On Information Rates of the Fading Wyner Cellular Model via the Thouless Formula for the Strip

Nathan Levy*†, Ofer Zeitouni‡§ and Shlomo Shamai (Shitz)†

* Département de Mathématiques et Applications, Ecole Normale Supérieure, Paris 75005, France

† Department of Electrical Engineering, Technion, Haifa 32000, Israel

‡ Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel

§ School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA

Abstract

We apply the theory of random Schrödinger operators to the analysis of multi-users communication channels similar to the Wyner model, that are characterized by short-range intra-cell broadcasting. With H the channel transfer matrix, $HH^\dagger$ is a narrow-band matrix and in many aspects is similar to a random Schrödinger operator. We relate the per-cell sum-rate capacity of the channel to the integrated density of states of a random Schrödinger operator; the latter is related to the top Lyapunov exponent of a random sequence of matrices via a version of the Thouless formula. Unlike related results in classical random matrix theory, limiting results do depend on the underlying fading distributions. We also derive several bounds on the limiting per-cell sum-rate capacity, some based on the theory of random Schrödinger operators, and some derived from information theoretical considerations. Finally, we get explicit results in the high-SNR regime for some particular cases.

I. Introduction

The growing demand for ubiquitous access to high-data rate services, has produced a huge amount of research analyzing the performance of wireless communications systems. Techniques for providing better service and coverage in cellular mobile communications are currently being investigated by industry and academia. In particular, the use of joint multi-cell processing (MCP), which allows the base-stations (BSts) to jointly process their signals, equivalently creating a distributed antenna array, has been identified as a key tool for enhancing system performance (see [1][2] and references therein for surveys of recent results on multi-cell processing).

Motivated by the fact that mobile users in a cellular system "see" only a small number of BSts, and by the desire to provide analytical results, an attractive analytically tractable model for a multi-cell system was suggested by Wyner in [3] (see also [4] for an earlier relevant work). In this model, the system's cells are ordered in either an infinite linear array, or in the familiar two-dimensional hexagonal pattern (also infinite). It is assumed that only adjacent-cell interference is present and characterized by a single parameter, a scaling factor $\alpha \in [0,1]$. Considering non-fading channels and a "wideband" (WB) transmission scheme, where all bandwidth is available for coding (as opposed to random spreading), the throughputs obtained with optimum and linear MMSE joint processing of the received signals from all cell-sites are derived in [3]. Since it was first presented, "Wyner-like" models have provided a framework for many works analyzing various transmission schemes in both the up-link and down-link channels (see [1][5] and references therein).

In this paper we consider a generalized "Wyner-like" cellular setup and study its per-cell sum-rate capacity. According to Wyner's setup, the cells are arranged on a circle (or a line), and the mobile users "see" only a fixed number of BSts which are located close to their cell's boundaries. All the BSts are assumed to be connected through an ideal back-haul network to a central multi-cell processor (MCP), that can *jointly* process the up-link received signals of all cell-sites, as well as pre-process the signals to be transmitted by all cell-sites in the down-link channel. The model is characterized by short-range intra-cell broadcasting. Thus, if we denote by H the channel transfer matrix, then $HH^{\dagger}$ is in many aspects similar to a random Schrödinger operator. More specifically, the per-cell sum-rate capacity of the channel

is a function of the integrated density of state of $HH^\dagger$, which in turn is related to the top Lyapunov exponent of a random sequence of matrices via a version of the Thouless formula. Unlike associated results in classical random matrix theory, limiting results do depend on the underlying fading distributions.

As an application of our result and motivated by the fact that future cellular systems implicitly assume high-SNR configurations mandatory for high data rate services, we get explicit results in the high-SNR regime for some particular cases.

The rest of the paper is organized as follows. In Section II, we present the problem statement. In Section III, we prove the convergence of the per-cell sum-rate capacity when the number of cells and BSts goes to infinity and we express the limit in terms of the Lyapunov exponent of a sequence of random matrices (Theorem III.1). In Section IV, we give several reformulations of this result that yields a particularly simple expression in the high-SNR regime. In Section V, we give different bounds on the per-cell sum-rate capacity, some of which are based on the theory of product of random matrices, and some on information theoretical considerations. In Section VI, we specialize the results and make them explicit in some particular cases. Finally in Section VII we discuss some open problems using numerical simulations. The relevant background on the theory of Lyapunov exponents is given in Appendix A1, and the relevant background on exterior products is given in Appendix C. Several proofs are postponed to Appendices A2, A3 and B. The per-cell sum-rate capacity of the non-fading channels is derived in Appendix D.

