Random Matrix Analysis of Large Relay Networks

Veniamin I. Morgenshtern and Helmut Bölcskei

***Abstract*— We analyze fading interference relay networks where M single-antenna source-destination terminal pairs communicate concurrently and in the same frequency band through a set of K single-antenna relays using half-duplex two-hop relaying. The relays do not have channel state information, perform amplify-and-forward (AF) relaying, and the destination terminals can cooperate and perform joint decoding. Our main results are as follows:**

- **We compute the per source-destination terminal pair capacity for $M, K \to \infty$, with $K/M \to \beta$ fixed, using tools from random matrix theory.**
- **We show that for $\beta \to \infty$, the AF relay network is turned into a point-to-point multiple-input multiple-output link and thus extend the result found previously for the finite $M, K \to \infty$ case in [1] to the $M, K \to \infty$ case.**

I. Introduction

This paper deals with interference fading relay networks where M single-antenna source-destination terminal pairs communicate concurrently and in the same frequency band through half-duplex two-hop relaying over a common set of K single-antenna relay terminals (see Fig. 1). The relays do not have channel state information (CSI), perform amplify-and-forward (AF) relaying, and the destination terminals can cooperate and perform joint decoding.

A. Contributions and Relation to Previous Work

Previous work in [1] demonstrated that for M fixed and $K \to \infty$, AF relaying turns the fading interference relay network into a fading point-to-point multiple-input multiple-output (MIMO) link, showing that the use of relays as active scatterers can recover spatial multiplexing gain in poor scattering environments. Our main contributions are as follows:

- The proof techniques in [1] rely heavily on M being finite. Building on results reported in [2], we compute the $M, K \to \infty$ (with $K/M \to \beta$ fixed) per source-destination terminal pair capacity using tools from random matrix theory [3], [4]. The limiting eigenvalue density function of the effective MIMO channel matrix between the source and destination terminals is characterized in terms of its Stieltjes transform as the unique solution of a fixed-point equation, which can be transformed into a fourth-order equation. Upon solving this fourth-order equation and applying the inverse Stieltjes transform, the remaining steps to computing the limiting eigenvalue density function, and, based on that, the asymptotic network capacity, need to be carried out numerically. We show that this can be accomplished in a straightforward fashion and provide a corresponding algorithm.
- We show that for $\beta \to \infty$, the fading AF relay network is turned into a fading point-to-point MIMO link (in a sense to be made precise in Section IV), thus establishing the large-M, K analog of the result found previously for the finite-$M, K \to \infty$ case in [1].

This research was supported by Nokia Research Center Helsinki, Finland and by the STREP project No. IST-027310 (MEMBRANE) within the Sixth Framework Programme of the European Commission.

Veniamin I. Morgenshtern and Helmut Bölcskei are with the Communication Theory Laboratory, ETH Zurich, Switzerland, Email: {vmorgens,boelcskei}@nari.ee.ethz.ch

B. Notation

The superscripts ^{T}, ^{H}, and * stand for transposition, conjugate transpose, and element-wise conjugation, respectively. $\log(\cdot)$ stands for the logarithm to base 2. $I[x] = 1$ if x is true and $I[x] = 0$ if x is false. The unit step function $u(x) = 0$ for $x < 0$ and $u(x) = 1$ for $x \geq 0$. $\mathbb{E}$ denotes the expectation operator. A circularly symmetric zero-mean complex Gaussian random variable (RV) is a RV $Z = X + j\,Y \sim \mathcal{CN}(0, \sigma^2)$, where X and Y are independent identically distributed (i.i.d.) $\mathcal{N}(0, \sigma^2/2)$. $\delta(x)$ is the Dirac delta distribution. $(x)^+ = x$ for $x > 0$ and 0 otherwise. Matrices and vectors (both deterministic and random) are denoted by uppercase and lowercase, respectively, boldface letters. The element of a matrix $\mathbf{X}$ in the nth row and mth column and the nth element of a vector $\mathbf{x}$ are denoted as $[\mathbf{X}]_{n,m}$ and $[\mathbf{x}]_n$, respectively. $\lambda_i(\mathbf{X})$, $\lambda_{\min}(\mathbf{X})$, and $\lambda_{\max}(\mathbf{X})$ stand for the ith, the minimum, and the maximum eigenvalue of a matrix $\mathbf{X}$, respectively. $\|\mathbf{x}\|$ denotes the ℓ^2-norm of the vector $\mathbf{x}$. $\Re z$ and $\Im z$ designate the real and imaginary part of $z \in \mathbb{C}$, respectively. $\mathbb{C}^+ \triangleq \{z \in \mathbb{C} : \Im z > 0\}$. For any $n, m \in \mathbb{N}$, $m \geq n$, $[n : m]$ denotes the natural numbers $\{n, n+1, \ldots, m\}$.

II. Channel and Signal Model

A. General Assumptions

We consider an interference relay network (see Fig. 1) consisting of $K + 2M$ single-antenna terminals with M designated source-destination terminal pairs $\{\mathcal{S}_m, \mathcal{D}_m\}$ $(m \in [1\!:\!M])$ and K relays $\mathcal{R}_k$ $(k \in [1\!:\!K])$. We assume that no direct link between the individual source-destination terminal pairs exists (e.g., caused by large separation). Transmission takes place in half-duplex fashion (the terminals cannot transmit and receive simultaneously) in two hops (a.k.a. two-hop relaying) over two separate time slots. In the first time slot, the source terminals simultaneously broadcast their information to all the relay terminals (i.e., each relay terminal receives a superposition of all source signals). After

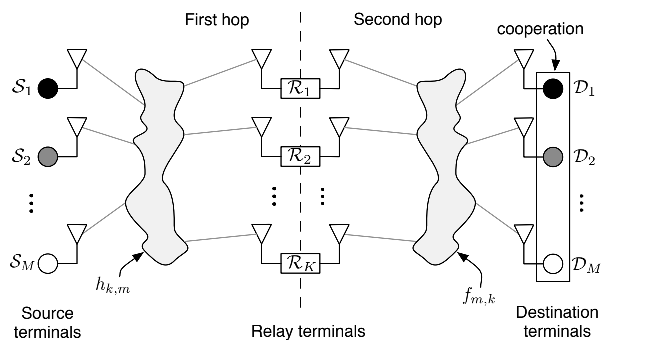


Fig. 1. Two-hop wireless relay network setup.

processing the received signals, the relay terminals simultaneously broadcast the processed data to all the destination terminals during the second time slot. Our setup can be considered as an interference channel [5] with dedicated relays, hence the terminology *interference relay network*.

