ON SLOW-FADING NON-SEPARABLE CORRELATION MIMO SYSTEMS

R. RASHIDI FAR, T. ORABY, W. BRYC*, AND R. SPEICHER+

Abstract. In a frequency selective slow-fading channel in a MIMO system, the channel matrix is of the form of a block matrix. We propose a method to calculate the limit of the eigenvalue distribution of block matrices if the size of the blocks tends to infinity. We will also calculate the asymptotic eigenvalue distribution of HH^*, where the entries of H are jointly Gaussian, with a correlation of the form $E[h_{pj}\bar{h}_{qk}] = \sum_{s=1}^{t} \Psi_{jk}^{(s)} \hat{\Psi}_{pq}^{(s)}$ (where t is fixed and does not increase with the size of the matrix). We will use an operator-valued free probability approach to achieve this goal. Using this method, we derive a system of equations, which can be solved numerically to compute the desired eigenvalue distribution.

Keywords: MIMO systems, channel models, eigenvalue distribution, fading channels, free probability, Cauchy transform, intersymbol interference, random matrices, channel capacity.

1. Introduction

With the introduction of some sophisticated communication techniques such as CDMA (Code-Division Multiple-Access) and MIMO (Multiple-Input Multiple-Output), the communications community has been looking into analyzing different aspects of these systems, ranging from the channel capacity to the structure of the receiver. It has been shown that the channel matrix plays a key role in the capacity of the channel [1, 2] as well as in the structure of the optimum receiver [3, 4].

* Research partially supported by NSF grant #DMS-0504198.

+ Research supported by a Discovery Grant and a Leadership Support Initiative Award from the Natural Sciences and Engineering Research Council of Canada and by a Killam Fellowship from the Canada Council for the Arts.

R. Rashidi Far and R. Speicher are with the Department of Mathematics and Statistics, Queen's University, Ontario, Canada K7L 3N6 reza, speicher@mast.queensu.ca , T. Oraby and W. Bryc are with the Department of Mathematical Sciences, University of Cincinnati, 28855, Campus Way PO Box 210025, Cincinnati, OH 45221-0025, USA, orabyt@math.uc.edu, wlodzimierz.bryc@uc.edu
.

More precisely, the eigenvalue distribution of the channel matrix is the factor of interest in different applications.

Free probability [5, 6, 7] and random matrix theory have proven to provide the right kind of tools in tackling such kind of problems [8, 9]. For example, Tse and Zeitouni [10] applied random matrix theory to study linear multiuser receivers, Moustakas *et. al.* [11] applied it to calculate the capacity of a MIMO channel. Müller Muller-02,Muller-02a employed it in calculating the eigenvalue distribution of a particular fading channel and later Debbah and Müller [14] applied it in MIMO channel modeling.

There are, however, also many interesting (more realistic) models for the channel matrix, which are not directly accessible with the usual free probability or random matrix techniques. Let us be a bit more specific on such examples. For a MIMO wireless system with n_T transmitter antenna and n_R receiver antenna, the received signal at time index n, $Y_n = [y_{1,n}, \cdots, y_{n_R,n}]^T$, will be as follows:

$$Y_n = HX_n + N_n, \tag{1}$$

where H is the channel matrix, $X_n = [x_{1,n}, \cdots, x_{n_T,n}]^T$ is the transmitted signal at time n and N_n is the noise signal. The channel matrix entries h_{ij} reflect the channel effect on the signal transmitted from antenna j in the transmitter and received at antenna i in the receiver. In a more realistic channel modeling, one may consider the Intersymbol-Interference (ISI) [15, Chapter 2]. In this case, the channel impulse response between the transmitter antenna j and the receiver antenna i is a vector $h_{ij} = \begin{bmatrix} h_1^{(ij)} & h_2^{(ij)} & \cdots & h_{L-1}^{(ij)} & h_L^{(ij)} \end{bmatrix}^T$ where L is the length of the impulse response of the channel (number of the taps). Consequently, the channel matrix for a signal frame of K will be as follows:

$$H = \begin{bmatrix} A_1 & A_2 & \cdots & A_L & \mathbf{0} & \mathbf{0} & & \cdots & \mathbf{0} \\ \mathbf{0} & A_1 & A_2 & \cdots & A_L & \mathbf{0} & & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_1 & A_2 & \cdots & A_L & \mathbf{0} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \cdots & & \ddots & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} & A_1 & A_2 & \cdots & A_L \end{bmatrix}, \tag{2}$$

where there are $K-1$ zero-matrices in each row and $A_l = (h_l^{(ij)})_{\substack{i=1,\cdots,n_R \\ j=1,\cdots,n_T}}$ (see Fig. 1 for the block diagram). To calculate the capacity of such a channel, one needs to know the eigenvalue distribution of the HH^* [8].

The above random matrix falls into the class of random matrices where one has correlations between the entries of H. Whereas random matrices with independent entries are quite well understood and there exist many analytic results on their asymptotic eigenvalue distribution, only very special cases of the correlated situation could be treated in the literature. The most prominent of those is the case of separable correlation, where the covariance between the entries of $H = (h_{ij})_{i,j}$ factorizes as $E[h_{pj}\bar{h}_{qk}] = \Psi^T_{jk}\Psi^R_{pq}$, where Ψ^T and Ψ^R are matrices describing the transmit and the receive correlation, respectively. Our block matrix H from Eq. (2) does not fall into this class.

In this paper, we will show how a more general version of free probability theory, so-called "operator valued free probability theory" allows to deal with more general situations of correlated entries. In particular, we will treat the case of block matrices, as the above H from Eq. (2), and also extend results from [16] from the case of separable correlations to the more general situation

$$E[h_{pj}\bar{h}_{qk}] = \sum_{s=1}^{t} \Psi^{(s)}_{jk}\hat{\Psi}^{(s)}_{pq}. \tag{3}$$

The results for the asymptotic eigenvalue distribution of HH^* for these two cases are stated in Section 2, in Theorems 1 and 2. The proof of Theorem 1 is given in Section 3, as a consequence of a corresponding statement, Theorem 3, for selfadjoint matrices X. We will show that these selfadjoint matrices are asymptotically described by operator-valued semicircular elements, and the equation describing the limiting Cauchy transform of X follows then from the general theory of semicircular elements. In Appendix I we state the main notions and results in relation with operator-valued semicircular elements. In Appendix II we prove Theorem 2, again by showing that a corresponding selfadjoint matrix X is asymptotically an operator-valued semicircular element. In Appendix III we state a more general version of Theorem 3, for the situation where the blocks are not necessarily square matrices.

2. Asymptotic eigenvalue distribution of HH^*

In this section we will present our main results on the asymptotic eigenvalue distribution for HH^* where H is a non-selfadjoint Gaussian random matrix with some specific kind of correlation between its entries. We will treat the block matrix case and the case of non-separable correlation. The proof of these theorems will be provided in the next section and in the appendix. An application to the asymptotic eigenvalue distribution of channel matrices of the form (2) will be given in Section 4.

2.1. HH^* **for block matrices.** Our main theorem on block channel matrices is the following. For used notation, see Section 5.1.

Theorem 1. *Fix natural* b *and* a *and a real valued "covariance function"* $\tau(i,k;j,l)$ *such that* $\tau(i,k;j,l) = \tau(j,l;i,k)$, $i,j = 1,\ldots,a; k,l = 1,\ldots,b$. *Assume, for* $N \in \mathbb{N}$, *that* $\{h_{rp}^{(i,k)} \mid i = 1,\ldots,a,\ k = 1,\ldots,b,\ r,p = 1,\ldots,N\}$ *are jointly Gaussian complex random variables, with the prescription of mean zero and covariance*

$$E[h_{rp}^{(i,k)} \bar{h}_{sq}^{(j,l)}] = \frac{1}{(b+a)N} \delta_{rs}\delta_{pq} \cdot \tau(i,k;j,l). \tag{4}$$

We also assume circular complex Gaussian law, i.e, $E[(h_{rp}^{(i,k)})^2] = 0$.

