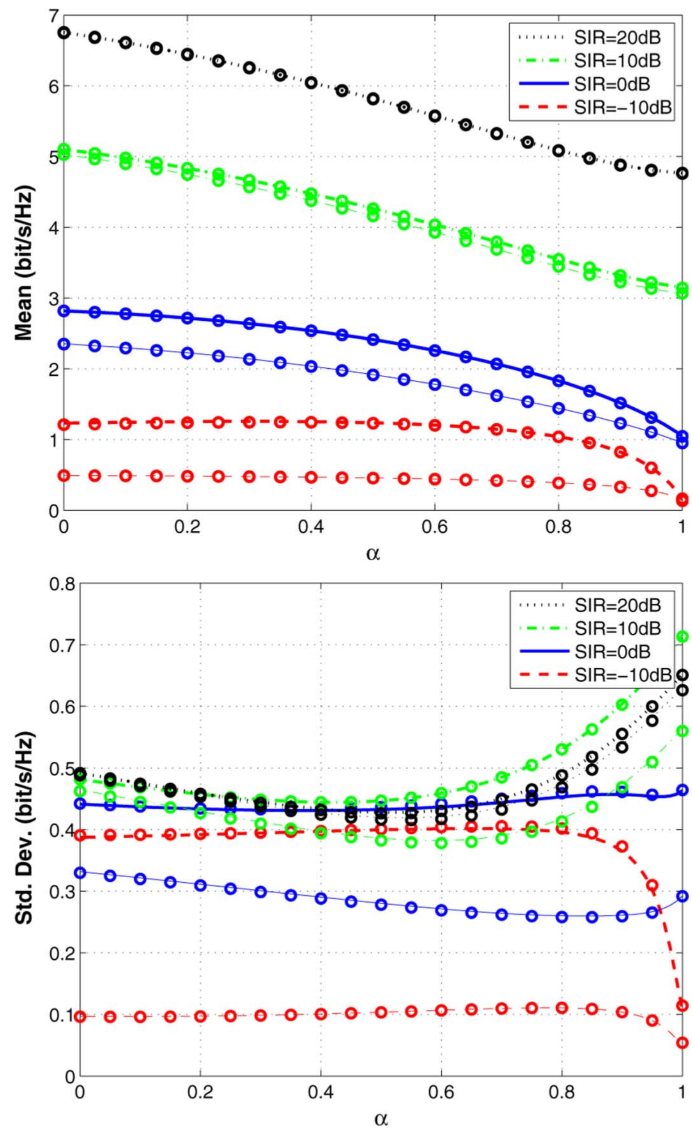

Fig. 10. Same as Fig. 8 but $\mathsf{SNR} = 10$ dB.

numerically by Monte Carlo simulations and show an empirical CDF plot. Solid red curves report the asymptotic CDF, which is determined by the asymptotic mean and variance calculated analytically with the method proposed in this paper. The curves show that there is a rather good agreement between simulation and asymptotic analytic results, supporting the applicability of the asymptotic analysis when the number of antennas is small.

VIII. Conclusion

In this paper, we addressed the problem of evaluating the probability distribution of the mutual information over a separately correlated Rician fading MIMO channel affected by the presence of interference. Previous works in the literature addressed the evaluation of either the mean and variance of the mutual information without interference [43], or only the mean (ergodic) mutual information with interference and Rician fading [44]. As a matter of fact, the ergodic mutual information is a representative figure for ergodic channels but it is insufficient to properly describe the capacity versus outage tradeoff in nonergodic systems. Second-order statistics are needed to find the outage probability or the outage capacity, for whose derivation the knowledge of the ergodic mutual information is not sufficient.

This work complements earlier results by Moustakas *et al.* [29], dealing with the case of Rayleigh fading. These results paved the way to further developments by focusing on a more specific (Rayleigh fading) channel model which does not encompass many practical situations occurring with MIMO systems. The main extension of the present analysis with respect to [29] stands in the consideration of mixed-valued matrices containing both Grassmann and complex variables. This type of matrices, introduced in the context of theoretical physics, are used to accomplish the derivation of the mean and variance of the mutual information in the presence of multiple-access interference for the Kronecker model with Rician fading. In order to deal with these mixed-valued matrices we resorted to some powerful tools developed in the contest of theoretical physics, namely: the *replica method* and *superanalysis*.

Therefore, our main result is the derivation of the asymptotic second-order mutual information statistics for a separately correlated Rician fading MIMO channel in the presence of