B. Channel and Signal Model

Throughout the paper, frequency-flat fading over the bandwidth of interest as well as perfectly synchronized transmission and reception between the terminals is assumed. The input-output (I-O) relation for the link between the source terminals and the relay terminals during the first time slot is given by

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{z} \tag{1}$$

where $\mathbf{r} = [r_1, r_2, \ldots, r_K]^T$ with r_k denoting the signal received at the kth relay terminal, $\mathbf{H} \in \mathbb{C}^{K\times M}$ with $[\mathbf{H}]_{k,m} = h_{k,m}$ $(k \in [1\!:\!K],\ m \in [1\!:\!M])$ where $h_{k,m}$ denotes the i.i.d. complex-valued zero-mean channel gains with variance one corresponding to the $\mathcal{S}_m \rightarrow \mathcal{R}_k$ links, $\mathbf{s} = [s_1, s_2, \ldots, s_M]^T$ where s_m is the zero-mean Gaussian signal transmitted by $\mathcal{S}_m$ and the vector $\mathbf{s}$ is i.i.d. temporally and spatially (across source terminals). Finally, $\mathbf{z} = [z_1, z_2, \ldots, z_K]^T$ where $z_k \sim \mathcal{CN}(0, \sigma^2)$ is temporally and spatially (across relay terminals) white noise. The kth relay terminal simply scales its received signal r_k to produce the output signal t_k. The collection of output signals t_k, organized in the vector $\mathbf{t} = [t_1, t_2, \ldots, t_K]^T$, is then broadcast to the destination terminals during the second time slot, while the source terminals remain silent. The mth destination terminal receives the signal y_m with $\mathbf{y} = [y_1, y_2, \ldots, y_M]^T$ given by

$$\mathbf{y} = \mathbf{F}\mathbf{t} + \mathbf{w} \tag{2}$$

where $\mathbf{F} \in \mathbb{C}^{M\times K}$ with $[\mathbf{F}]_{m,k} = f_{m,k}$ $(m \in [1\!:\!M],\ k \in [1\!:\!K])$ where $f_{m,k}$ denotes the i.i.d. complex-valued zero-mean channel gains with variance one corresponding to the $\mathcal{R}_k \rightarrow \mathcal{D}_m$ links, and $\mathbf{w} = [w_1, w_2, \ldots, w_M]^T$ with $w_m \sim \mathcal{CN}(0, \sigma^2)$ being temporally and spatially (across destination terminals) white noise. We impose a per-source-terminal power constraint $\mathbb{E}\big[|s_m|^2\big] \leq 1/M$ $(m \in [1\!:\!M])$, which results in the total transmit power trivially satisfying the constraint $\mathbb{E}\big[\|\mathbf{s}\|^2\big] \leq 1$. Furthermore, we impose a per-relay-terminal power constraint $\mathbb{E}\big[|t_k|^2\big] \leq P_{\mathrm{rel}}/K$ $(k \in [1\!:\!K])$ by setting $t_k = \sqrt{P_{\mathrm{rel}}/((1+\sigma^2)K)}\, r_k$; this results trivially in the total power transmitted by the relay terminals satisfying $\mathbb{E}\big[\|\mathbf{t}\|^2\big] \leq P_{\mathrm{rel}}$.

Throughout the paper, we assume that the source and relay terminals do not have CSI and the destination terminals perform *joint decoding* and have access to the realizations of $\mathbf{H}$ and $\mathbf{F}$. In fact, as the analysis below shows, knowledge of $\mathbf{FH}$ and $\mathbf{F}$ is sufficient.

We conclude by noting that the entries in $\mathbf{H}$ and $\mathbf{F}$ were assumed to all have the same variance, which implies that our signal model does not account for pathloss. This assumption is conceptual as the proof technique used to derive the main result of the paper does not seem to extend to the case of general (finite) variances of the entries in $\mathbf{H}$ and $\mathbf{F}$. On the other hand, we do not require $\mathbf{H}$ and $\mathbf{F}$ to have Gaussian entries.

III. Asymptotic Network Capacity

The overall I-O relation is obtained by inserting (1) into (2) and reads

$$\mathbf{y} = \frac{d}{\sqrt{K}}\mathbf{F}\mathbf{H}\mathbf{s} + \frac{d}{\sqrt{K}}\mathbf{F}\mathbf{z} + \mathbf{w} \tag{3}$$

where $d = \sqrt{P_{\mathrm{rel}}/(1+\sigma^2)}$. Based on the I-O relation (3), we shall next study the behavior of $I(\mathbf{y};\mathbf{s}\,|\,\mathbf{FH},\mathbf{F})$ when $M, K \rightarrow \infty$ with $K/M \rightarrow \beta$. We start by noting that

$$I(\mathbf{y};\mathbf{s}\,|\,\mathbf{FH},\mathbf{F}) = \\ = \log\det\left(\mathbf{I} + \frac{d^2}{\sigma^2 MK}\mathbf{H}^H\mathbf{F}^H\left(\frac{d^2}{K}\mathbf{F}\mathbf{F}^H + \mathbf{I}\right)^{-1}\mathbf{F}\mathbf{H}\right).$$

Since the destination terminals perform joint decoding, the ergodic capacity per source-destination terminal pair is given by[1]

$$C_{\mathrm{AF}} = \frac{1}{2}\,\mathbb{E}\left[\frac{1}{M}\sum_{k=1}^{K}\log\left(1 + \frac{1}{\sigma^2}\lambda_k\left(\frac{1}{M}\mathbf{H}\mathbf{H}^H\mathbf{T}\right)\right)\right] \tag{4}$$

where

$$\mathbf{T} \triangleq \frac{d^2}{K}\mathbf{F}^H\left(\mathbf{I} + \frac{d^2}{K}\mathbf{F}\mathbf{F}^H\right)^{-1}\mathbf{F}$$

and the factor $1/2$ in (4) results from the fact that data is transmitted over two time slots.