Consider now block matrices $H_N = (H^{(i,k)})_{\substack{i=1,\ldots,a \\ k=1,\ldots,b}}$, *where, for each* $i = 1,\ldots,a$ *and* $k = 1,\ldots,b$, *the blocks are given by* $H^{(i,k)} = \left(h_{rp}^{(i,k)}\right)_{r,p=1}^N$.

Then, for $N \to \infty$, *the* $aN \times aN$ *matrix* $H_N H_N^*$ *has almost surely a limiting eigenvalue distribution whose Cauchy transform* $G(z)$ *is determined by*

$$G(z) = \mathrm{tr}_a(\mathcal{G}_1(z)),$$

where $\mathcal{G}_1(z)$ *is an* $M_a(\mathbb{C})$*-valued analytic function on the upper complex half plane, which is uniquely determined by the facts that*

$$\lim_{|z|\to\infty, \Im(z)>0} z\mathcal{G}_1(z) = I_a, \tag{5}$$

and that it satisfies for all z *in the upper complex half plane the matrix equation*

$$z\mathcal{G}_1(z) = I_a + \eta_1\Big(\big(I_b - \eta_2(\mathcal{G}_1(z))\big)^{-1}\Big) \cdot \mathcal{G}_1(z), \tag{6}$$

where

$$\eta_1 : M_b(\mathbb{C}) \to M_a(\mathbb{C}) \qquad \textit{and} \qquad \eta_2 : M_a(\mathbb{C}) \to M_b(\mathbb{C})$$

are the covariance mappings given by

$$[\eta_1(D)]_{ij} := \frac{1}{b+a} \sum_{k,l=1}^{b} \tau(i,k;j,l) \cdot [D]_{kl} \tag{7}$$

and

$$[\eta_2(D)]_{kl} := \frac{1}{b+a} \sum_{i,j=1}^{a} \tau(i,k;j,l) \cdot [D]_{ji}. \tag{8}$$

The proof of this theorem will be given in Section 3, by reducing it to Theorem 3.

In Section 4 we will use this to analyze the asymptotic eigenvalue distribution of HH^* for the channel matrix from Eq. (2).

2.2. HH^* **for non-separable correlated fading.** In [16], MIMO wireless systems under correlated fading were analyzed by asymptotic analysis of the eigenvalue distribution of $H_n H_n^*$, where the entries of the $n \times n$ random matrix $H = 1/\sqrt{n}(h_{ij})_{i,j=1}^n$ were assumed as jointly Gaussian with the following covariance structure:

$$E[h_{pj}\bar{h}_{qk}] = \Psi^T_{jk}\Psi^R_{pq},$$

where Ψ^T and Ψ^R are Hermitian positive-definite matrices describing the transmit and the receive correlation, respectively. The assumption on Ψ^T and Ψ^R is that both have a limiting eigenvalue distribution.

We will now show how operator-valued free probability theory can be used to analyze a generalization of this to the case

$$E[h_{pj}\bar{h}_{qk}] = \sum_{s=1}^{t} \Psi^{(s)}_{jk}\hat{\Psi}^{(s)}_{pq}. \tag{9}$$

The number t of summands is here fixed and does not depend on n. As before, one needs the existence of the limiting joint distribution of the Ψ's and the limiting joint distribution of the $\hat{\Psi}$'s. Mixed moments in Ψ and $\hat{\Psi}$ do not play a role for the result on HH^*.

This situation is treated in the next theorem, which we will prove in Appendix II, Section 6.2. Some of the basic notions from free probability which are used in the formulation of the theorem are defined in Appendix I. As in [16] we will restrict here, for notational simplicity, to the case of a square H. By invoking ideas from [17], one can also extend the results to rectangular H.

Theorem 2. *Assume that h_{ij} $(i, j \in \mathbb{N})$ are jointly Gaussian complex random variables with mean zero and covariance given by (9) for some $t \geq 1$ and some positive-definite matrices $\Psi^{(s)}_{ij}$ and $\hat{\Psi}^{(s)}_{ij}$ $(s = 1, \dots, t)$. We also assume circular complex Gaussian law, i.e, $E[(h_{rp})^2] = 0$. We assume that, as $n \to \infty$, the $\big((\Psi^{(s)}_{i,j}), (\hat{\Psi}^{(s)}_{i,j})\big)_{s=1,\dots,t}$ converge in distribution to some elements $(\Psi_s, \hat{\Psi}_s)_{s=1,\dots,t}$ in some non-commutative probability space $(\mathcal{B}, \varphi)$.*

We denote by $\mathcal{B}_1 \subset \mathcal{B}$ the algebra generated by $\Psi_1, \dots, \Psi_t$ and by $\mathcal{B}_2 \subset \mathcal{B}$ the algebra generated by $\hat{\Psi}_1, \dots, \hat{\Psi}_t$. Furthermore we define

$$\eta_1 : \mathcal{B}_1 \to \mathcal{B}_2, \qquad \eta_1(b) := \sum_{s=1}^{t} \hat{\Psi}_s \varphi(b\Psi_s) \tag{10}$$

and

$$\eta_2 : \mathcal{B}_2 \to \mathcal{B}_1, \qquad \eta_2(b) := \sum_{s=1}^{t} \Psi_s \varphi(b\hat{\Psi}_s). \tag{11}$$

We consider now

$$H_n := \frac{1}{\sqrt{n}} \big(h_{ij}\big)_{i,j=1}^{n}$$

Then the eigenvalue distribution of $H_n H_n^*$ converges almost surely to a limiting distribution whose Cauchy transform G is given by $G(z) = \varphi(\mathcal{G}_1(z))$, where $\mathcal{G}_1$ is the solution of the equation

$$z\mathcal{G}_1(z) = \mathbf{id} + \sum_{s_1=1}^{t} \Psi_{s_1} \varphi\left(\Big(\mathbf{id} - \sum_{s_2=1}^{t} \hat{\Psi}_{s_2} \varphi\big(\mathcal{G}_1(z)\Psi_{s_2}\big)\Big)^{-1} \hat{\Psi}_{s_1} \right) \cdot \mathcal{G}_1(z). \tag{12}$$

One should note that the solution $\mathcal{G}_1(z)$ of the above fixed point equation lies in the algebra $\mathcal{B}_1$ and that its value does not depend on mixed moments between the Ψ_s's and the $\hat{\Psi}_s$'s. By results from [18] (as outlined in Appendix I, Section 5.4), there exists, for each $z \in \mathbb{C}^+$ a unique solution of equation (12) with the right positivity property.

Note also that Eq. (12) reduces in the case $t = 1$ to the fixed point equation in Theorem IV.2 in [16].

3. Asymptotic eigenvalue distribution for selfadjoint block matrices

Our Theorem 1 on the asymptotic eigenvalue distribution of HH^* for a block matrix H follows from a corresponding statement for a selfadjoint block matrix X, which also has Gaussian entries with correlations. The reduction to the selfadjoint case can be achieved by the well-known trick of going over to

$$X = \begin{bmatrix} 0 & H \\ H^* & 0 \end{bmatrix}. \tag{13}$$

In this section we will state the selfadjoint version of Theorem 1 and show how it implies the result for HH^*.