II. Problem statement

In this paper we consider the following setup. $m+d$ cells with K single antenna users per cell are arranged on a line, where the m single antenna BSts are located in the cells. Starting with the WB transmission scheme where all bandwidth is devoted for coding and all K users are transmitting simultaneously each with average power ρ, and assuming synchronized communication, a vector baseband representation of the signals received at the system's BSts is given for an arbitrary time index i by

$$\boldsymbol{y}(i) = H_m(i)\boldsymbol{x}(i) + \boldsymbol{z}(i), \tag{II.1}$$

where $\boldsymbol{x}(i)$ is the $(m+d)K$ complex Gaussian symbols vector, $\boldsymbol{z}(i)$ is the unitary complex Gaussian additive noise vector. Note that the SNR is ρ. From now on, we omit the time

index i. H_m is the following $m \times K(m+d)$ channel transfer matrix, which is a $d+1$ block diagonal matrix defined by

$$H_m = \begin{pmatrix} \zeta_{1,1} & \zeta_{1,2} & \cdots & \zeta_{1,d+1} & 0 & \cdots & 0 \\ 0 & \zeta_{2,2} & \cdots & \zeta_{2,d+1} & \zeta_{2,d+2} & & \vdots \\ \vdots & & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & \zeta_{m,m} & \zeta_{m,m+1} & \cdots & \zeta_{m,d+m} \end{pmatrix},$$

where $\zeta_{i,j}$ are $1 \times K$ row vectors. For $s \in \mathbb{N}^*$, we will denote by ζ^s the vector $(\zeta_{s-d,s}, \ldots, \zeta_{s,s})$ and we denote by π it distribution. We assume in the rest of the paper that for $n \in \mathbb{N}^*$ and $0 \leq i \leq d$ the vectors $(\zeta_{n-i,n})$ are distributed according to π_i. We define $\Omega = (\zeta^n)_{n \in \mathbb{N}^*}$ and $\mathbb{P}$, the probability distribution on Ω associated to the above problem. We denote by $\mathbb{E}$ the associated expectation. We also use the 2 norm for vectors and matrices. For matrices, it is the Froebenius norm, which is a sub-multiplicative norm.

Throughout this paper, we assume a subset of the following hypotheses.

(H1) The vectors $(\zeta^j)_{j \in \mathbb{N}^*}$ form a stationary ergodic sequence.

(H2) There exists $\varepsilon > 0$ such that for $0 \leq i \leq d$, $\mathbb{E}_{\pi_i} \left|\log |x|\right|^{1+\varepsilon} < \infty$.

(H3) If $(x_0, \ldots, x_d)$ is distributed according to π, then almost surely, $x_0 x_d^\dagger \neq 0$.

For $m \in \mathbb{N}^*$ and $\lambda > 0$, we set $G_m = H_m H_m^\dagger + \lambda \operatorname{Id}_m$, where Id_m is the $m \times m$ identity matrix. Although G_m depends on λ, we will not write that dependence unless there is an ambiguity. Under the assumption that $H_m(i)$ is ergodic with respect to the time index i, that the Channel State Information (CSI) is known at the receiver whereas the users know the statistics of the CSI, and that the channel varies fast enough so as to allow each transmitted codeword to experience a large number of fading states, we follow [1] and study the per-cell sum-rate capacity that is given by the following formula ([6])

$$Cap_m(\rho) = \frac{1}{m} \mathbb{E} \log \det \left(\operatorname{Id} + \rho H_m H_m^\dagger \right) = \log \rho + \frac{1}{m} \mathbb{E} \left(\log \det G_m(\lambda) \right), \qquad \text{(II.2)}$$

where $\lambda = 1/\rho$.

III. Main result

We set for $i \in \mathbb{N}^*$

$$C_i = \begin{pmatrix} \zeta_{d(i-1)+1,d(i-1)+1} & \zeta_{d(i-1)+1,d(i-1)+2} & \cdots & \zeta_{d(i-1)+1,di} \\ 0 & \zeta_{d(i-1)+2,d(i-1)+2} & \cdots & \zeta_{d(i-1)+2,di} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \zeta_{di,di} \end{pmatrix}$$

and

$$D_i = \begin{pmatrix} {\zeta_{d(i-2)+1,d(i-1)+1}}^\dagger & {\zeta_{d(i-2)+2,d(i-1)+1}}^\dagger & \cdots & {\zeta_{d(i-1),d(i-1)+1}}^\dagger \\ 0 & {\zeta_{d(i-2)+2,d(i-1)+2}}^\dagger & \cdots & {\zeta_{d(i-1),d(i-1)+2}}^\dagger \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & {\zeta_{d(i-1),di}}^\dagger \end{pmatrix}.$$

For all $i \in \mathbb{N}^*$, C_i are $d \times dK$ matrices and D_i are $dK \times d$ matrices. We fix $\zeta_{i,j}$ with $i \leq 0$ or $j \leq 0$ so that $C_1 D_1 = \mathrm{Id}_{d,d}$.