To compute C_{AF} in the $M, K \rightarrow \infty$ limit with $K/M \rightarrow \beta$, we start by analyzing the corresponding asymptotic behavior of $\lambda_k\big((1/M)\mathbf{H}\mathbf{H}^H\mathbf{T}\big)$. To this end, we define the empirical spectral distribution (ESD) of a matrix (random or deterministic) according to

Definition 1: Let $\mathbf{X} \in \mathbb{C}^{N\times N}$ be a Hermitian matrix. The ESD of $\mathbf{X}$ is defined as

$$F^N_{\mathbf{X}}(x) \triangleq \frac{1}{N}\sum_{n=1}^{N} I\left[\lambda_n(\mathbf{X}) \leq x\right].$$

For random $\mathbf{X}$, the quantity $F^N_{\mathbf{X}}(x)$ is random as well, i.e., it is a RV for each x. In the following, our goal is to prove the convergence (in the sense defined below), when $M, K \rightarrow$

[1] In the finite-M, K case, we need $\mathbf{H}$ and $\mathbf{F}$ to be ergodic for (4) to be well-defined.

∞ with $K/M \to \beta$ and $\beta \in (0,\infty)$, of $F^{K}_{(1/M)\mathbf{H}\mathbf{H}^H\mathbf{T}}(x)$ to a deterministic limit and find the corresponding limiting eigenvalue distribution.

Definition 2: We say that the ESD $F^{N}_{\mathbf{X}}(x)$ of a random Hermitian matrix $\mathbf{X} \in \mathbb{C}^{N\times N}$ converges a.s. to a deterministic limiting function $F_{\mathbf{X}}(x)$, when $N \to \infty$, if for any $\epsilon > 0$ there exists an $N_0 > 0$ s.t. $\forall N \geq N_0$ a.s.

$$\sup_{x\in\mathbb{R}} \left| F^{N}_{\mathbf{X}}(x) - F_{\mathbf{X}}(x) \right| \leq \epsilon.$$

To prove the convergence of $F^{K}_{(1/M)\mathbf{H}\mathbf{H}^H\mathbf{T}}(x)$ to a deterministic limiting function, we start by analyzing $F^{K}_{\mathbf{T}}(x)$.

Lemma 1: For $M,K \to \infty$ with $K/M \to \beta$, the ESD $F^{K}_{\mathbf{T}}(x)$ converges a.s. to a nonrandom limiting distribution $F_{\mathbf{T}}(x)$ with corresponding density given by[2]

$$\begin{aligned} f_{\mathbf{T}}(x) &= \\ &= \frac{\sqrt{(1+\gamma_1)(1+\gamma_2)}}{2\pi d^2 x(1-x)^2} \sqrt{\left(\frac{\gamma_2}{1+\gamma_2} - x\right)^{+} \left(x - \frac{\gamma_1}{1+\gamma_1}\right)^{+}} \\ &\quad + \left[1 - \frac{1}{\beta}\right]^{+} \delta(x) \end{aligned} \tag{5}$$

where $\gamma_1 \triangleq d^2(1 - 1/\sqrt{\beta})^2$ and $\gamma_2 \triangleq d^2(1 + 1/\sqrt{\beta})^2$.

Proof: We start with the singular value decomposition

$$\frac{d}{\sqrt{K}}\mathbf{F} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}$$

where the columns of $\mathbf{U} \in \mathbb{C}^{M,M}$ are the eigenvectors of the matrix $(d^2/K)\mathbf{F}\mathbf{F}^H$, the columns of $\mathbf{V}^H \in \mathbb{C}^{K,K}$ are the eigenvectors of $(d^2/K)\mathbf{F}^H\mathbf{F}$, and the matrix $\mathbf{\Sigma} \in \mathbb{R}^{M,K}$ contains $R = \min(M,K)$ nonzero entries $\Sigma_{11}, \Sigma_{22}, \ldots, \Sigma_{RR}$, which are the positive square roots of the nonzero eigenvalues of the matrix $(d^2/K)\mathbf{F}\mathbf{F}^H$. Defining $\mathbf{\Lambda} \triangleq \mathbf{\Sigma}\mathbf{\Sigma}^H \in \mathbb{R}^{M,M}$, we have

$$\mathbf{T} = \mathbf{V}^H \mathbf{\Sigma}^H \left(\mathbf{I} + \mathbf{\Lambda}\right)^{-1} \mathbf{\Sigma}\mathbf{V}.$$

By inspection, it follows that

$$F^{K}_{\mathbf{\Sigma}^H(\mathbf{I}+\mathbf{\Lambda})^{-1}\mathbf{\Sigma}}(x) = \frac{M}{K} F^{M}_{\mathbf{\Lambda}}\left(\frac{x}{1-x}\right) + \left(1 - \frac{M}{K}\right) u(x). \tag{6}$$

As $F^{M}_{\mathbf{\Lambda}}(x) = F^{M}_{(d^2/K)\mathbf{F}\mathbf{F}^H}(x)$, by the Marčenko-Pastur law (see Theorem 2 in Appendix A), we conclude that $F^{M}_{\mathbf{\Lambda}}(x)$ converges a.s. to a limiting nonrandom distribution $F_{\mathbf{\Lambda}}(x)$ with corresponding density

$$f_{\mathbf{\Lambda}}(x) = \frac{\beta}{2\pi x d^2} \sqrt{(\gamma_2 - x)^{+} (x - \gamma_1)^{+}} + [1-\beta]^{+}\delta(x). \tag{7}$$

From (6) we can, therefore, conclude that $F^{K}_{\mathbf{\Sigma}^H(\mathbf{I}+\mathbf{\Lambda})^{-1}\mathbf{\Sigma}}(x)$ converges a.s. to a nonrandom limit given by

$$F_{\mathbf{\Sigma}^H(\mathbf{I}+\mathbf{\Lambda})^{-1}\mathbf{\Sigma}}(x) = \frac{1}{\beta} F_{\mathbf{\Lambda}}\left(\frac{x}{1-x}\right) + \left(1 - \frac{1}{\beta}\right) u(x). \tag{8}$$

[2]Note that (5) implies that $f_{\mathbf{T}}(x)$ is compactly supported in the interval $[\gamma_1/(1+\gamma_1), \gamma_2/(1+\gamma_2)]$.