3.1. **Selfadjoint block matrices.** Let us consider the selfadjoint version of Theorem 1. We will here restrict to the situation where all blocks are square matrices of the same size. For some applications it might actually be better to allow also blocks of a rectangular size (which, of course, have to fit together to form a big square matrix). There is a straightforward generalization of the following theorem to that situation; we will state this in Appendix III.

Theorem 3. *Fix a natural d and a "covariance function" σ which satisfies*

$$\sigma(i,j;k,l) = \overline{\sigma(k,l;i,j)} \tag{14}$$

for all $i,j,k,l = 1,\dots,d$. Assume, for $N \in \mathbb{N}$, that $\{a_{rp}^{(i,j)} \mid i,j = 1,\dots,d,\ r,p = 1,\dots,N\}$ are jointly Gaussian random variables, with

$$a_{rp}^{(i,j)} = \overline{a_{pr}^{(j,i)}} \qquad \text{for all } i,j = 1,\dots,d,\ r,p = 1,\dots,N$$

and the prescription of mean zero and covariance

$$E[a_{rp}^{(i,j)} a_{qs}^{(k,l)}] = \frac{1}{dN}\delta_{rs}\delta_{pq} \cdot \sigma(i,j;k,l). \tag{15}$$

Consider now block matrices $X_N = (A^{(i,j)})_{i,j=1}^d$, where, for each $i,j = 1,\dots,d$, the blocks are given by $A^{(i,j)} = \big(a_{rp}^{(i,j)}\big)_{r,p=1}^N$.

Then, for $N \to \infty$, the $dN \times dN$ matrix X_N has almost surely a limiting eigenvalue distribution whose Cauchy transform $G(z)$ is determined by

$$G(z) = \mathrm{tr}_d(\mathcal{G}(z)), \tag{16}$$

where $\mathcal{G}(z)$ is an $M_d(\mathbb{C})$-valued analytic function on the upper complex half plane, which is uniquely determined by the facts that

$$\lim_{|z|\to\infty, \Im(z)>0} z\mathcal{G}(z) = I_d, \tag{17}$$

and that it satisfies for all z in the upper complex half plane the matrix equation

$$z\mathcal{G}(z) = I_d + \eta(\mathcal{G}(z)) \cdot \mathcal{G}(z), \tag{18}$$

where $\eta : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ is the covariance mapping

$$\big[\eta(D)_{i,j=1}^d\big]_{ij} := \frac{1}{d}\sum_{k,l=1}^d \sigma(i,k;l,j) \cdot [D]_{kl}. \tag{19}$$

The proof of Theorem 3 is given in Appendix II. Let us here just point out that the determining equation (18) is actually the equation for an operator-valued semicircular element; thus, one essentially has to realize that X_N converges to a suitably chosen operator-valued semicircular element.

Theorem 3 has also some interest of its own; for an application to some selfadjoint block matrix problems from [19] see [20]. Here we will just use it to prove our Theorem 1

3.2. **Proof of Theorem 1.** Let us consider matrices H_N as in Theorem 1. For clarity of notation, we will in the following suppress the index N. The calculation of the eigenvalue distribution of HH^* can be reduced to the situation treated in the previous section by the following trick. Consider

$$X = \begin{bmatrix} 0 & H \\ H^* & 0 \end{bmatrix}.$$

With $d = b+a$, this is a selfadjoint $dN \times dN$-matrix and can be viewed as a $d \times d$-block matrix of the form considered in Theorem 3; thus we can use this to get the asymptotic eigenvalue distribution of X.

The only remaining question is how to relate the eigenvalues of X with those of HH^*. This is actually quite simple, we only have to note that all the odd moments of X are zero and

$$X^2 = \begin{bmatrix} HH^* & 0 \\ 0 & H^*H \end{bmatrix}$$

Thus the eigenvalues of X^2 are the eigenvalues of HH^* together with the eigenvalues of H^*H. (One might also note HH^* is an $aN \times aN$ and H^*H is an $bN \times bN$ matrix. Assuming that $a < b$ (otherwise exchange the role of H and H^*) we have then that the eigenvalues of H^*H are the eigenvalues of HH^* plus $(b-a)N$ additional zeros. However, we will not need this information in the following.)

So we should rewrite our equation for the Cauchy transform G_X of X in terms of the Cauchy transform G_{X^2} of X^2. Since X is even, both are related by

$$z \cdot G_{X^2}(z^2) = G_X(z).$$

By noting that the operator-valued Cauchy transform $\mathcal{G}(z)$ of X depends, up to an overall factor $1/z$, only on z^2, we can introduce a quantity $\mathcal{H}$ by

$$z \cdot \mathcal{H}(z^2) = \mathcal{G}(z).$$

Then with $n = dN$ we have

$$\lim_{n\to\infty} G_{X^2}(z) = \mathrm{tr}_d[\mathcal{H}(z)], \tag{20}$$

and the equation (18) for $\mathcal{G}$ becomes

$$z\mathcal{H}(z) = I_d + z\eta\big(\mathcal{H}(z)\big) \cdot \mathcal{H}(z). \tag{21}$$

It is fairly easy to see that the covariance mapping $\eta : M_{b+a}(\mathbb{C}) \to M_{b+a}(\mathbb{C})$ of X splits according to

$$\eta : \begin{bmatrix} D_1 & D_3 \\ D_4 & D_2 \end{bmatrix} \mapsto \begin{bmatrix} \eta_1(D_2) & 0 \\ 0 & \eta_2(D_1) \end{bmatrix},$$

where

$$\eta_1 : M_b(\mathbb{C}) \to M_a(\mathbb{C}) \qquad \text{and} \qquad \eta_2 : M_a(\mathbb{C}) \to M_b(\mathbb{C})$$

are the two covariance mappings for H as in Theorem 1. Therefore, our $(b+a) \times (b+a)$ matrix $\mathcal{H}$ decomposes as a 2×2-block matrix

$$\mathcal{H}(z) = \begin{bmatrix} \mathcal{G}_1(z) & 0 \\ 0 & \mathcal{G}_2(z) \end{bmatrix}$$

where $\mathcal{G}_1$ and $\mathcal{G}_2$ are $M_a(\mathbb{C})$-valued and $M_b(\mathbb{C})$-valued, respectively, analytic functions in the upper complex half plane. Then one has

$$\lim_{N\to\infty} G_{HH^*}(z) = \mathrm{tr}_a\big(\mathcal{G}_1(z)\big). \tag{22}$$

and

$$\lim_{N\to\infty} G_{H^*H}(z) = \mathrm{tr}_b\big(\mathcal{G}_2(z)\big). \tag{23}$$

The equation (21) for $\mathcal{H}$ splits now into the two equations

$$z\mathcal{G}_1(z) = I_a + z\eta_1\big(\mathcal{G}_2(z)\big) \cdot \mathcal{G}_1(z)$$

and

$$z\mathcal{G}_2(z) = I_b + z\eta_2\big(\mathcal{G}_1(z)\big) \cdot \mathcal{G}_2(z).$$

One can eliminate $\mathcal{G}_2$ from those equations by solving the second equation for $\mathcal{G}_2$ and inserting this into the first equation, yielding Eq. (6).

4. Results and Discussion

Our theorems give us the Cauchy transform G of the asymptotic eigenvalue distribution HH^* of the considered matrices in the form $G(z) = \mathrm{tr}_r(\mathcal{G}_1(z))$, where $\mathcal{G}_1(z)$ is a solution to the matrix equation (6) or (12). Usually, it is more convenient to deal with the equation (21) for the corresponding selfadjoint matrix X.