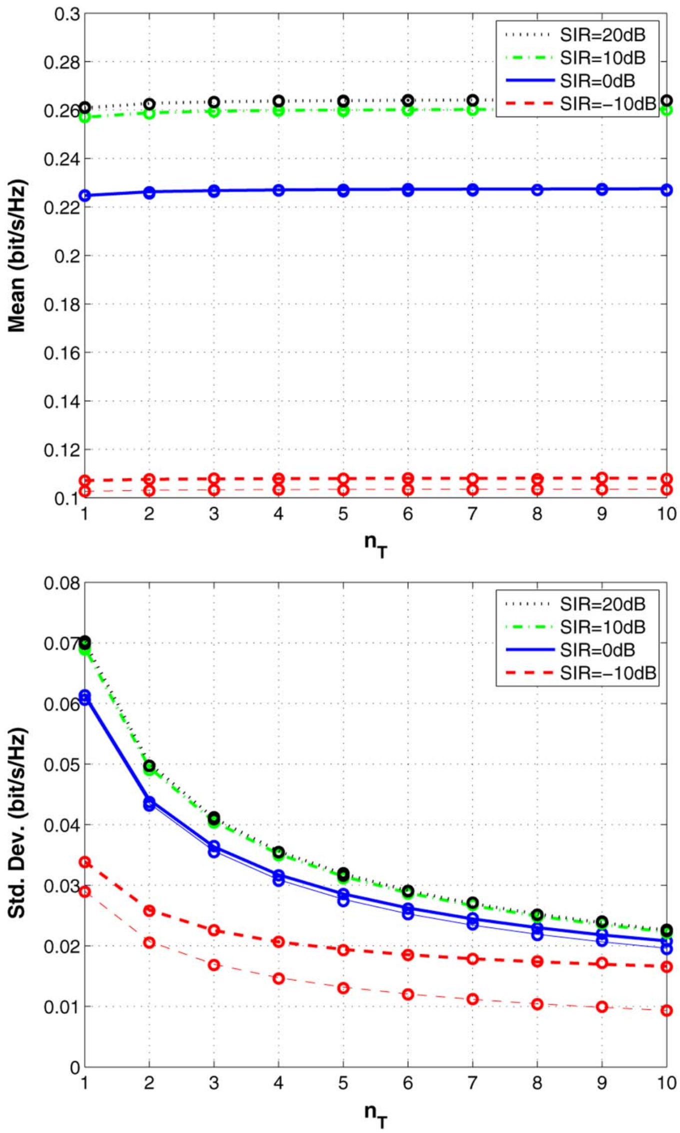

Fig. 11. Mean and standard deviation of mutual information as a function of n_T with $\mathsf{SNR} = -10$ dB.

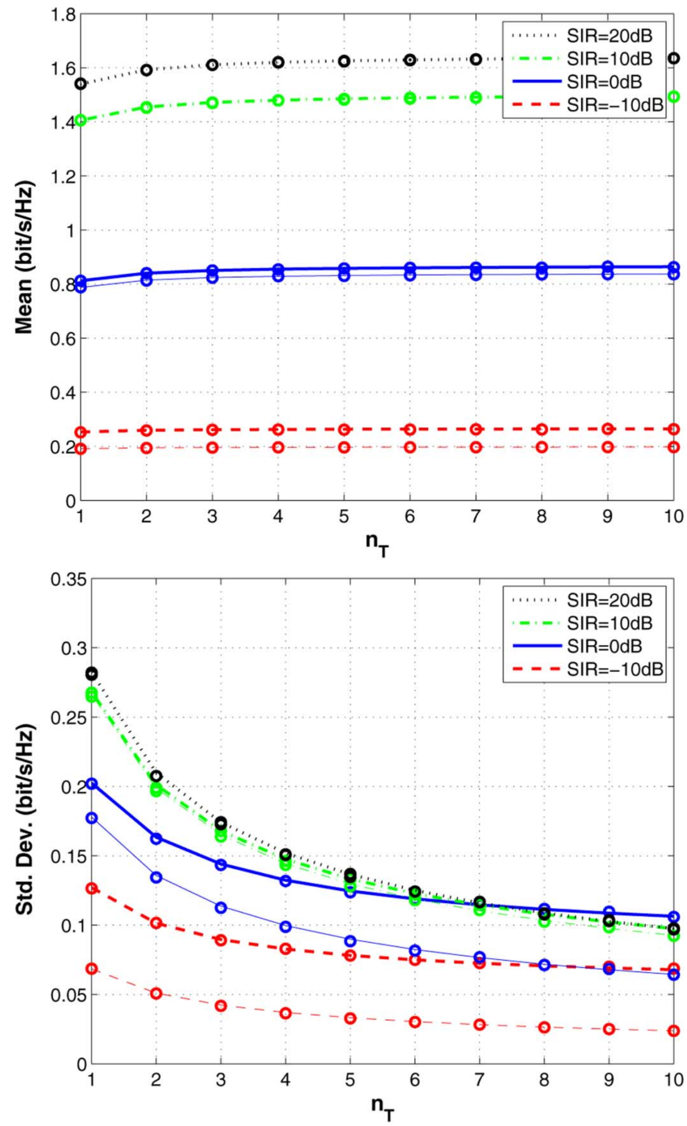

Fig. 12. Same as Fig. 11 but $\mathsf{SNR} = 0$ dB.

co-channel interference. Our analysis shows that the mutual information probability distribution approaches the Gaussian distribution asymptotically (as the number of transmit, interfering, and receive antennas grow large, with their ratios approaching finite constants).

From an applicative standpoint, our analytic results turn out to be useful in the analysis of MIMO systems with a moderately large number of antennas. In fact, we showed by Monte Carlo simulation that the difference between the asymptotic analytic results and their numerically obtained counterparts are limited in the case of as few as 4 antennas. Similarly, we compared outage probability curves obtained by asymptotic analysis and by Monte Carlo simulation, and this comparison showed that a good agreement exists also in this case but with a slightly larger number of antennas.

Since in this paper we focused on the mutual information, further developments are called for in order to obtain approximations of the outage capacity.

Appendix A
Integral Identities for Complex Matrices

Identity A.1: (Completing the Square) Let $\boldsymbol{A}$, $\boldsymbol{B}$ be a complex, Hermitian, positive definite $m \times m$ and $n \times n$ matrix, respectively, and $\boldsymbol{C}$, $\boldsymbol{D}$ be $n \times m$ complex matrices. Then

$$\int D_c\boldsymbol{U}\mathrm{etr}(-\pi(\boldsymbol{A}\boldsymbol{U}^H\boldsymbol{B}\boldsymbol{U} + \boldsymbol{C}^H\boldsymbol{U} + \boldsymbol{U}^H\boldsymbol{D})) = \det(\boldsymbol{A})^{-n}\det(\boldsymbol{B})^{-m}\mathrm{etr}(\pi\boldsymbol{A}^{-1}\boldsymbol{C}^H\boldsymbol{B}^{-1}\boldsymbol{D})$$

with integration over the domain $\boldsymbol{U} \in \mathbb{C}^{n\times m}$ and $D_c\boldsymbol{U} \triangleq \prod_{i=1}^{n}\prod_{j=1}^{m} d(\boldsymbol{U})_{ij}$.