Taking the derivative w.r.t. x on both sides of (8), the density corresponding to $F_{\mathbf{\Sigma}^H(\mathbf{I}+\mathbf{\Lambda})^{-1}\mathbf{\Sigma}}(x)$ is obtained as

$$\begin{aligned} f_{\mathbf{\Sigma}^H(\mathbf{I}+\mathbf{\Lambda})^{-1}\mathbf{\Sigma}}(x) &= \\ &= \frac{1}{\beta} f_{\mathbf{\Lambda}}\left(\frac{x}{1-x}\right) \frac{1}{(1-x)^2} + \left(1 - \frac{1}{\beta}\right)\delta(x). \end{aligned} \tag{9}$$

We obtain the final result in (5) by noting that $f_{\mathbf{T}}(x) = f_{\mathbf{\Sigma}^H(\mathbf{I}+\mathbf{\Lambda})^{-1}\mathbf{\Sigma}}(x)$ because of the unitarity of $\mathbf{V}$ and by inserting (7) into (9) and carrying out straightforward algebraic manipulations. ■

Based on Lemma 1, we can now apply Theorem 1 (Appendix A) to conclude that $F^{K}_{(1/M)\mathbf{H}\mathbf{H}^H\mathbf{T}}(x)$ converges a.s. to a deterministic function $F_{(1/M)\mathbf{H}\mathbf{H}^H\mathbf{T}}(x)$ as $M, K \to \infty$ with $K/M \to \beta$. The corresponding limiting density $f_{(1/M)\mathbf{H}\mathbf{H}^H\mathbf{T}}(x)$ is obtained by applying the Stieltjes inversion formula (21) to the solution of the fixed-point equation

$$G(z) = \underbrace{\int_{-\infty}^{\infty} \frac{f_{\mathbf{T}}(x)dx}{x(1-\beta-\beta z G(z)) - z}}_{I}, \quad z \in \mathbb{C}^{+} \tag{10}$$

in the set

$$\left\{ G(z) \in \mathbb{C} \,\middle|\, -(1-\beta)/z + \beta G(z) \in \mathbb{C}^{+} \right\}, \quad z \in \mathbb{C}^{+} \tag{11}$$

where we used the symbol $G(z)$ to denote the Stieltjes transform $G_{F_{(1/M)\mathbf{H}\mathbf{H}^H\mathbf{T}}}(z)$. In the following, for brevity, we write G instead of $G(z)$. To solve (10), we first compute the integral I on the right-hand side (RHS) of (10). We substitute $f_{\mathbf{T}}(x)$ from (5) into (10) and define

$$\eta_1 \triangleq \frac{\gamma_1}{1+\gamma_1}, \; \eta_2 \triangleq \frac{\gamma_2}{1+\gamma_2}, \; \rho \triangleq \frac{\sqrt{(1+\gamma_1)(1+\gamma_2)}}{2\pi d^2}$$

to obtain

$$\begin{aligned} I = &-\frac{1}{z}\left[1 - \frac{1}{\beta}\right]^{+} \\ &+ \frac{1}{z} \underbrace{\int_{\eta_1}^{\eta_2} \frac{\rho\sqrt{(\eta_2 - x)(x - \eta_1)}\,dx}{x(1-x)^2 \left(x\left(\frac{1-\beta}{z} - \beta G\right) - 1\right)}}_{\hat{I}}. \end{aligned} \tag{12}$$

The integral $\hat{I}$ is computed in Appendix B. Employing the notation introduced in Appendix B, we can finally write the fixed point equation (10) as

$$Gz = -\left[1 - \frac{1}{\beta}\right]^{+} + \chi A_1 \hat{I}_1 + \chi A_2 \hat{I}_2 + \chi A_3 \hat{I}_3 + \chi A_4 \hat{I}_4. \tag{13}$$

It is tedious, but straightforward, to show that for any $\beta > 0$

$$-\left[1 - \frac{1}{\beta}\right]^{+} + \chi A_1 \hat{I}_1 = -\frac{\beta - 1}{2\beta}$$

so that (13) can be written as

$$Gz + \frac{\beta - 1}{2\beta} - \chi A_2 \hat{I}_2 - \chi A_3 \hat{I}_3 = \chi A_4 \hat{I}_4. \tag{14}$$

Next, multiplying (14) by $2d^2\beta(G\beta z + z + \beta - 1)^2$, squaring both sides, introducing the auxiliary variable

$$\hat{G} \triangleq -\frac{1-\beta}{z} + \beta G$$

we obtain after straightforward, but tedious, manipulations that $\hat{G}$ must satisfy the following quartic equation

$$\hat{G}^4 + a_3\hat{G}^3 + a_2\hat{G}^2 + a_1\hat{G} + a_0 = 0 \tag{15}$$

with the coefficients

$$a_3 = \frac{1}{z}(2z - \beta + 1) \qquad a_2 = \frac{1}{z}\left(z - \beta + 3 - \frac{\beta}{d^2}\right)$$
$$a_1 = \frac{1}{z^2}\left(2z - \beta + 1 - \frac{\beta}{d^2}\right) \quad a_0 = \frac{1}{z^2}.$$

The quartic equation (15) can be solved analytically. The resulting expressions are, however, very lengthy, do not lead to interesting insights, and will therefore be omitted. It is important to note, however, that (15) has two pairs of complex conjugate roots. The solutions of (15) will henceforth be denoted as $\hat{G}_1, \hat{G}_1^*, \hat{G}_2$, and $\hat{G}_2^*$. We recall that our goal is to find the unique solution G of the fixed point equation (10) s.t. $\hat{G} = -(1-\beta)/z + \beta G \in \mathbb{C}^+$, $\forall z \in \mathbb{C}^+$. Therefore, in each point $z \in \mathbb{C}^+$ we can immediately eliminate the two solutions (out of the four) that have a negative imaginary part. In practice, this can be done conveniently by constructing the functions $\hat{G}_1' \triangleq \Re\hat{G}_1 + j\left|\Im\hat{G}_1\right|$ and $\hat{G}_2' \triangleq \Re\hat{G}_2 + j\left|\Im\hat{G}_2\right|$, which can be computed analytically, satisfy (15), and are in $\mathbb{C}^+$ for any $z \in \mathbb{C}^+$. Next, note that (14) has a unique solution in the set (11), which is also the unique solution of (10). We can obtain this solution $G(z)$, $z \in \mathbb{C}^+$, by substituting $G_1 = (1/\beta)(\hat{G}_1' - (\beta-1)/z)$ and $G_2 = (1/\beta)(\hat{G}_2' - (\beta-1)/z)$ into (14) and checking which of the two satisfies the equation. Unfortunately, it seems that this verification cannot be formalized in the sense of identifying the unique solution of (14) in analytic form. The primary reason for this is that to check *algebraically* if G_1 or G_2 satisfy (14), we have to perform a noninvertible transformation (squaring) of (14), which doubles the number of solutions of this equation, and results in G_1 and G_2 both satisfying the resulting equation. The second reason is that, depending on the values of the parameters $\beta > 0, d > 0$, the correct solution is either G_1 or G_2, and the dependence between G_1, G_2, β, and d has a complicated structure. Starting from the analytical expressions for G_1 and G_2, we can identify, however, for any fixed $\beta > 0, d > 0$, the density function $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x) = (1/\pi)\lim_{y\to 0^+}\Im[G(x+jy)]$ corresponding to the unique solution of (14) [and hence of (10)] numerically. This is accomplished as follows. We know that, for given x, $\lim_{y\to 0^+}\Im\left[G(x+jy)\right]$ is either equal to