We recover the corresponding eigenvalue distribution μ from G in the usual way, by invoking Stieltjes inversion formula

$$d\mu(x) = -\frac{1}{\pi} \lim_{\varepsilon \searrow 0} \Im G(x + i\varepsilon) dx, \tag{24}$$

where the limit is weak convergence of measures.

Usually, there is no explicit solution for our matrix equations, so that we have to rely on numerical methods for solving those. Note that we do not get directly an equation for G. We first have to solve the matrix equation, then take the trace of this solution. Thus, in terms of the entries of our matrix $\mathcal{G}_1$ or $\mathcal{H}$, we face a system of quadratic equations which we solve numerically, either by using Newton's algorithm [21] or by iterations as in Eq. (31).

4.1. **Example: ISI channel matrix.** In this section we want to specify our general theorems to the case of the ISI channel matrices as appearing in Eq. (2). For simplicity we treat the case of square blocks.

4.1.1. *Proposition.* Let H_N be the channel matrix from Eq. (2) with $n_R = n_T =: N$, such that each entry h_l^{ij} has variance 1. Put $d = 2K + L - 1$, $n = Nd$. As $N \to \infty$ the spectral law of $H_N H_N^*/n$ converges with probability one to a deterministic probability measure which is a mixture of K densities with Cauchy transform

$$\lim_{N\to\infty} G_{HH^*/n}(z) = \frac{1}{K}\sum_{j=1}^{K} f_j(z). \tag{25}$$

Functions f_j are each a Cauchy transform of a probability measure and the following conditions hold.

(1) $f_j = f_{K+1-j}$ for $1 \le j \le K$
(2) The diagonal matrix $\mathcal{G}_1 = \operatorname{diag}(f_1, \ldots, f_K)$ satisfies equation (6) with $\eta_1 : M_{K+L-1}(\mathbb{C}) \to M_K(\mathbb{C})$ given by

$$[\eta_1(D)]_{ij} = \frac{1}{L+2K-1}\sum_{k=1}^{K}[D]_{i+k,j+k},\ 1 \le i,j \le K \tag{26}$$

and with $\eta_2 : M_K(\mathbb{C}) \to M_{K+L-1}(\mathbb{C})$ such that on the diagonal we have

$$[\eta_2(D)]_{jj} = \frac{1}{L+2K-1}\sum_{k=\max\{1,j-L+1\}}^{\min\{j,K\}}[D]_{j+k,j+k},\ 1 \le j \le K+L-1. \tag{27}$$

Proof. We note that the only non-zero values of τ are $\tau(i,j;k,j+k-i) = 1$ when $1 \le i,k \le K$, $i \le j \le i+L-1$. Therefore (7) gives (26) and (8) gives (27).

From (26) and (27) we see that η maps diagonal matrices into diagonal matrices, so the solution $\mathcal{H}$ of equation (21) must be diagonal,

$$\mathcal{H}(z) = \operatorname{diag}(f_1, f_2, \ldots, f_K, g_1, g_2, \ldots, g_{K+L-1}).$$

We now note that the symmetry conditions $f_j = f_{K+1-j}$ and $g_j = g_{K+L-j}$ are preserved under the mapping $D \mapsto I_d + \eta(D)\cdot D$, therefore the same symmetries must be satisfied by the solution $\mathcal{H}$. Thus $\mathcal{G}_1 = \operatorname{diag}(f_1, \ldots, f_K)$ satisfies (6) and $f_j = f_{K+1-j}$ as claimed. □

4.1.2. *Example.* As a concrete example we consider a MIMO system with ISI ($L = 4$) and frame size of 4 ($K = 4$):

$$H_N = \begin{bmatrix} A & B & C & D & 0 & 0 & 0 \\ 0 & A & B & C & D & 0 & 0 \\ 0 & 0 & A & B & C & D & 0 \\ 0 & 0 & 0 & A & B & C & D \end{bmatrix}, \tag{28}$$

where A, B, C, D are independent non-selfadjoint Gaussian $N \times N$-random matrices. It is also assumed that the impulse response of the channel from any transmitter antenna to any receiver antenna is identical and equal to $\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$. In this case $K = L = 4$,

$$\mathcal{G}_1(z) = \operatorname{diag}(f_1(z), f_2(z), f_2(z), f_1(z)),$$

$$[\eta_1(D)]_{ii} = \frac{1}{11}\sum_{j=i}^{i+3}[D]_{jj} \quad \text{for } i = 1, 2,$$

$$\eta_2(\mathcal{G}_1) = \frac{1}{11}\operatorname{diag}(f_1, f_1 + f_2, f_1 + 2f_2, 2f_1 + 2f_2)$$

and (6) yields the following system of equations.

$$\begin{aligned} z &= \frac{1}{f_1} + \frac{1}{11 - f_1} + \frac{1}{11 - f_1 - f_2} + \frac{1}{11 - f_1 - 2f_2} + \frac{1}{11 - 2f_1 - 2f_2}, \\ z &= \frac{1}{f_2} + \frac{1}{11 - f_1 - f_2} + \frac{2}{11 - f_1 - 2f_2} + \frac{1}{11 - 2f_1 - 2f_2}. \end{aligned}$$

The limiting Cauchy transform is $G_{HH^*}(z) = (f_1 + f_2)/2$. We use Newton's algorithm to solve this quadratic system of equations; the match between this solution and simulations is shown in Fig. 2.

4.2. **Convergence speed of capacity.** The results developed in this manuscript are good assets to study the asymptotic behaviour of slow-fading non-separable correlation MIMO channels when $N \to \infty$ but the authenticity of these results for limited N is also of interest in practice.

In this subsection, the asymptotic capacity of a slow-fading MIMO channel with $L = 2$ and the frame length of $K = 2$ in different SNR is compared with the capacity of such a channel for several N. The channel matrix for this system is as follows:

$$H = \begin{bmatrix} A & B & 0 \\ 0 & A & B \end{bmatrix},$$

and the results are depicted in Fig. 3. As the figure shows, with increasing the size of the blocks, the system capacity fast approaches the asymptotic capacity, suggesting a reasonable match between the asymptotic capacity and the capacity with a block size of 10 and bigger.

5. Appendix I: Prerequisites

5.1. **Notations.** The following notations are used in the paper:

$M_d(\mathbb{C})$	complex $d \times d$ matrices
$M_d(\mathcal{A})$	$d \times d$ matrices with entries from the algebra $\mathcal{A}$
$[D]_{ij}$	i, j-entry of the matrix D
tr_d	normalized trace on $M_d(\mathbb{C})$
$\Im(X)$	Imaginary part of X
I_d	$d \times d$ Identity matrix
$\mathbf{id}$	identity operator on a Hilbert space
$\overline{X}$	complex conjugate of X
δ_{ij}	Dirac delta function
X^*	Hermitian conjugate of matrix X
$\mathbb{C}^+$	complex upper half plane

The Cauchy transform of a probability measure μ on $\mathbb{R}$ is defined by

$$G(z) = \int_{\mathbb{R}} \frac{1}{z-t} d\mu(t) \qquad (z \in \mathbb{C}^+).$$

5.2. **(Operator-valued) non-commutative probability spaces and freeness.** A pair $(\mathcal{A}, \varphi)$ consisting of a unital algebra and a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ with $\varphi(1) = 1$ is called a non-commutative probability space. If $\mathcal{B}$ is a subalgebra of $\mathcal{A}$, then a mapping $E : \mathcal{A} \to \mathcal{B}$ is called a conditional expectation if we have for all $a \in \mathcal{A}$ and $b_1, b_2 \in \mathcal{B}$ that

$$E[b_1 a b_2] = b_1 E[a] b_2.$$

An algebra $\mathcal{A}$ with a conditional expectation onto a subalgebra $\mathcal{B}$ is called a $\mathcal{B}$-valued probability space.