Proof: see [43, Identity B.1] ■

Identity A.2: (Hubbard-Stratonovich Transformation) Let $\boldsymbol{A}, \boldsymbol{B}$ be complex $m \times m$ matrices. Then

$$\mathrm{etr}(-\boldsymbol{A}\boldsymbol{B}) = \int d\mu(\boldsymbol{W}, \boldsymbol{Z})\mathrm{etr}(\boldsymbol{Z}\boldsymbol{W} - \boldsymbol{A}\boldsymbol{W} - \boldsymbol{Z}\boldsymbol{B})$$

where $d\mu(\boldsymbol{W}, \boldsymbol{Z}) \triangleq \prod_{a,b=1}^{m}(\frac{1}{2\pi i}d(\boldsymbol{W})_{ab}d(\boldsymbol{Z})_{ab})$ and the contours $\mu(\boldsymbol{W}, \boldsymbol{Z})$ for the elements of $\boldsymbol{W}$ and $\boldsymbol{Z}$ are along the real and parallel to the imaginary axis, respectively.

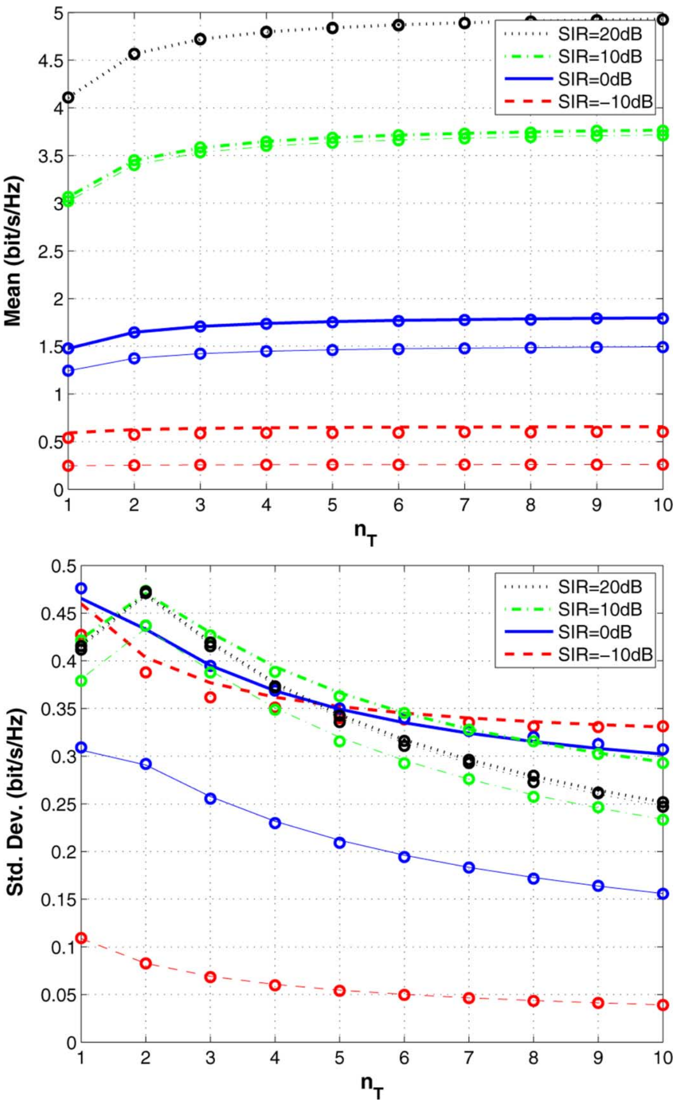

Fig. 13. Same as Fig. 11 but $\mathsf{SNR} = 10$ dB.

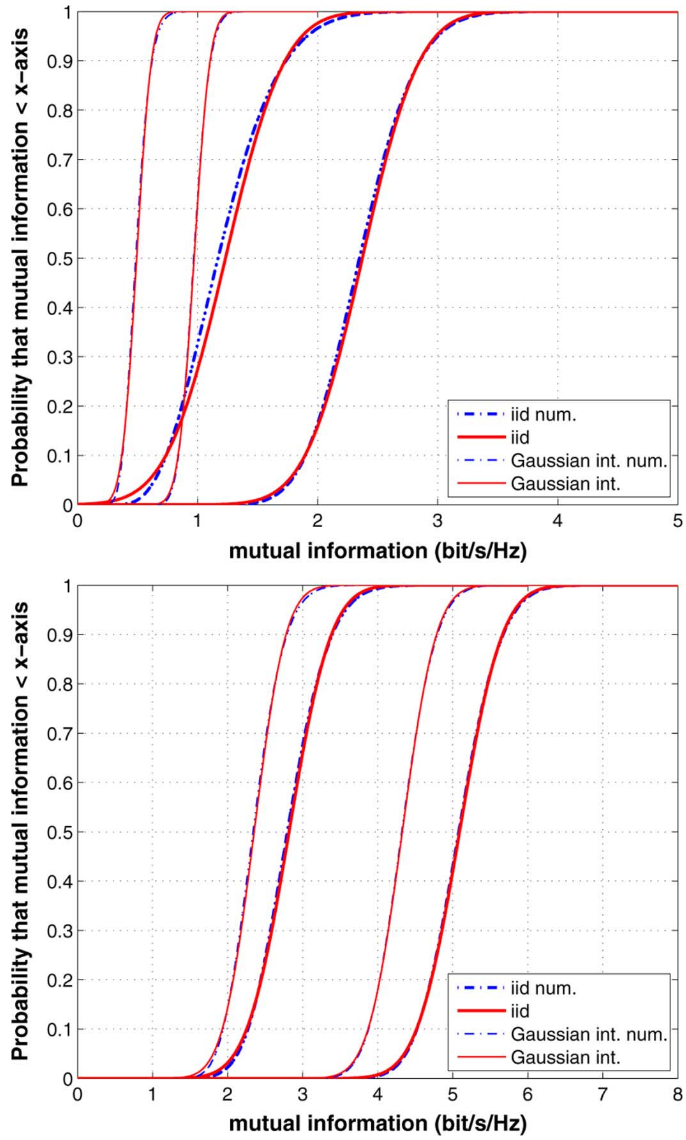

Fig. 14. CDF of random mutual information for $\mathsf{SIR} = -10$ dB (figure on the top) and $\mathsf{SIR} = 0$ dB (figure at the bottom). $n_T = n_R = n_I \in \{4, 8\}$.