$$L_1(x) \triangleq \lim_{y\to 0^+}\Im\left[G_1(x+jy)\right]$$

or

$$L_2(x) \triangleq \lim_{y\to 0^+}\Im\left[G_2(x+jy)\right].$$

Even though the functions $L_1(x)$ and $L_2(x)$ can be computed analytically (with the resulting expressions being very lengthy and involved), it seems that for any fixed $x > 0$ the correct choice between the values $L_1(x)$ and $L_2(x)$ can only be made numerically. The following algorithm constitutes one possibility to solve this problem.

Algorithm—Choice of the Limit
Input: $x > 0$

1) Choose a small enough $y > 0$
2) Substitute $G_1(x + jy)$ and $G_2(x + jy)$ into (14)
3) *If* $G_1(x + jy)$ satisfies (14), *then*
   return $L_1(x)$
   otherwise
   return $L_2(x)$

As any other numerical procedure, this algorithm includes a heuristic element. The following comments are therefore in order.

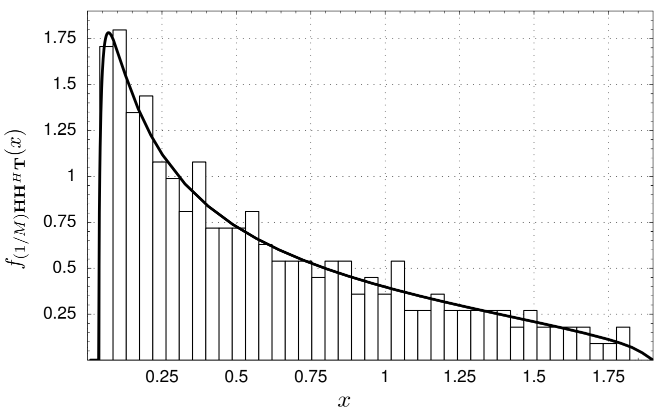
- In Step 1, the choice of y cannot be formalized in the sense of giving an indication of how small y has to be as a function of β and d. On the one hand, y has to be strictly greater than zero, because (14), in general, holds in $\mathbb{C}^+$ only and does not need to hold neither for $G_1(x+j0)$ nor for $G_2(x+j0)$. On the other hand, y should be small enough for $G_1(x+jy)$ to be close to $L_1(x)$ and $G_2(x+jy)$ to be close to $L_2(x)$. The correctness of the output of the algorithm is justified by the fact that $G(z)$ is analytic in $\mathbb{C}^+$ (see Definition 3).
- In Step 3 the check whether $G_1(x+jy)$ satisfies (14) is performed numerically. Therefore, rounding errors will arise. It turns out, however, that in practice, unless $\left|L_1(x) - L_2(x)\right|$ is very small (in this case it does not matter which of the two values we choose), the solution of (14) yields a clear indication whether $G_1(x+jy)$ or $G_2(x+jy)$ is the correct choice.
- To compute the density $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ using the proposed algorithm, we need to run Steps 1–3 for every x. It will be proved below that $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ is always compactly supported and bounds for its support will be given in analytic form (as a function of β and d). Since the algorithm consists of very basic arithmetic operations only, it is very fast and can easily be run on a dense grid inside the support region of $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$.

As an example, for $\beta = 1/2$ and $d = 1$, Fig. 2(a) shows the density $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ obtained by the algorithm formulated above along with the histogram of the same density obtained through Monte-Carlo simulation. We can see that the two curves match very closely and that our method allows to obtain a much more refined picture of the limiting density. Fig. 2(b) shows $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ for $\beta = 2$, 1, $1/2$ and $d = 1$ obtained through our algorithm.

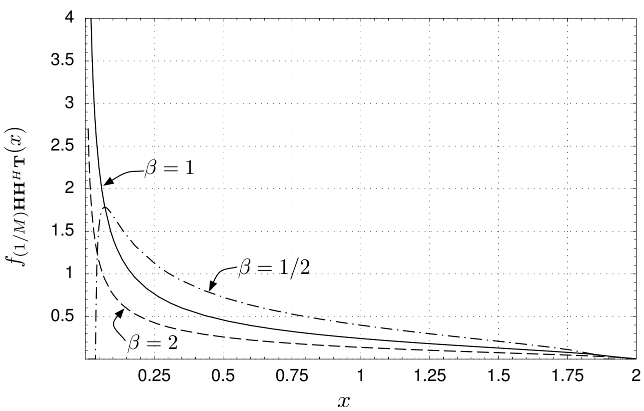
The final step in computing the asymptotic capacity of the AF relay network is to take the limit $K, M \to \infty$ with $K/M \to \beta$ in (4) and to evaluate the resulting integral

$$C_{\mathrm{AF}}^{\beta} \triangleq \frac{\beta}{2}\int_0^\infty \log\left(1 + \frac{x}{\sigma^2}\right) f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)\,dx \tag{16}$$

numerically. The evaluation of (16) is drastically simplified by taking into account that $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ is compactly supported. The corresponding interval boundaries (or, more specifically, bounds thereon) can be computed analytically as a function of β and d. We start by noting that the



(a)



(b)

Fig. 2. Limiting density $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ (a) for $\beta = 1/2$ and $d = 1$ along with its histogram (Monte-Carlo) and (b) for $\beta = 2, 1, 1/2$ and $d = 1$.