If we are given such a $\mathcal{B}$-valued probability space then we say that unital subalgebras $\mathcal{A}_i \subset \mathcal{A}$ $(i \in I)$ are free over $\mathcal{B}$ (or with respect to E) if the following is satisfied: whenever we have $a_1, \dots, a_n \in \mathcal{A}$ such that $a_j \in \mathcal{A}_{i(j)}$ $(i(1), \dots, i(n) \in I)$ with $i(1) \neq i(2)$, $i(2) \neq i(3)$, $\dots$, $i(n-1) \neq i(n)$ and with $E[a_j] = 0$ for all $j = 1, \dots, n$, then we also have that $E[a_1 \cdots a_n] = 0$. In the case that $\mathcal{B} = \mathbb{C}$ (i.e., E is just a linear functional φ) we say that the $\mathcal{A}_i$ are free.

Elements in $\mathcal{A}$ are called free (over $\mathcal{B}$), if the algebras generated by them are free (over $\mathcal{B}$); they are called $*$-free (over $\mathcal{B}$), if the $*$-algebras generated by them are free (over $\mathcal{B}$).

5.3. **Convergence in distribution.** Let $(\mathcal{A}_N, \varphi_N)$ $(N \in \mathbb{N})$ and $(\mathcal{A}, \varphi)$ be non-commutative probability spaces. Let I be an index set and consider for each $i \in I$ random variables $a_N^{(i)} \in \mathcal{A}_N$ and $a_i \in \mathcal{A}$. We say that $(a_N^{(i)})_{i\in I}$ converges in distribution to $(a_i)_{i\in I}$ and denote this by

$$(a_N^{(i)})_{i\in I} \xrightarrow{\text{distr}} (a_i)_{i\in I},$$

if we have that each joint moment of $(a_N^{(i)})_{i\in I}$ converges to the corresponding joint moment of $(a_i)_{i\in I}$, i.e. if we have for all $n\in\mathbb{N}$ and all $i(1),\dots,i(n)\in I$

$$\lim_{N\to\infty} \varphi_N(a_N^{(i(1))}\cdots a_N^{(i(n))}) = \varphi(a_{i(1)}\cdots a_{i(n)}). \tag{29}$$

We say that $(a_N^{(i)})_{i\in I}$ converges in $*$-distribution to $(a_i)_{i\in I}$ if $(a_N^{(i)}, a_N^{(i)*})_{i\in I}$ converges in distribution to $(a_i, a_i^*)_{i\in I}$.

5.4. **Operator-valued semicircular elements.** Let $(\mathcal{A}, E:\mathcal{A}\to\mathcal{B})$ be a $\mathcal{B}$-valued probability space and let, in addition, be given a linear mapping $\eta:\mathcal{B}\to\mathcal{B}$. Then an element $s\in\mathcal{A}$ is called a $\mathcal{B}$-valued operator-valued semicircular element with covariance mapping η if one has $E[sbs]=\eta(b)$ for all $b\in\mathcal{B}$ and, more generally, for all $m\in\mathbb{N}$ and all $b_1,\dots,b_{m-1}\in\mathcal{B}$ that

$$E[sb_1s\cdots sb_{m-1}s] = \sum_{\pi\in NC_2(m)} \eta_\pi[b_1,\dots,b_{m-1}],$$

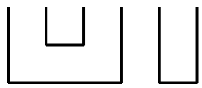
where $NC_2(m)$ are the non-crossing pairings of m elements (for details on non-crossing pairings in the context of free probability see [7]) and where η_π is given by an iterated application of the mapping η according to the nesting of the blocks of π. If one identifies a non-crossing pairing with a putting of brackets at the positions of the s's, then the way that η has to be iterated is quite obvious. To make this clear, let us consider as an example just the contribution of the five non-crossing pairings of six elements to the sixth moment. The latter is given by

$$\begin{aligned} E[sb_1sb_2sb_3sb_4sb_5s] &= \eta(b_1)\cdot b_2\cdot\eta(b_3)\cdot b_4\cdot\eta(b_5)\\ &\quad+\eta(b_1)\cdot b_2\cdot\eta\big(b_3\cdot\eta(b_4)\cdot b_5\big)+\eta\Big(b_1\cdot\eta\big(b_2\cdot\eta(b_3)\cdot b_4\big)\cdot b_5\Big)\\ &\quad+\eta\big(b_1\cdot\eta(b_2)\cdot b_3\big)\cdot b_4\cdot\eta(b_5)+\eta\big(b_1\cdot\eta(b_2)\cdot b_3\cdot\eta(b_4)\cdot b_5\big), \end{aligned}$$

corresponding to:

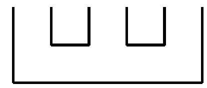
$\eta(b_1)\cdot b_2\cdot\eta(b_3)\cdot b_4\cdot\eta(b_5)$ $\qquad$ $\eta(b_1)\cdot b_2\cdot\eta\big(b_3\cdot\eta(b_4)\cdot b_5\big)$

$sb_1sb_2sb_3sb_4sb_5s$



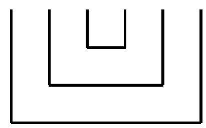
$\eta\big(b_1 \cdot \eta(b_2) \cdot b_3\big) \cdot b_4 \cdot \eta(b_5)$

$sb_1sb_2sb_3sb_4sb_5s$



$\eta\big(b_1 \cdot \eta(b_2) \cdot b_3 \cdot \eta(b_4) \cdot b_5\big)$

$sb_1sb_2sb_3sb_4sb_5s$



$\eta\Big(b_1 \cdot \eta\big(b_2 \cdot \eta(b_3) \cdot b_4\big) \cdot b_5\Big)$

For the rigorous definition of η_π and more details on operator-valued semicircular elements, we refer to [22].

In the situations which are relevant to us, the algebras $\mathcal{A}$ and $\mathcal{B}$ are operator algebras of bounded operators on Hilbert spaces. In such a case, the main statement about an operator-valued semicircular element s is the following description of its operator-valued Cauchy transform. Define

$$\mathcal{G} : \mathbb{C}^+ \to \mathcal{B}, \qquad \mathcal{G}(z) := E[\frac{1}{z-s}].$$

This is an analytic map in the upper half plane and it is, for any $z \in \mathbb{C}^+$, determined by the operator-valued quadratic equation

$$z\mathcal{G}(z) = \mathbf{id} + \eta\big(\mathcal{G}(z)\big) \cdot \mathcal{G}(z). \tag{30}$$

For a derivation and details on this, see [23, 22]. In [18] it is shown that Eq. (30) has for fixed $z \in \mathbb{C}^+$ exactly one solution $\mathcal{G}$ with negative imaginary part; furthermore, this solution is the limit of iterates $\mathcal{G}_n = \mathcal{F}_z^n(\mathcal{G}_0)$ for any initial point $\mathcal{G}_0$ with negative imaginary part. $\mathcal{F}_z$ is here the mapping

$$\mathcal{F}_z(\mathcal{G}) = \big(z \cdot \mathbf{id} - \eta(\mathcal{G})\big)^{-1}. \tag{31}$$

6. Appendix II: Proof of the main theorems

6.1. **Proof of Theorem 3.** There are several alternative methods of proof of Theorem 3. It can be derived from Girko [24] by specializing his Theorem to a block matrix with N^2 blocks of size $d \times d$ obtained from our matrix X_N by a suitable similarity transformation. It can be derived by elementary method of moments, see [20]. We choose here to give a proof by using various results from the theory of operator-valued free probability, thus showing how this result fits conceptually into the frame of operator-valued free probability. The connection between random matrices and operator-valued free probability (for "band" random matrices, with independent entries but variances depending on the position of the entry) was made by Shlyakhtenko in [25, 26].