Proof: This identity is obtained as the limit of a sequence of Gaussian integrals, see [43, Identity B.2]. ■

Appendix B
Integral Identities for Grassmann Valued Matrices

Identity B.1: (Completing the Square) Let $\boldsymbol{A}$, $\boldsymbol{B}$ be a complex, nonsingular $m \times m$ and $n \times n$ matrix, respectively, and $\boldsymbol{\Phi}$, $\bar{\boldsymbol{\Theta}}$ be Grassmann valued $n \times m$ and $m \times n$ matrices, respectively. Then

$$\int D_g(\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}) \mathrm{etr}(\boldsymbol{A}\bar{\boldsymbol{\Psi}}\boldsymbol{B}\boldsymbol{\Psi} + \bar{\boldsymbol{\Theta}}\boldsymbol{\Psi} + \bar{\boldsymbol{\Psi}}\boldsymbol{\Phi}) = \det(\boldsymbol{A})^n \det(\boldsymbol{B})^m \mathrm{etr}(-\boldsymbol{A}^{-1}\bar{\boldsymbol{\Theta}}\boldsymbol{B}^{-1}\boldsymbol{\Phi})$$

with integration over $\boldsymbol{\Psi} \in \mathcal{I}G^{n\times m}$ and $\bar{\boldsymbol{\Psi}} \in \mathbb{G}^{m\times n}$ and $D_g(\boldsymbol{\Psi}, \bar{\boldsymbol{\Psi}}) \triangleq \prod_{i=1}^{m}\prod_{j=1}^{n} d(\boldsymbol{\Psi})_{ji} d(\bar{\boldsymbol{\Psi}})_{ij}$.

Proof: see [29, Identity 5]. ■

Appendix C
Derivation of (18)

Setting $\boldsymbol{A} = \boldsymbol{I}_a, \boldsymbol{B} = \boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H$, and $\boldsymbol{C} = \boldsymbol{D} = \boldsymbol{0}$ in Identity (A.1), we get

$$\det\left(\boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\right)^{-a} = \int D_c\boldsymbol{U}\mathrm{etr}(-\pi(\boldsymbol{U}^H\boldsymbol{U} + \boldsymbol{U}^H\boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H\boldsymbol{U} + \boldsymbol{U}^H\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\boldsymbol{U})).$$

Using Identity (A.1) with $\boldsymbol{A} = \boldsymbol{I}_a$, $\boldsymbol{B} = \boldsymbol{I}_{n_T}$, $\boldsymbol{C}^H = i\boldsymbol{U}^H\boldsymbol{H}\boldsymbol{Q}^{1/2}$, and $\boldsymbol{D} = i\boldsymbol{Q}^{1/2}\boldsymbol{H}^H\boldsymbol{U}$ we further find that

$$\mathrm{etr}(-\pi\boldsymbol{U}^H\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\boldsymbol{U}) = \int D_c\boldsymbol{V}\mathrm{etr}(-\pi(\boldsymbol{V}^H\boldsymbol{V} + i\boldsymbol{U}^H\boldsymbol{H}\boldsymbol{Q}^{1/2}\boldsymbol{V} + i\boldsymbol{V}^H\boldsymbol{Q}^{1/2}\boldsymbol{H}^H\boldsymbol{U})).$$

Using Identity (A.1) with $\boldsymbol{A} = \boldsymbol{I}_a$, $\boldsymbol{B} = \boldsymbol{I}_{n_I}$, $\boldsymbol{C}^H = i\boldsymbol{U}^H\boldsymbol{H}_I\boldsymbol{Q}_I^{1/2}$, and $\boldsymbol{D} = i\boldsymbol{Q}_I^{1/2}\boldsymbol{H}_I^H\boldsymbol{U}$ yields

$$\begin{aligned}&\operatorname{etr}\left(-\pi\boldsymbol{U}^H\boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H\boldsymbol{U}\right)\\&\quad=\int D_c\boldsymbol{W}\operatorname{etr}\Big(-\pi(\boldsymbol{W}^H\boldsymbol{W}+i\boldsymbol{U}^H\boldsymbol{H}_I\boldsymbol{Q}_I^{1/2}\boldsymbol{W}\\&\qquad\qquad+i\boldsymbol{W}^H\boldsymbol{Q}_I^{1/2}\boldsymbol{H}_I^H\boldsymbol{U})\Big).\end{aligned}$$

Finally, inserting $\boldsymbol{H}$ and $\boldsymbol{H}_I$ from (1) yields the functions defined in (19)–(22).

Appendix D
Derivation of (23)

Setting $\boldsymbol{A} = \boldsymbol{I}_a, \boldsymbol{B} = \boldsymbol{I}_{n_R} + \boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H$ and $\bar{\boldsymbol{\Theta}} = \boldsymbol{\Phi} = \boldsymbol{0}$ in Identity (B.1) gives

$$\begin{aligned}&\det\left(\boldsymbol{I}_{n_R}+\boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H\right)^a\\&\qquad=\int D_g(\boldsymbol{\Psi},\bar{\boldsymbol{\Psi}})\operatorname{etr}(\bar{\boldsymbol{\Psi}}\boldsymbol{\Psi}+\bar{\boldsymbol{\Psi}}\boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H\boldsymbol{\Psi}).\end{aligned}$$

Using again Identity (B.1) with $\boldsymbol{A} = \boldsymbol{I}_a, \boldsymbol{B} = \boldsymbol{I}_{n_I}$, $\boldsymbol{\Phi} = \boldsymbol{Q}_I^{1/2}\boldsymbol{H}_I^H\boldsymbol{\Psi}$, and $-\bar{\boldsymbol{\Theta}} = \bar{\boldsymbol{\Psi}}\boldsymbol{H}_I\boldsymbol{Q}_I^{1/2}$ gives

$$\begin{aligned}&\operatorname{etr}\left(\bar{\boldsymbol{\Psi}}\boldsymbol{H}_I\boldsymbol{Q}_I\boldsymbol{H}_I^H\boldsymbol{\Psi}\right)\\&\quad=\int D_g(\boldsymbol{\Omega},\bar{\boldsymbol{\Omega}})\operatorname{etr}\left(\bar{\boldsymbol{\Omega}}\boldsymbol{\Omega}+\bar{\boldsymbol{\Omega}}\boldsymbol{Q}_I^{1/2}\boldsymbol{H}_I^H\boldsymbol{\Psi}-\bar{\boldsymbol{\Psi}}\boldsymbol{H}_I\boldsymbol{Q}_I^{1/2}\boldsymbol{\Omega}\right).\end{aligned}$$

Finally, inserting $\boldsymbol{H}_I$ from (1) yields the functions defined in (24)–(25).