We thereby get the following block description of H

$$H_{dn} = \left(\begin{array}{c|c|c|c|c} C_1 & D_2^\dagger & 0_{d,dK} & \cdots & 0_{d,dK} \\ \hline 0_{d,dK} & C_2 & D_3^\dagger & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \ddots & 0_{d,dK} \\ \hline 0_{d,dK} & \cdots & 0_{d,dK} & C_n & D_{n+1}^\dagger \end{array}\right),$$

where $0_{d,dK}$ is the $d \times dK$ zero matrix.

Under the hypothesis (H2), in order to study the limit in m of $Cap_m(\rho)$, it is enough to study $Cap_{nd}(\rho)$ (see Remark A.13 following the proof of Lemma A.6). We get the following block representation of G_{dn}:

$$G_{dn} = \left(\begin{array}{c|c|c|c} C_1 C_1^\dagger + D_2^\dagger D_2 + \lambda \,\mathrm{Id}_d & (C_2 D_2)^\dagger & 0_d & 0_d \\ \hline C_2 D_2 & \ddots & \ddots & 0_d \\ \hline 0_d & \ddots & \ddots & (C_n D_n)^\dagger \\ \hline 0_d & 0_d & C_n D_n & C_n C_n^\dagger + D_{n+1}^\dagger D_{n+1} + \lambda \,\mathrm{Id}_d \end{array}\right).$$

Note that under (H3), for all $i \in \mathbb{N}^*$, $C_i D_i$ is a $d \times d$ invertible matrix.

For $i \in \mathbb{N}^*$, we denote by M_i the following matrix

$$\left(\begin{array}{c|c} 0_d & \mathrm{Id}_d \\ \hline -(C_{i+1}D_{i+1})^{-1\dagger} C_i D_i & -(C_{i+1}D_{i+1})^{-1\dagger} \left(C_i C_i^\dagger + D_{i+1}^\dagger D_{i+1} + \lambda \,\mathrm{Id}_d \right) \end{array}\right)$$

and denote $N_i = \bigwedge^d M_i$. Moreover, $\gamma(N)$ denotes the top Lyapunov exponent associated with $\{N_i\}$, i.e.

$$\gamma(N) \triangleq \lim_{n\to\infty} \frac{1}{n} \log \|N_n \cdots N_1\| .$$

Note that by Theorem A.1, $\gamma(N)$ is deterministic. See Appendix A1 for the definitions concerning the Lyapunov exponents and Appendix C for the relevant background on exterior products. Recall that $(M_i)_{i\in\mathbb{N}^*}$ and $(N_i)_{i\in\mathbb{N}^*}$ depend on λ.

Theorem III.1. *Assume (H1), (H2) and (H3), and set $\lambda = 1/\rho$.*

1. *We have*

$$Cap_m(\rho) \xrightarrow[m\to\infty]{} \log\rho + \mathbb{E}_\pi \log\left|\zeta_0\zeta_d^\dagger\right| + \frac{1}{d}\gamma\left(N\right) \triangleq Cap(\rho),$$

   where the expectation is taken such that $(\zeta_0, \ldots, \zeta_d)$ is distributed according to π.

2. *As ρ goes to infinity,*

$$Cap(\rho) = \log\rho + \mathbb{E}_\pi \log\left|\zeta_0\zeta_d^\dagger\right| + \frac{1}{d}\gamma\left(N(\lambda=0)\right) + o(1).$$

The theorem is proved in Appendix A.

As an alternative to deriving exact analytical results we will also be interested in extracting parameters that characterize the channel rate in the high-SNR regime [7]; such parameters are the high-SNR slope (also referred to as the "multiplexing gain")

$$\mathcal{S}_\infty \triangleq \lim_{\rho\to\infty} \frac{Cap(\rho)}{\log(K\rho)},$$

and the high-SNR power offset

$$\mathcal{L}_\infty \triangleq \lim_{\rho\to\infty} \frac{1}{\log 2}\left(\log(K\rho) - \frac{Cap(\rho)}{\mathcal{S}_\infty}\right),$$

yielding the following affine capacity approximation

$$Cap(\rho) \approx \frac{\mathcal{S}_\infty \log 2}{3|_{\mathrm{dB}}}\left(K\rho|_{\mathrm{dB}} - 3|_{\mathrm{dB}}\mathcal{L}_\infty\right).$$

A direct consequence of Theorem III.1 is the following high-SNR characterization.