second part of Theorem 2 in Appendix A implies (under the additional assumption that the entries of $\mathbf{H}$ have finite fourth moments) that a.s. $\lim_{M\to\infty} \lambda_{\max}\big((1/M)\mathbf{HH}^H\big) = (1+\sqrt{\beta})^2$. From (9) and Theorem 2, it follows that a.s. $\lambda_{\max}(\mathbf{T}) = d^2(1+\sqrt{\beta})^2/(\beta + d^2(1+\sqrt{\beta})^2)$. For any realization of $\mathbf{H}$ and $\mathbf{T}$ and any M, K, by the submultiplicativity of the spectral norm, we have

$$\lambda_{\max}\big((1/M)\mathbf{HH}^H\mathbf{T}\big) \le \lambda_{\max}\big((1/M)\mathbf{HH}^H\big)\,\lambda_{\max}(\mathbf{T})$$

which implies that for $M, K \to \infty$ with $K/M \to \beta$ a.s.

$$\lambda_{\max}\big((1/M)\mathbf{HH}^H\mathbf{T}\big) \le \frac{d^2(1+\sqrt{\beta})^4}{\beta + d^2(1+\sqrt{\beta})^2} \triangleq x_{\max}.$$

We can thus conclude that $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ is compactly supported on the interval[3] $[0, x_{\max}]$. Consequently, the integral in (16) becomes

$$C_{\mathrm{AF}}^{\beta} = \frac{\beta}{2}\int_0^{x_{\max}} \log\Big(1+\frac{x}{\sigma^2}\Big)\, f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)\, dx$$

which can be computed numerically, using any standard method for numerical integration and employing the algorithm described above to evaluate $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ at the required grid points. Using this procedure, we computed C_{AF}^{β} as a function of β for $d = 1$ with the result depicted in Fig. 3. We can see that for $\beta < 1$ (i.e., $K < M$), C_{AF}^{β} increases very quickly with β, which is due to the fact that the corresponding effective MIMO channel matrix (i.e., the matrix between the $\mathcal{S}_m$ and the $\mathcal{D}_m$) is building up rank and hence spatial multiplexing gain. For $\beta > 1$ (i.e., $K > M$), when the effective MIMO channel matrix is already full rank with high probability, the curve flattens out and for $\beta \to \infty$, the capacity C_{AF}^{β} seems to converge to a finite value. In the next section, we prove that C_{AF}^{β} indeed converges to a finite limit as $\beta \to \infty$. This result has an interesting interpretation as it allows to relate the AF relay network to a point-to-point MIMO channel.

[3]The actual supporting interval of $f_{(1/M)\mathbf{HH}^H\mathbf{T}}(x)$ may, in fact, be smaller.

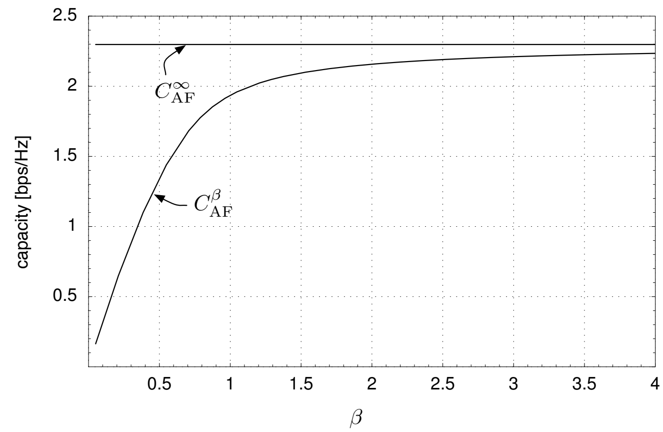


Fig. 3. Capacity C_{AF}^{β} as a function of β for $d = 1$ and $\sigma^2 = 0.01$.

IV. Convergence to Point-to-Point MIMO Channel

In [1], it was shown that for finite M, as $K \to \infty$, the two-hop AF relay network capacity converges to half the capacity of a point-to-point MIMO link; the factor $1/2$ penalty comes from the fact that communication takes place over two time slots. In the following, we demonstrate that the result in [1] can be generalized to the $M, K \to \infty$ case. More specifically, we show that for $\beta \to \infty$ the asymptotic ($M, K \to \infty$) capacity of the two-hop AF relay network is equal to half the asymptotic ($M \to \infty$) capacity of a point-to-point MIMO channel with M transmit and M receive antennas. We start by dividing (15) by β and taking the limit[4] $\beta \to \infty$, which yields the quadratic equation

$$z\hat{G}^2 + z\left(1+\frac{1}{d^2}\right)\hat{G} + \left(1+\frac{1}{d^2}\right) = 0. \tag{17}$$

The two solutions of (17) are given by

$$\hat{G}_{1,2}(z) = \frac{-z\big(1+\frac{1}{d^2}\big) \pm \sqrt{z^2\left(1+\frac{1}{d^2}\right)^2 - 4z\left(1+\frac{1}{d^2}\right)}}{2z}. \tag{18}$$

[4]It is important that we take the limit $M, K \to \infty$ with $K/M \to \beta$ first and afterwards let $\beta \to \infty$.

Applying the Stieltjes inversion formula (21) to (18) and choosing the solution that yields a positive density function, we obtain

$$\beta f_{(1/M)\mathbf{H}\mathbf{H}^H\mathbf{T}}(x) = \\ = \frac{1}{\pi} \lim_{y\to 0^+} \Im\left[\beta G(x+jy)\right] \\ = \frac{1}{\pi} \lim_{y\to 0^+} \Im\left[\hat{G}(x+jy)\right] \\ = \frac{1}{2\pi x}\sqrt{\left[4x\left(1+\frac{1}{d^2}\right) - x^2\left(1+\frac{1}{d^2}\right)^2\right]^+}. \tag{19}$$

Inserting (19) into (16) and changing the integration variable according to $u \triangleq x\big(1+1/d^2\big)$, we find that $C_{\mathrm{AF}}^{\beta} \xrightarrow{\beta\to\infty} C_{\mathrm{AF}}^{\infty}$, where

$$C_{\mathrm{AF}}^{\infty} \triangleq \frac{1}{4\pi}\int_0^4 \sqrt{\frac{4}{u}-1}\,\log\left(1+\frac{d^2}{(d^2+1)\sigma^2}u\right) du. \tag{20}$$

Comparing (20) with [6, Eq. (13)], it follows that for $\beta\to\infty$ the asymptotic $M,K\to\infty$ with $K/M\to\beta$ per source-destination terminal pair capacity in the two-hop AF relay network is equal to half the asymptotic $(M\to\infty)$ per-antenna capacity in a point-to-point MIMO link with M transmit and M receive antennas, provided the SNR in the relay case is defined as $\mathrm{SNR} \triangleq d^2/\big((d^2+1)\sigma^2\big)$. For M and K large, it is easy to verify that this choice corresponds to the SNR at each destination terminal in the AF relay network. In this sense, we can conclude that for $\beta\to\infty$ the AF relay network "converges" to a point-to-point MIMO link with the same received SNR.