Appendix E
Derivation of (33)

We define the following vectors of complex variables

$$\boldsymbol{u} \triangleq \operatorname{vec}\boldsymbol{U}\quad \boldsymbol{v} \triangleq \operatorname{vec}\boldsymbol{V}\quad \boldsymbol{w} \triangleq \operatorname{vec}\boldsymbol{W} \tag{46}$$

and Grassmann variables

$$\begin{cases}\boldsymbol{\psi} \triangleq \operatorname{vec}\boldsymbol{\Psi} & \bar{\boldsymbol{\psi}} \triangleq \operatorname{vec}\bar{\boldsymbol{\Psi}}^T\\ \boldsymbol{\omega} \triangleq \operatorname{vec}\boldsymbol{\Omega} & \bar{\boldsymbol{\omega}} \triangleq \operatorname{vec}\bar{\boldsymbol{\Omega}}^T.\end{cases} \tag{47}$$

First we integrate out the Grassmann variables in (47). Using vectorization the Grassmann integrals can be written as in (48),

$$\begin{aligned}&\int D_g(\boldsymbol{\psi},\bar{\boldsymbol{\psi}})\int D_g(\boldsymbol{\omega},\bar{\boldsymbol{\omega}})\\&\times\exp\left((\bar{\boldsymbol{\psi}}^T,\bar{\boldsymbol{\omega}}^T)\boldsymbol{M}\begin{pmatrix}\boldsymbol{\psi}\\\boldsymbol{\omega}\end{pmatrix}+\sqrt{\pi}(\bar{\boldsymbol{\psi}}^T,\bar{\boldsymbol{\omega}}^T)\boldsymbol{\Gamma}\begin{pmatrix}\boldsymbol{u}\\\boldsymbol{v}\\\boldsymbol{w}\end{pmatrix}\right.\\&\qquad\qquad\left.+\sqrt{\pi}(\boldsymbol{u}^H,\boldsymbol{v}^H,\boldsymbol{w}^H)\bar{\boldsymbol{\Gamma}}\begin{pmatrix}\boldsymbol{\psi}\\\boldsymbol{\omega}\end{pmatrix}\right)\\&\quad=\det(\boldsymbol{M})\exp\left(-\pi(\boldsymbol{u}^H,\boldsymbol{v}^H,\boldsymbol{w}^H)\bar{\boldsymbol{\Gamma}}\boldsymbol{M}^{-1}\boldsymbol{\Gamma}\begin{pmatrix}\boldsymbol{u}\\\boldsymbol{v}\\\boldsymbol{w}\end{pmatrix}\right)\end{aligned} \tag{48}$$

where we used Identity B.1 and defined $\boldsymbol{M}$, $\boldsymbol{\Gamma}$, and $\bar{\boldsymbol{\Gamma}}$ according to (34). Next we integrate out the complex variables. Using vectorization, the complex valued integrals can then be written as

$$\begin{aligned}&\int D_c\boldsymbol{u}\int D_c\boldsymbol{v}\int D_c\boldsymbol{w}\\&\quad\cdot\exp\left(-\pi(\boldsymbol{u}^H,\boldsymbol{v}^H,\boldsymbol{w}^H)(\boldsymbol{K}+\bar{\boldsymbol{\Gamma}}\boldsymbol{M}^{-1}\boldsymbol{\Gamma})\begin{pmatrix}\boldsymbol{u}\\\boldsymbol{v}\\\boldsymbol{w}\end{pmatrix}\right)\\&\qquad=\det(\boldsymbol{K}+\bar{\boldsymbol{\Gamma}}\boldsymbol{M}^{-1}\boldsymbol{\Gamma})^{-1}\end{aligned}$$

where we used Identity A.1 and the definition of the matrix $\boldsymbol{K}$ given by (34).

Appendix F
Asymptotic Vanishing of Higher-Order Terms

Recall from (39) that

$$F_k = \frac{(-1)^k}{k}\operatorname{str}((\boldsymbol{\mathfrak{X}}^{-1}|_{\mathcal{S}}\delta\boldsymbol{\mathfrak{X}})^k)$$

for $k > 2$. Let $\boldsymbol{S}$ be an arbitrary positive semi definite $n_R \times n_R$ matrix with the property that $\|\boldsymbol{S}\|_{sp} = O(1)$. Matrix $\boldsymbol{S}$ is a placeholder for $\boldsymbol{R}$ or $\boldsymbol{R}_I$, which share the same asymptotic properties. From the asymptotic setting defined in Section III-B we then get the bounds for the matrices from (15) given by (49).