First, one has to observe that the blocks of X_N converge almost surely to a semi-circular family (see [5, 6, 27, 28]), thus the wanted limit distribution of X_N is the same as the one of a $d\times d$-matrix S, where the entries of S are from a semi-circular family, with covariance σ. By using the description of operator-valued cumulants of this matrix in terms of the cumulants of the entries of the matrix (see [29]), it is readily seen that S is a $M_d(\mathbb{C})$-valued semi-circular element, with covariance η. The equation for $\mathcal{G}(z)$ follows then from the basic R-transform or cumulant theory of operator-valued free probability theory, see Sect. 5.4 above, in particular Eq. (30).

6.2. **Proof of Theorem 2.** If we decompose the positive definite matrices $\Psi^{(s)}$ and $\hat{\Psi}^{(s)}$ as $\hat{\Psi}^{(s)} = A_s^2$ and $\Psi^{(s)} = B_s^2$ (where we take the positive square roots $A_s = \sqrt{\hat{\Psi}^{(s)}}$ and $B_s = \sqrt{\Psi^{(s)}}$) then our $n \times n$ matrix H_n can be written as

$$H_n = \sum_{s=1}^{t} A_s Z_s B_s$$

where $Z_1, ..., Z_t$ are independent $n\times n$ matrices of independent complex Gaussian variables.

By our assumption on the convergence of $(\Psi^{(s)}, \hat{\Psi}^{(s)})_{s=1,\dots,t}$ we know that also $(A_s, B_s)_{s=1,\dots,t}$ converges in distribution to $(a_s, b_s)_{s=1,\dots,t}$, where $a_s = \sqrt{\Psi_s}$ and $b_s = \sqrt{\hat{\Psi}_s}$. By the asymptotic freeness of Gaussian random matrices from non-random matrices [5, 7, 6, 28] we know then that $(A_s, B_s, Z_s)_{s=1,\dots,t}$ converges in $*$-distribution to $(a_s, b_s, c_s)_{s=1,\dots,t}$ in some $(\mathcal{A}, \varphi)$, where $c_1, \dots, c_t$ are $*$-free circular elements such that $a_1, b_1, \dots, a_t, b_t$ is $*$-free from $c_1, \dots, c_t$. (A circular element is of the form $c = s_1 + is_2$ where s_1, s_2 are free semicircular elements.) By $\mathcal{B}$ we denote, as in our theorem, the subalgebra of $\mathcal{A}$ which is generated by all $a_1, b_1, \dots, a_t, b_t$.

Then H_n converges in $*$-distribution to

$$H = \sum_{s=1}^{t} a_s c_s b_s,$$

and $H_nH_n^*$ converges to HH^*. To calculate the distribution of HH^* we go again over to the selfadjoint 2×2 matrix

$$X = \begin{bmatrix} 0 & H \\ H^* & 0 \end{bmatrix} = \sum_{s=1}^{t} \begin{bmatrix} 0 & a_s c_s b_s \\ b_s c_s^* a_s & 0 \end{bmatrix} = \sum_{s=1}^{t} \begin{bmatrix} a_s & 0 \\ 0 & b_s \end{bmatrix} \begin{bmatrix} 0 & c_s \\ c_s^* & 0 \end{bmatrix} \begin{bmatrix} a_s & 0 \\ 0 & b_s \end{bmatrix}.$$

The relation between the distribution of X and the distribution of HH^* is as in Section 3.2, thus it remains essentially to determine the distribution of X.

Put

$$A_s := \begin{bmatrix} a_s & 0 \\ 0 & b_s \end{bmatrix}$$

and

$$S_s := \begin{bmatrix} 0 & c_s \\ c_s^* & 0 \end{bmatrix},$$

so that we have

$$X = \sum_{s=1}^{t} A_s S_s A_s$$

The main problem is now that the different terms $A_sS_sA_s$ in X are not free and thus one cannot reduce the situation directly to the case $t = 1$. However, we have operator-valued freeness with respect to a suitably chosen conditional expectation. Namely, let us first take the conditional expectation E from $\mathcal{A}$ to $\mathcal{B}$ (which exists by general arguments, because we are in a tracial situation, see, e.g., [23]) and then we go over to 2×2 matrices by taking this E entrywise, i.e. we consider

$$1 \otimes E : M_2(\mathcal{A}) \to M_2(\mathcal{B})$$

given by

$$1 \otimes E \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = \begin{bmatrix} E(a_1) & E(a_2) \\ E(a_3) & E(a_4) \end{bmatrix}.$$

From Theorem 3.5 in [29] it follows now that, for each $s = 1, \dots, t$, $A_sS_sA_s$ is a semicircular element over $M_2(\mathcal{B})$, and furthermore, that all $A_1S_1A_1, \dots, A_tS_tA_t$ are free over $M_2(\mathcal{B})$. But this implies that also their sum X is a semicircular element over $M_2(\mathcal{B})$. It remains to calculate its covariance function. We have

$$\eta(D) = 1 \otimes E(XDX),$$

i.e., for $d_1, d_2, d_3, d_4 \in \mathcal{B}$

$$\eta \begin{bmatrix} d_1 & d_3 \\ d_4 & d_2 \end{bmatrix} = \sum_{i,j=1}^{t} \begin{bmatrix} E[a_i c_i b_i d_2 b_j c_j^* a_j] & E[a_i c_i b_i d_4 a_j c_j b_j] \\ E[b_i c_i^* a_i d_3 b_j c_j^* a_j] & E[b_i c_i^* a_i d_1 a_j c_j b_j] \end{bmatrix}$$

It is quite easy to see that the conditional expectation $E : \mathcal{A} \to \mathcal{B}$ acts for all $b \in \mathcal{B}$ as

$$\begin{aligned} E[c_i b c_j^*] &= \delta_{ij}\varphi(b) \\ E[c_i^* b c_j] &= \delta_{ij}\varphi(b) \\ E[c_i^* b c_j^*] &= 0 \\ E[c_i b c_j] &= 0 \end{aligned}$$

Thus

$$\eta \begin{bmatrix} d_1 & d_3 \\ d_4 & d_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^t a_i a_i \varphi(b_i d_2 b_i) & 0 \\ 0 & \sum_{i=1}^t b_i b_i \varphi(a_i d_1 a_i) \end{bmatrix} = \begin{bmatrix} \eta_2(d_2) & 0 \\ 0 & \eta_1(d_1) \end{bmatrix}$$

where $\mathcal{B}_1$ is the algebra generated by $a_1^2, \dots, a_t^2$ (i.e., the algebra generated by $\Psi_1, \dots, \Psi_t$) and $\mathcal{B}_2$ is the algebra generated by $b_1^2, \dots, b_t^2$ (i.e, the algebra generated by $\hat{\Psi}_1, \dots, \hat{\Psi}_t$) and $\eta_1 : \mathcal{B}_1 \to \mathcal{B}_2$ and $\eta_2 : \mathcal{B}_2 \to \mathcal{B}_1$ are the mappings as in our Theorem.