V. Conclusion

For a two-hop AF relay network with joint decoding at the destination terminals, we computed the asymptotic (in M and K with $K/M\to\beta$ fixed) network capacity using tools from random matrix theory. To the best of our knowledge, this is the first application of random matrix theory to characterize the capacity behavior of large fading networks. We furthermore demonstrated that for $\beta\to\infty$ the relay network converges to a point-to-point MIMO link. This generalizes the finite-M result in [1] and shows that the use of relays as active scatterers can recover spatial multiplexing gain in poor scattering environments even if the number of transmit and receive antennas grows large. More importantly, our result shows that linear increase in the number of relays as a function of transmit-receive antenna pairs is sufficient for this to happen.

Appendix A
Some Essentials from Random Matrix Theory

In this section, we briefly summarize the basic definitions and results from random matrix theory used in this paper. An excellent tutorial on this subject is [3].

Definition 3 (Stieltjes transform): Let $F(x)$ be a distribution function with density $f(x)$. The analytic function

$$G_F(z) \triangleq \int \frac{f(x)}{x-z}\,dx, \quad z\in\mathbb{C}^+$$

is called the Stieltjes transform of $F(x)$.

Lemma 2 (Inversion formula): Let $G_F(z)$ be the Stieltjes transform of a distribution function $F(x)$. The corresponding density function can be obtained as

$$f(x) = \frac{1}{\pi}\lim_{y\to 0+}\Im\left[G_F(x+jy)\right]. \tag{21}$$

Theorem 1 (Silverstein [2]): Define the following quantities on a common probability space:

- The random matrix $\mathbf{X}\in\mathbb{C}^{N\times N'}$ has i.i.d. zero-mean entries with variance one.
- The random matrix $\mathbf{Y}\in\mathbb{C}^{N\times N}$ is Hermitian nonnegative definite with $F_{\mathbf{Y}}^{N}(x)$, for $N\to\infty$, converging on $[0,\infty)$ a.s. to a nonrandom distribution function $F_{\mathbf{Y}}(x)$ with corresponding density $f_{\mathbf{Y}}(x)$.

Assume that the matrices $\mathbf{X}$ and $\mathbf{Y}$ are statistically independent. Then, for $N,N'\to\infty$ with $N/N'\to\beta$, $F^N_{(1/N')\mathbf{X}\mathbf{X}^H\mathbf{Y}}(x) \xrightarrow{\text{a.s.}} F_{(1/N')\mathbf{X}\mathbf{X}^H\mathbf{Y}}(x)$ with its Stieltjes transform $G_{F_{(1/N')\mathbf{X}\mathbf{X}^H\mathbf{Y}}}(z)$ satisfying $(z\in\mathbb{C}^+)$

$$G_{F_{(1/N')\mathbf{X}\mathbf{X}^H\mathbf{Y}}}(z) = \\ = \int_{-\infty}^{\infty} \frac{f_{\mathbf{Y}}(x)dx}{x(1-\beta-\beta z\, G_{F_{(1/N')\mathbf{X}\mathbf{X}^H\mathbf{Y}}}(z)) - z}.$$

The solution of this fixed-point equation is unique in the set

$$\left\{ G_{F_{(1/N')\mathbf{X}\mathbf{X}^H\mathbf{Y}}} \in\mathbb{C} \,\middle|\, -\frac{1-\beta}{z} + \beta\, G_{F_{(1/N')\mathbf{X}\mathbf{X}^H\mathbf{Y}}} \in \mathbb{C}^+ \right\}.$$

We shall furthermore use the Marčenko-Pastur law as stated in [7].

Theorem 2 (Marčenko-Pastur [8]): Assume that the matrix $\mathbf{X}\in\mathbb{C}^{N,N'}$ has i.i.d. zero-mean entries with variance d^2. Then, for $N,N'\to\infty$ with $N'/N\to\beta$, the ESD of $(1/N)\mathbf{X}\mathbf{X}^H$ converges a.s. to a limiting distribution function with density

$$f_{(1/N)\mathbf{X}\mathbf{X}^H}(x) = \\ = \frac{\beta}{2\pi x d^2}\sqrt{(\gamma_2-x)^+(x-\gamma_1)^+} + [1-\beta]^+\delta(x)$$

where $\gamma_1 = d^2(1-1/\sqrt{\beta})^2$ and $\gamma_2 = d^2(1+1/\sqrt{\beta})^2$.

Under the same assumptions as in the first statement, if, in addition, the entries of $\mathbf{X}$ have finite fourth moments, then a.s.

$$\lim_{N'\to\infty}\lambda_{\min}\left(\frac{1}{N}\mathbf{X}\mathbf{X}^H\right) = \gamma_1$$
$$\lim_{N'\to\infty}\lambda_{\max}\left(\frac{1}{N}\mathbf{X}\mathbf{X}^H\right) = \gamma_2.$$

Appendix B
Computation of the Integral $\hat{I}$ in (12)

In the following, we detail the computation of the integral

$$\hat{I} \triangleq \rho\int_{\eta_1}^{\eta_2} \frac{\sqrt{(\eta_2-x)(x-\eta_1)}\,dx}{x(1-x)^2\left(x\left(\frac{1-\beta}{z}-\beta G\right)-1\right)}$$

on the RHS of (12). With the change of variables

$$t = \sqrt{\frac{x - \eta_1}{\eta_2 - x}}$$

and the notation

$$\begin{aligned}\mu_1 &\triangleq 1 - \eta_1 \\ \mu_2 &\triangleq 1 - \eta_2 \\ \nu_1 &\triangleq \eta_1 \left(\frac{1-\beta}{z} - \beta G \right) - 1 \\ \nu_2 &\triangleq \eta_2 \left(\frac{1-\beta}{z} - \beta G \right) - 1\end{aligned}$$

the integral $\hat{I}$ can be written as

$$\hat{I} = 2(\eta_2 - \eta_1)^2 \rho \int_0^\infty \frac{t^2(t^2+1)dt}{(\eta_2 t^2 + \eta_1)(\mu_2 t^2 + \mu_1)^2 (\nu_2 t^2 + \nu_1)}.$$