$$\begin{aligned}\|\boldsymbol{A}_{\boldsymbol{K}}\boldsymbol{S}\|_{sp} &\le \|\boldsymbol{S}\|_{sp} = O(1)\\ \|\boldsymbol{I}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}_I\|_{sp} &= \Bigg\|\Big(\hat{\boldsymbol{I}}+\tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1/2}\\&\qquad\times(\boldsymbol{I}_{n_R}-\hat{\boldsymbol{R}}^{-1/2}\tilde{\boldsymbol{H}}(\hat{\boldsymbol{T}}+\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}})^{-1}\\&\qquad\times\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1/2})\hat{\boldsymbol{R}}^{-1/2}\tilde{\boldsymbol{H}}_I\Big)^{-1}\tilde{\boldsymbol{T}}_I\Bigg\|_{sp}\\&\le\|\hat{\boldsymbol{I}}^{-1}\|_{sp}\|\tilde{\boldsymbol{T}}_I\|_{sp}\\&\le\|\tilde{\boldsymbol{T}}_I\|_{sp} = O(1/n_R)\\ \|\boldsymbol{E}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}\|_{sp} &= \Bigg\|(\hat{\boldsymbol{T}}+\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1/2}\\&\qquad\times(\boldsymbol{I}_{n_R}-\hat{\boldsymbol{R}}^{-1/2}\tilde{\boldsymbol{H}}_I\left(\hat{\boldsymbol{I}}+\tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I\right)^{-1}\\&\qquad\times\tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1/2})\hat{\boldsymbol{R}}^{-1/2}\tilde{\boldsymbol{H}})^{-1}\tilde{\boldsymbol{T}}\Bigg\|_{sp}\\&\le\|\hat{\boldsymbol{T}}^{-1}\|_{sp}\|\tilde{\boldsymbol{T}}\|_{sp}\\&\le\|\tilde{\boldsymbol{T}}\|_{sp} = O(1/n_R)\end{aligned} \tag{49}$$

Here, we used the fact that $\boldsymbol{X}(\boldsymbol{I}+\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H \le \boldsymbol{I}$ for $\boldsymbol{X} \ge \boldsymbol{0}$. We further find that

$$\begin{aligned}\|\boldsymbol{H}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}\|_{sp} &= \left\|\left(\hat{\boldsymbol{I}}+\tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I\right)^{-1}\tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}\boldsymbol{E}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}\right\|_{sp}\\&\le\left\|\left(\hat{\boldsymbol{I}}+\tilde{\boldsymbol{H}}_I^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I\right)^{-1}\right\|_{sp}\\&\qquad\cdot\|\tilde{\boldsymbol{H}}_I\|_{sp}\|\hat{\boldsymbol{R}}^{-1}\|_{sp}\|\tilde{\boldsymbol{H}}\|_{sp}\|\boldsymbol{E}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}\|_{sp}\\&\le\|\tilde{\boldsymbol{H}}_I\|_{sp}\|\tilde{\boldsymbol{H}}\|_{sp}\|\boldsymbol{E}_{\boldsymbol{K}}\tilde{\boldsymbol{T}}\|_{sp}\\&=O(1/n_R)\\ \|B_{\boldsymbol{K}}\tilde{\boldsymbol{T}}\|_{sp} &= \|\hat{\boldsymbol{R}}^{-1}(\tilde{\boldsymbol{H}}\boldsymbol{E}_{\boldsymbol{K}}+\tilde{\boldsymbol{H}}_I\boldsymbol{H}_{\boldsymbol{K}})\tilde{\boldsymbol{T}}\|_{sp}\end{aligned}$$

$$
\begin{aligned}
&\leq \|\hat{\boldsymbol{R}}^{-1}\|_{sp}\|(\|\tilde{\boldsymbol{H}}\|_{sp}\|\boldsymbol{E_K}\tilde{\boldsymbol{T}}\|_{sp} \\
&\quad + \|\tilde{\boldsymbol{H}}_I\|\|\boldsymbol{H_K}\tilde{\boldsymbol{T}}\|_{sp}) \\
&\leq \|\tilde{\boldsymbol{H}}\|_{sp}\|\boldsymbol{E_K}\tilde{\boldsymbol{T}}\|_{sp} + \|\tilde{\boldsymbol{H}}_I\|\|\boldsymbol{H_K}\tilde{\boldsymbol{T}}\|_{sp} \\
&= O(1/n_R) \\
\|\boldsymbol{F_K}\tilde{\boldsymbol{T}}_I\|_{sp} &= \|(\hat{\boldsymbol{T}} + \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}})^{-1}\tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}}_I\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I\|_{sp} \\
&\leq \|(\hat{\boldsymbol{T}} + \tilde{\boldsymbol{H}}^H\hat{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{H}})^{-1}\|_{sp}\|\tilde{\boldsymbol{H}}\|_{sp} \\
&\quad \cdot \|\hat{\boldsymbol{R}}^{-1}\|_{sp}\|\tilde{\boldsymbol{H}}_I\|_{sp}\|\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I\|_{sp} \\
&\leq \|\tilde{\boldsymbol{H}}\|_{sp}\|\tilde{\boldsymbol{H}}_I\|_{sp}\|\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I\|_{sp} \\
&= O(1/n_R) \\
\|\boldsymbol{C_K}\tilde{\boldsymbol{T}}_I\|_{sp} &= \|\hat{\boldsymbol{R}}^{-1}(\tilde{\boldsymbol{H}}\boldsymbol{F_K} + \tilde{\boldsymbol{H}}_I\boldsymbol{I_K})\tilde{\boldsymbol{T}}_I\|_{sp} \\
&\leq \|\hat{\boldsymbol{R}}^{-1}\|_{sp}(\|\tilde{\boldsymbol{H}}\|_{sp}\|\boldsymbol{F_K}\tilde{\boldsymbol{T}}_I\|_{sp} \\
&\quad + \|\tilde{\boldsymbol{H}}_I\|_{sp}\|\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I\|_{sp}) \\
&\leq \|\tilde{\boldsymbol{H}}\|_{sp}\|\boldsymbol{F_K}\tilde{\boldsymbol{T}}_I\|_{sp} + \|\tilde{\boldsymbol{H}}_I\|_{sp}\|\boldsymbol{I_K}\tilde{\boldsymbol{T}}_I\|_{sp} \\
&= O(1/n_R) \\
\|\boldsymbol{D_K}\boldsymbol{S}\|_{sp} &= \|\hat{\boldsymbol{T}}^{-1}\tilde{\boldsymbol{H}}^H\boldsymbol{A_K}\boldsymbol{S}\|_{sp} \\
&\leq \|\hat{\boldsymbol{T}}^{-1}\|_{sp}\|\tilde{\boldsymbol{H}}\|_{sp}\|\boldsymbol{A_K}\boldsymbol{S}\|_{sp} \\
&\leq \|\tilde{\boldsymbol{H}}\|_{sp}\|\boldsymbol{A_K}\boldsymbol{S}\|_{sp} = O(1) \\
\|\boldsymbol{G_K}\boldsymbol{S}\|_{sp} &= \|\hat{\boldsymbol{I}}^{-1}\tilde{\boldsymbol{H}}_I^H\boldsymbol{A_K}\boldsymbol{S}\|_{sp} \\
&\leq \|\hat{\boldsymbol{I}}^{-1}\|_{sp}\|\tilde{\boldsymbol{H}}_I\|_{sp}\|\boldsymbol{A_K}\boldsymbol{S}\|_{sp} \\
&\leq \|\tilde{\boldsymbol{H}}_I\|_{sp}\|\boldsymbol{A_K}\boldsymbol{S}\|_{sp} = O(1).
\end{aligned}
$$