Denote by $\mathcal{D}$ the subalgebra of $M_2(\mathcal{B})$ of the form

$$\mathcal{D} = \begin{bmatrix} \mathcal{B}_1 & 0 \\ 0 & \mathcal{B}_2 \end{bmatrix}.$$

We see that η maps $\mathcal{D}$ to itself. Then it follows by Theorem 3.1 of [29] that X is also a semicircular element over $\mathcal{D}$, with the same η. This implies then that our corresponding operator-valued Cauchy transform $\mathcal{G}(z)$ lies in $\mathcal{D}$, thus is of the form

$$\mathcal{G}(z) = \begin{bmatrix} \mathcal{G}_1(z) & 0 \\ 0 & \mathcal{G}_2(z) \end{bmatrix},$$

where $\mathcal{G}_1(z) \in \mathcal{B}_1$ and $\mathcal{G}_2(z) \in \mathcal{B}_2$. The rest follows then from analyzing the corresponding operator-valued quadratic equation (30) for $\mathcal{G}(z)$ and relating it with the Cauchy transform of HH^* as in Section 3.2.

7. Appendix III: Selfadjoint case with rectangular blocks

In some applications one encounters situations where the blocks themselves might not be square matrices, but more general rectangular matrices. Of course, the sizes of the blocks must fit together to make up a big square matrix. This means that in Theorem 1 we replace $n = dN$ by a decomposition $n = N_1 + \cdots + N_d$, and the block $A^{(i,j)}$ will then be a $N_i \times N_j$-matrix. We are interested in the limit that N_i/n converges to some number α_i.

Let us first introduce the generalizations of our relevant notations from the square case. Note that dependent rectangular blocks can be re-cut into different nonequivalent configurations of dependent blocks. We will assume that such repartitioning has already been done and resulted in the covariance function $\sigma(i,j;k,l)$ that can only be different from zero if the size of the block $A^{(i,j)}$ fits (at least in the limit $n \to \infty$) with the size of the block $A^{(k,l)}$.

Notation 4. Fix a natural number d and a d-tuple $\alpha = (\alpha_1, \dots, \alpha_d)$ with $0 < \alpha_i < 1$ for all $i = 1, \dots, d$ and $\alpha_1 + \dots + \alpha_d = 1$. Furthermore, let a covariance function $\sigma = \big(\sigma(i,j;k,l)\big)_{i,j,k,l=1}^{d}$ be given with the property that (14) holds and in addition $\sigma(i,j;k,l) = 0$ unless $\alpha_i = \alpha_l$ and $\alpha_j = \alpha_k$. Then we use the following notations.

1) $M_\alpha(\mathbb{C})$ are those matrices from $M_d(\mathbb{C})$ which correspond to square blocks,

$$M_\alpha(\mathbb{C}) := \{D \in M_d(\mathbb{C}) \mid [D]_{ij} = 0 \text{ unless } \alpha_i = \alpha_j\}.$$

2) We define the *weighted covariance mapping*

$$\eta_\alpha : M_\alpha(\mathbb{C}) \to M_\alpha(\mathbb{C})$$

as follows:

$$[\eta_\alpha(D)]_{ij} := \sum_{k,l=1}^{d} \sigma(i,k;l,j) \cdot \alpha_k \cdot [D]_{kl}.$$

3) Furthermore, the weighted trace

$$\operatorname{tr}_\alpha : M_\alpha(\mathbb{C}) \to \mathbb{C}$$

is given by

$$\operatorname{tr}_\alpha(D) := \sum_{i=1}^{d} \alpha_i \cdot [D]_{ii}.$$

The following statement for rectangular blocks can be reduced to the case of square blocks by cutting the rectangular blocks into smaller square blocks (at least asymptotically); for a more direct combinatorial proof, see [20].

Theorem 5. *With the above notation, for $\{N_1, \dots, N_d\} \subset \mathbb{N}$ consider block matrices*

$$X_{N_1,\dots,N_d} = \big(A^{(i,j)}\big)_{i,j=1}^{d}.$$

For each $i,j = 1, \dots, d$, the $A^{(i,j)}$ are Gaussian $N_i \times N_j$ random matrices, $A^{(i,j)} = \big(a_{rp}^{(i,j)}\big)_{\substack{r=1,\dots,N_i \\ p=1,\dots,N_j}}$. The latter are such that the collection of all entries $\{a_{rp}^{(i,j)} \mid i,j = 1, \dots, d,\ r = 1, \dots, N_i, p = 1, \dots, N_j\}$ of the matrix $X_{N_1,\dots,N_d}$ forms a Gaussian family which is determined by

$$a_{rp}^{(i,j)} = \overline{a_{pr}^{(j,i)}} \qquad \text{for all } i,j = 1, \dots, d,\ r = 1, \dots, N_i,\ p = 1, \dots, N_j$$

and the prescription of mean zero and covariance

$$E[a_{rp}^{(i,j)}a_{qs}^{(k,l)}] = \frac{1}{n}\delta_{rs}\delta_{pq}\cdot\sigma(i,j;k,l),$$

where we put

$$n := N_1 + \cdots N_d.$$

Then, for $n\to\infty$ *such that*

$$\lim_{n\to\infty}\frac{N_i}{n} = \alpha_i \qquad \text{for all } i = 1,\dots,d,$$

the matrix $X_{N_1,\dots,N_d}$ *has almost surely a limiting eigenvalue distribution whose Cauchy transform* $G(z)$ *is determined by*

$$G(z) = \mathrm{tr}_\alpha(\mathcal{G}(z)),$$

where $\mathcal{G}(z)$ *is an* $M_\alpha(\mathbb{C})$*-valued analytic function on the upper complex half plane, which is uniquely determined by the facts that* (17) *holds and that it satisfies for all* z *in the upper complex half plane the matrix equation*

$$z\mathcal{G}(z) = I_d + \eta_\alpha(\mathcal{G}(z))\cdot\mathcal{G}(z). \tag{32}$$

Acknowledgement

We thank Ralf Müller for bringing the problem of block matrices in the context of MIMO to our attention. We also acknowledge the very constructive feedback and suggestions of two reviewers of an earlier version of this paper.

References

[1] G. J. Foschini and M. Gans, "On Limits of wireless communications in fading environment when using multi-element antennas," *Wireless Personal Communications*, Vol. 6, No. 3, pp. 331-335, 1998.

[2] S. Verdú, "Capacity region of Gaussian CDMA channels: The symbol synchronous case," *Proc. Allerton Conf. on Communication, control and computing*, Monticello, IL, pp.1025-1034, 1986.

[3] U. Madhow and M. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Communications*, Vol. 42, pp.3178-3188, Dec. 1994.

[4] S. Verdú, *Multiuser Detection*, Cambridge, UK: Cambridge University Press, 1998.

[5] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free random variables*, ser. CRM Monograph Series. Providence, RI: American Mathematical Society, 1992, vol. 1.

[6] F. Hiai and D. Petz, *The semicircle law, free random variables and entropy*, ser. Mathematical Surveys and Monographs. Providence, RI: American Mathematical Society, 2000, vol. 77.