To simplify further, we introduce the notation

$$\kappa_1 \triangleq -\frac{\eta_1}{\eta_2}, \quad \kappa_2 \triangleq -\frac{\mu_1}{\mu_2}, \quad \kappa_3 \triangleq -\frac{\nu_1}{\nu_2}, \quad \chi \triangleq \frac{2(\eta_2 - \eta_1)^2}{\eta_2\, \mu_2^2\, \nu_2} \rho$$

so that

$$\hat{I} = \chi \int_0^\infty \frac{t^2(t^2+1)dt}{(t^2 - \kappa_1)(t^2-\kappa_2)^2(t^2-\kappa_3)}. \tag{22}$$

Upon partial fraction expansion of the integrand in (22), we obtain

$$\hat{I} = \chi(A_1 \hat{I}_1 + A_2 \hat{I}_2 + A_3 \hat{I}_3 + A_4 \hat{I}_4)$$

where

$$\hat{I}_1 \triangleq \int_0^\infty \frac{dt}{t^2 - \kappa_1} \qquad \hat{I}_2 \triangleq \int_0^\infty \frac{dt}{(t^2 - \kappa_2)^2}$$
$$\hat{I}_3 \triangleq \int_0^\infty \frac{dt}{t^2 - \kappa_2} \qquad \hat{I}_4 \triangleq \int_0^\infty \frac{dt}{t^2 - \kappa_3} \tag{23}$$

with

$$A_1 = \frac{\kappa_1(\kappa_1+1)}{(\kappa_1-\kappa_2)^2(\kappa_1-\kappa_3)} \tag{24}$$
$$A_2 = \frac{\kappa_2(\kappa_2+1)}{(\kappa_2-\kappa_1)(\kappa_2-\kappa_3)} \tag{25}$$
$$A_3 = \frac{-\kappa_2^2 - \kappa_1\kappa_2^2 + \kappa_1\kappa_3 + 2\kappa_1\kappa_2\kappa_3 - \kappa_2^2\kappa_3}{(\kappa_2-\kappa_1)^2(\kappa_2-\kappa_3)^2} \tag{26}$$
$$A_4 = \frac{\kappa_3(\kappa_3+1)}{(\kappa_3-\kappa_1)(\kappa_3-\kappa_2)^2}. \tag{27}$$

The integrals in (23) can be evaluated as follows

$$\hat{I}_1 = \frac{1}{\sqrt{-\kappa_1}} \arctan \frac{t}{\sqrt{-\kappa_1}} \bigg|_0^\infty = \frac{\pi}{2\sqrt{-\kappa_1}} \tag{28}$$
$$\begin{aligned}\hat{I}_2 &= -\frac{t}{2\kappa_2(t^2-\kappa_2)}\bigg|_0^\infty - \frac{1}{2\kappa_2\sqrt{-\kappa_2}} \arctan \frac{t}{\sqrt{-\kappa_2}}\bigg|_0^\infty = \\ &= -\frac{\pi}{4\kappa_2\sqrt{-\kappa_2}}\end{aligned} \tag{29}$$
$$\hat{I}_3 = \frac{1}{\sqrt{-\kappa_2}} \arctan \frac{t}{\sqrt{-\kappa_2}} \bigg|_0^\infty = \frac{\pi}{2\sqrt{-\kappa_2}} \tag{30}$$
$$\hat{I}_4 = \frac{1}{\sqrt{-\kappa_3}} \arctan \frac{t}{\sqrt{-\kappa_3}} \bigg|_0^\infty = \frac{\pi}{2\sqrt{-\kappa_3}}. \tag{31}$$

The quantity κ_3 is complex-valued, and the $\arctan$ and square root in (30) are understood as the principal values of these functions in $\mathbb{C}$ as defined in [9].

Finally, by inspection, combining (28)–(31) with (24)–(27) and resubstituting the values of the parameters $\kappa_1, \kappa_2, \kappa_3, \chi, \rho, \mu_1, \mu_2, \eta_1, \eta_2, \nu_1, \nu_2, \gamma_1$, and γ_2, after straightforward but tedious simplifications, we find

$$\chi A_1 \hat{I}_1 = \frac{\left(\sqrt{\beta}+1\right)\left|\sqrt{\beta}-1\right|}{2\beta}$$
$$\chi A_2 \hat{I}_2 = -\frac{z}{\sqrt{\beta}(G\beta z + z + \beta - 1)}$$
$$\chi A_3 \hat{I}_3 = \frac{-z\left(\sqrt{\beta}-1\right)^2}{2\beta(G\beta z + z + \beta - 1)} + \frac{z(G\beta z + \beta - 1)}{2d^2(G\beta z + z + \beta - 1)^2}$$
$$\begin{aligned}\chi A_4 \hat{I}_4 = &-\frac{(G\beta z + \beta - 1)\sqrt{\frac{d^2(G\beta z+z+\beta-1)\left(\sqrt{\beta}-1\right)^2+z\beta}{d^2(G\beta z+z+\beta-1)\left(\sqrt{\beta}+1\right)^2+z\beta}}}{2d^2\beta(G\beta z + z + \beta - 1)^2} \times \\ &\times \left(d^2 (G\beta z + z + \beta - 1)\left(\sqrt{\beta}+1\right)^2 + z\beta \right).\end{aligned}$$

References

[1] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, Jun. 2006.

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

[3] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Commun. and Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[4] R. R. Müller, "Applications of large random matrices in communications engineering," in *Proc. Int. Conf. on Advances Internet, Process., Syst., Interdisciplinary Research (IPSI)*, Sveti Stefan, Montenegro, Oct. 2003.

[5] A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–71, Jan. 1978.

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

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

[8] V. A. Marčenko and L. A. Pastur, "Distribution of some sets of random matrices," *Math. USSR-Sb*, vol. 1, pp. 457–483, 1967.

[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. Dover Publications, 1964.