An almost identical calculation yields the following bounds for the matrices defined in (14):

$$
\begin{aligned}
\|\boldsymbol{A_M}\boldsymbol{S}\|_{sp} &= O(1) \quad \|\boldsymbol{B_M}\tilde{\boldsymbol{T}}_I\|_{sp} = O(1/n_R) \\
\|\boldsymbol{C_M}\boldsymbol{S}\|_{sp} &= O(1) \quad \|\boldsymbol{D_M}\tilde{\boldsymbol{T}}_I\|_{sp} = O(1/n_R).
\end{aligned}
$$

Changing variables

$$
\begin{aligned}
\delta\boldsymbol{Z_1} &\to \frac{1}{\sqrt{n_R}}\delta\boldsymbol{Z_1} \\
\delta\boldsymbol{Z_2} &\to \frac{1}{\sqrt{n_R}}\delta\boldsymbol{Z_2} \\
\delta\boldsymbol{Z_3} &\to \frac{1}{\sqrt{n_R}}\delta\boldsymbol{Z_3} \\
\delta\boldsymbol{W_1} &\to \sqrt{n_R}\delta\boldsymbol{W_1} \\
\delta\boldsymbol{W_2} &\to \sqrt{n_R}\delta\boldsymbol{W_2} \\
\delta\boldsymbol{W_3} &\to \sqrt{n_R}\delta\boldsymbol{W_3} \\
\delta\boldsymbol{\Theta_1} &\to \frac{1}{\sqrt{n_R}}\delta\boldsymbol{\Theta_1} \\
\delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{2}} &\to \frac{1}{\sqrt{n_R}}\delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{2}} \\
\delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{1}} &\to \sqrt{n_R}\delta\bar{\boldsymbol{\Theta}}_{\boldsymbol{1}} \\
\delta\boldsymbol{\Theta_2} &\to \sqrt{n_R}\delta\boldsymbol{\Theta_2}
\end{aligned}
$$

plainly leaves the measure in (35) invariant and has the effect that each matrix block in $\boldsymbol{\mathfrak{X}}^{-1}|_{\mathcal{S}}\delta\boldsymbol{\mathfrak{X}}$ is $O(1/\sqrt{n_R})$. Therefore, $\mathrm{str}((\boldsymbol{\mathfrak{X}}^{-1}|_{\mathcal{S}}\delta\boldsymbol{\mathfrak{X}})^k) = O(r^{1-k/2})$ which approaches 0 as $n_R \to \infty$ for all $k > 2$.

References

[1] A. Abdi and M. Kaveh, "A space-time correlation model for multi-element antenna systems in mobile fading channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 550–560, Apr. 2002.

[2] F. A. Berezin and A. A. Kirillov, *Introduction to Superanalysis*. Dordrecht, Netherlands: Reidel, 1987.

[3] R. S. Blum, J. H. Winters, and N. R. Sollenberger, "On the capacity of cellular systems with MIMO," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 242–244, Jun. 2002.

[4] R. S. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, Jun. 2003.

[5] R. W. Butler, *Saddlepoint Approximations With Applications*. Cambridge: Cambridge Series in Statistical and Probabilistic Mathematics, 2007.

[6] S. Catreux, P. F. Driessen, and L. J. Greenstein, "Simulation results for an interference-limited multiple-input multiple-output cellular system," *IEEE Commun. Lett.*, vol. 4, no. 11, pp. 334–336, Nov. 2000.

[7] M. Chiani, M. Z. Win, and H. Shin, "Capacity of MIMO systems in the presence of interference," in *IEEE GLOBECOM 2006*, San Francisco, CA, Nov. 27–Dec. 1 2006.

[8] M. Chiani, M. Z. Win, and H. Shin, MIMO Networks: The Effects of Interference arXiv:0802.0738v1.

[9] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[11] J. Dumont, P. Loubaton, S. Lasaulce, and M. Debbah, "On the asymptotic performance of MIMO correlated Ricean channels," in *Proc. IEEE ICASSP 2005*, Philadelphia, PA, Mar. 18–23, 2005.

[12] J. Dumont, P. Loubaton, and S. Lasaulce, "On the capacity achieving transmit covariance matrices of MIMO correlated Rician channels: A large system approach," in *Proc. IEEE GLOBECOM 2006*, San Francisco, CA, Nov. 27–Dec. 1 2006.

[13] J. Dumont, W. Hachem, P. Loubaton, and J. Najim, On the Capacity Achieving Covariance Matrix for Rician MIMO Channels: An Asymptotic Approach arXiv:0710.4051.

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

[15] K. H. Fischer and J. A. Hertz, *Spin Glasses*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[16] G. J. Foschini, "Layered space-time architecture wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, Autumn 1996.

[17] D. Gesbert, D. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

[18] V. L. Girko, *Theory of Stochastic Canonical Equations*. Boston, MA: Kluwer Academic, 2001.