[7] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, vol. 335 of London Mathematical Society Lecture Notes Series, Cambridge University Press, 2006.
[8] A. M. Tulino and S. Verdú, "Random matrices and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, 2004.
[9] N. R. Rao and A. Edelman, "The polynomial method for random matrices," preprint, 2006, math.PR/0601389.
[10] D.N.C. Tse and O. Zeitouni, "Linear multiuser receivers in random environments," *IEEE Trans. Information Theory*, Vol. 46, No.1, pp. 171-188, Jan. 2000.
[11] 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. Information Theory*, Vol. 49, No. 10, pp.2545-2561, Oct. 2003.
[12] R.R. Müller, "On the asymptotic eigenvalue distribution of concatenated vector-valued fading channels,"*IEEE Trans. Information Theory*, Vol.48, No.7, pp.2086-2091, July 2002.
[13] R.R. Müller, "A random matrix model of communication via antenna arrays,"*IEEE Trans. Information Theory*, Vol.48, No.9, pp.2495-2506, September 2002.
[14] M. Debbah and R.R. Müller, "MIMO channel modeling and the principle of maximum entropy," *IEEE Trans. Information Theory*, Vol.51, No.5, pp.1667-1690, May 2005.
[15] E. G. Larsson and P. Stoica, *Space-time block coding for wireless communications*, Cambridge 2003.
[16] C.-N. Chuah, D. Tse, J. Kahn, R. Valenzuela, "Capacity Scaling in MIMO Wireless Systems Under Correlated Fading," *IEEE Trans. Information Theory*, Vol. 48, No. 3, March 2002.
[17] F. Benaych-Georges, "Rectangular random matrices, related free entropy and free Fisher's information," 2005, math.OA/0512081.
[18] W. Helton, R. Rashidi Far, R. Speicher "Operator-valued semicircular elements: Solving a quadratic matrix equation with positivity constraints" preprint 2007, math.0A/0703510.
[19] T. Oraby, "The spectral laws of symmetric block-matrices with large random blocks," 2005, submitted.
[20] R. Rashidi Far, T. Oraby, W.Bryc, R. Speicher: "Spectra of large block matrices," preprint 2006, cs.IT/0610045.
[21] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C*, 2nd ed. Cambridge: Cambridge University Press, 1992, the art of scientific computing.
[22] R. Speicher, "Combinatorial theory of the free product with amalgamation and operator-valued free probability theory," *Mem. Amer. Math. Soc.*, vol. 132, no. 627, pp. x+88, 1998.
[23] D. Voiculescu, "Operations on certain non-commutative operator-valued random variables," *Astérisque*, no. 232, pp. 243–275, 1995, recent advances in operator algebras (Orléans, 1992).

[24] V. L. Girko, "A matrix equation for resolvents of random matrices with independent blocks," *Theory Probab. Appl.*, Vol. 40 , No. 4, pp.635-644, 1995.

[25] D. Shlyakhtenko, "Random Gaussian band matrices and freeness with amalgamation," *Internat. Math. Res. Notices*, no. 20, pp. 1013–1025, 1996.

[26] D. Shlyakhtenko, "Gaussian random band matrices and operator-valued free probability theory," in *Quantum probability (Gdańsk, 1997)*, ser. Banach Center Publ. Warsaw: Polish Acad. Sci., 1998, vol. 43, pp. 359–368.

[27] W. Bryc, A. Dembo, and T. Jiang, "Spectral measure of large random Hankel, Markov and Toeplitz matrices," *Ann. Probab.*, vol. 34, pp. 1–38, 2005, expanded version: arxiv.org/abs/math.PR/0307330.

[28] F. Hiai and D. Petz, "Asymptotic freeness almost everywhere for random matrices," *Acta Sci. Math. (Szeged)*, vol. 66, no. 3-4, pp. 809–834, 2000.

[29] A. Nica, D. Shlyakhtenko, and R. Speicher, "Operator-valued distributions. I. Characterizations of freeness," *Int. Math. Res. Not.*, no. 29, pp. 1509–1538, 2002.

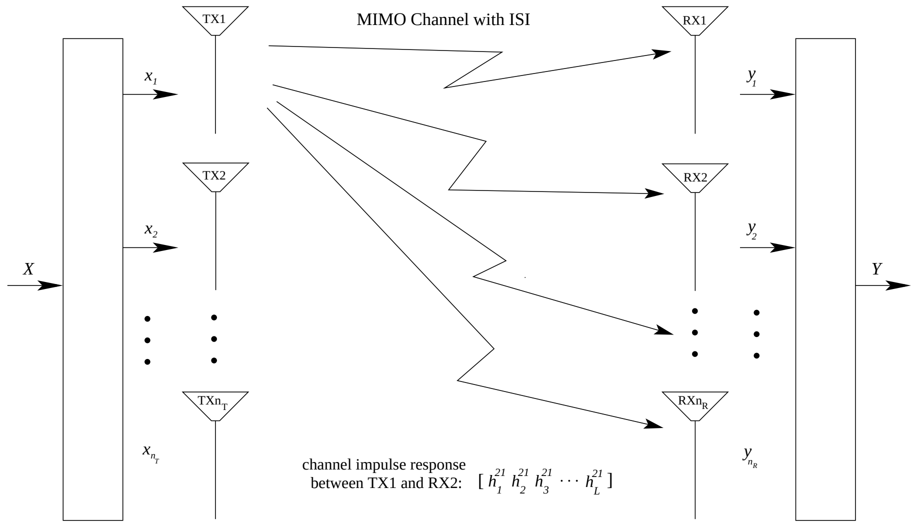


FIGURE 1. Block diagram of a MIMO system with ISI.

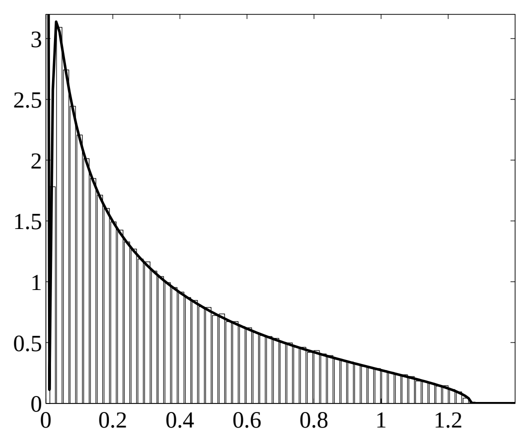


FIGURE 2. Superimposed theoretical density of the eigenvalues of complex normal $H_n H_n^*/n$ for a channel with ISI $L = 4$ and a MIMO system $n_R = n_T$ with frame length of $K = 4$ over its histogram for $N = 100$, based on 100 realizations.

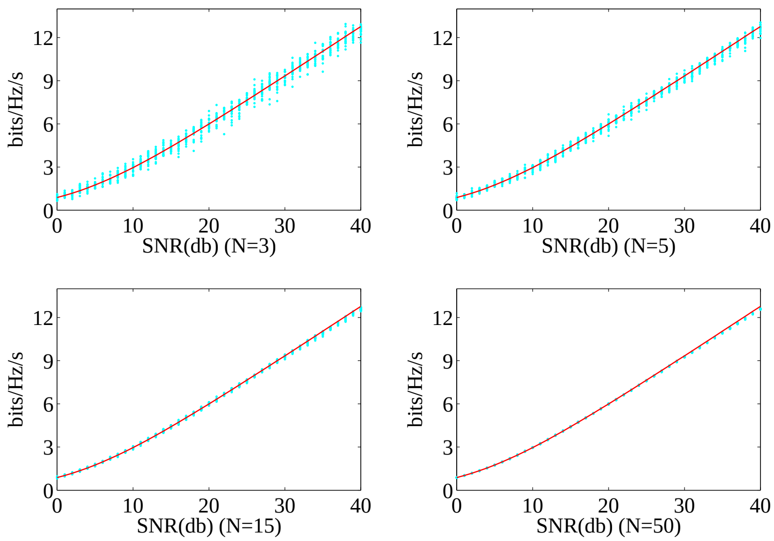


FIGURE 3. Asymptotic capacity (solid line) of the channel with ISI, L=2, in a MIMO system with frame length K=2 compared with the capacity of the same channel for different block sizes (dots).

List of Figures