[19] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[20] L. J. Greenstein, J. B. Andersen, H. L. Bertoni, S. Kozono, D. G. Michelson, and W. H. Tranter, "Channel and propagation models for wireless system design I and II," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3/6, Apr./Aug. 2002.

[21] 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.

[22] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *Ann. Appl. Probabil.*, vol. 17, no. 3, pp. 875–930, 2007.

[23] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for mutual information analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987–4004, Sep. 2008.

[24] D. Hösli, Y.-H. Kim, and A. Lapidoth, "Monotonicity results for coherent MIMO Rician channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4334–4339, Dec. 2005.

[25] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge University Press, 1985.

[26] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1175, Jul. 2004.

[27] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1102–1111, Mar. 2005.

[28] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sept. 2001.

[29] 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, pp. 2545–2561, Oct. 2003.

[30] A. L. Moustakas and S. H. Simon, "Random matrix theory of multi-antenna communications: The Ricean channel," *J. Phys. A: Math. Gen.*, vol. 38, no. 49, pp. 10859–10872, Dec. 2005.

[31] A. L. Moustakas and S. H. Simon, "On the outage capacity of correlated multiple-path MIMO channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3887–3903, Nov. 2007.

[32] R. 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.

[33] H. Nishimori*, Statistical Physics of Spin Glasses and Information Processing: An Introduction*. Oxford , U.K.: Oxford Univ. Press, 2001.

[34] E. Riegler and G. Taricco, "Second-order statistics of the mutual information of the asymptotic separately correlated Rician fading MIMO channel with interference," in *Proc. IEEE GLOBECOM 2007*, Wash., DC, Nov. 26–30, 2007.

[35] E. Riegler and G. Taricco, "On the ergodic capacity region of the separately correlated Rician fading multiple access MIMO channel," in *Proc. IEEE GLOBECOM 2007*, Wash., DC, Nov. 26–30, 2007.

[36] E. Riegler and G. Taricco, "Asymptotic ergodic capacity region and rate optimization of a multiple access OFDM MIMO channel with separately correlated Rician fading," in *Proc. IEEE GLOBECOM 2008*, New Orleans, LA, Nov. 30–Dec. 4 2008.

[37] A. M. Sayeed, "Deconstructing multi-antenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[38] J. Schwinger, "A note to the quantum dynamical principle," *Philos. Mag.*, vol. 44, pp. 1171–1179, 1953.

[39] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[40] 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.

[41] T. Tanaka, "Moment problem in replica method," *Interdisciplin. Inf. Sci.*, vol. 13, no. 1, pp. 17–23, 2007.

[42] G. Taricco, "On the capacity of separately correlated MIMO Rician fading channels," in *Proc. IEEE GLOBECOM 2006*, San Francisco, CA, Nov. 27–Dec. 1 2006.

[43] G. Taricco, "Asymptotic mutual information statistics of separately correlated Rician fading MIMO channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3490–3504, Aug. 2008.

[44] G. Taricco and E. Riegler, "On the ergodic capacity of the asymptotic separately correlated Rician fading MIMO channel with interference," in *Proc. IEEE ISIT 2007*, Nice, France, June 24–30, 2007.

[45] G. Taricco and E. Riegler, "On the ergodic capacity of correlated Rician fading MIMO channels with interference," *IEEE Trans. Inf. Theory.*, submitted for publication.

[46] G. Taricco and E. Riegler, "Asymptotic ergodic capacity of wideband MIMO channels with separately correlated Rician fading," in *Proc. IEEE GLOBECOM 2008*, New Orleans, LA, Nov. 30–Dec. 4 2008.

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

[48] A. M. Tulino, A. Lozano, and S. Verdu, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 662–671, Mar. 2006.

[49] S. Venkatesan, S. H. Simon, and R. A. Valenzuela, "Capacity of a Gaussian MIMO channel with nonzero mean," in *IEEE VTC 2003*, Orlando, FL, Oct. 6–9, 2003.

[50] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.

[51] C.-K. Wen and K.-K. Wong, "Asymptotic analysis of spatially correlated MIMO multiple-access channels with arbitrary signaling inputs for joint and separate decoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 252–268, Jan. 2007.

[52] J. H. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 5, pp. 871–878, Jun. 1987.

Erwin Riegler (M'07) was born in Neunkirchen, Austria. He received the Dipl-Ing. degree in technical physics (*summa cum laude*) in 2001 and the Dr. techn. degree in technical physics (*summa cum laude*) in 2004 from the Vienna University of Technology, Vienna, Austria.

He visited the Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany (September 2004–February 2005). From 2005 to 2006, he accepted a postdoctoral position at the Institute for Analysis and Scientific Computing, Vienna University of Technology, in the special research program (SFB) AURORA—Advanced Models, Applications and Software Systems for High Performance Computing. Since 2007, he has been a Senior Researcher with the Telecommunications Research Center Vienna (FTW) and in 2009, he also became affiliated with the Vienna University of Technology working in the research project Distributed Information Processing for Spatio-Temporal Fields in Wireless Sensor Networks. His research interests include MIMO communications, interference management, large system analysis, and transceiver design.

Giorgio Taricco (M'91–SM'03–F'10) was born in Torino, Italy. He received the Electronic Engineer degree (*cum laude*) from the Politecnico di Torino, Torino, Italy, in 1985.

He was a Research Associate with the Italian Telecom Labs from 1985 to 1987, where he was involved in the design of the GSM mobile telephony standard (channel coding). In 1996, he was a Research Fellow at ESTEC, Noordwijk, The Netherlands. He is now a Full Professor with the Politecnico di Torino. His research interests include error-control coding, multiuser detection, space-time coding, MIMO communications, and cognitive radio networks. He has coauthored approximately 60 papers in international journals, around 130 international conference contributions, several book chapters, and holds two international patents.

Prof. Taricco is an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY and the *Journal on Communications and Networks*. From 2003 to 2008, he was an Associate Editor for the IEEE COMMUNICATIONS LETTERS. He has been the Treasurer of the Conferences IEEE ISIT 2000, ISITA 2004, and IEEE ITW 2009.