Corollary III.2. *Assume (H1), (H2) and (H3). Then $\mathcal{S}_\infty = 1$ and*

$$\mathcal{L}_\infty = \frac{1}{\log 2}\left[\log K - \mathbb{E}_\pi \log\left|\zeta_0\zeta_d^\dagger\right| - \frac{1}{d}\gamma(N(\lambda=0))\right].$$

IV. Reformulations

We now derive alternative formulations for $\gamma(N)$ in Subsection IV-A and for $\gamma(N(\lambda=0))$ (which characterizes the hign-SNR regime), in Subsection IV-B.

A. Non-asymptotic results

In order to study $\gamma(N)$, we express it as the Lyapunov exponent of simpler matrices. For $i \geq d+1$, we define the following random matrices.

$$\mathfrak{m}_i = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \\ -\frac{\widetilde{\zeta}_{i,i-d}}{\widetilde{\zeta}_{i,i+d}} & \cdots & \cdots & -\frac{\widetilde{\zeta}_{i,i+d-1}}{\widetilde{\zeta}_{i,i+d}} \end{pmatrix},$$

where $\widetilde{\zeta}_{i,l}$ is the coefficient in position (i,l) in G_{dn}, and set

$$\mathfrak{n}_i = \bigwedge^{d} \mathfrak{m}_i.$$

Note that $(\mathfrak{m}_i)_{i\geq d+1}$ and $(\mathfrak{n}_i)_{i\geq d+1}$ depend on λ. We get the following proposition, whose proof is given in Appendix B1.

Proposition IV.1. *Assume (H1), (H2) and (H3). Then, $N_i = \mathfrak{n}_{id}\cdots\mathfrak{n}_{(i-1)d+1}$. Therefore, for every $\lambda \geq 0$, $\gamma(N) = d\gamma(\mathfrak{n})$, hence,*

$$Cap(\rho) = \log\rho + \mathbb{E}_{\pi_0,\pi_d} \log\left|\zeta_0\zeta_d^{\dagger}\right| + \gamma\left(\mathfrak{n}\right).$$

Note that for a given $i \in \mathbb{N}$, N_i depends on $\zeta^{d(i-1)+1},\ldots,\zeta^{(d+1)i}$, that is, the fading coefficients of $2d$ different cells. We now want to reduce the product of the N_i to a product of random matrices (that we denote by Ξ_i) depending on the fading coefficients of only d cells. Then we reduce it further to a product of random matrices (that we denote by ξ_i) depending on the fading coefficients of only one cell.

By doing so, we achieve two goals: first, we express $\gamma(N)$ as the Lyapunov exponent of simpler matrices. Second, if the fading coefficients are i.i.d for different cells, then the Ξ_i and the ξ_i are i.i.d. Products of i.i.d random matrices have been studied extensively (see for example [8]), moreover, their study can be reduced to the study of a Markov chain on an

appropriate space, which can lead to actual analytic expressions (see [5] for an example of study of such a Markov chain).

For $i \in \mathbb{N}^*$, we denote by Δ_i the following matrix

$$\left(\begin{array}{c|c} -\left(C_iC_i^\dagger + \lambda\,\mathrm{Id}_d\right)(C_iD_i)^{-1\dagger} & -C_iD_i + \left(C_iC_i^\dagger + \lambda\,\mathrm{Id}_d\right)(C_iD_i)^{-1\dagger} D_i^\dagger D_i \\ \hline (C_iD_i)^{-1\dagger} & -(C_iD_i)^{-1\dagger} D_i^\dagger D_i \end{array}\right)$$

and define $\Xi_i = \bigwedge^d \Delta_i$.

For $i \geq d+1$, we denote by δ_i the following matrix

$$\left(\begin{array}{c|c|c} -\frac{\zeta_{i-d+1,i}}{\zeta_{i-d,i}} & & \\ \vdots & \mathrm{Id}_{d-1} & 0_{d-1,d} \\ -\frac{\zeta_{i-1,i}}{\zeta_{i-d,i}} & & \\ \hline -\frac{\lambda+|\zeta_{i,i}|^2}{\zeta_{i-d,i}\zeta_{i,i}^\dagger} & 0_{1,d-1} & \lambda\frac{\zeta_{i-d,i}^\dagger}{\zeta_{i,i}^\dagger} \quad \cdots \quad \lambda\frac{\zeta_{i-1,i}^\dagger}{\zeta_{i,i}^\dagger} \\ \hline 0_{d-1,1} & 0_{d-1,d-1} & 0_{d-1,1} \quad \mathrm{Id}_{d-1} \\ \hline \frac{1}{\zeta_{i-d,i}\zeta_{i,i}^\dagger} & 0_{1,d-1} & -\frac{\zeta_{i-d,i}^\dagger}{\zeta_{i,i}^\dagger} \quad \cdots \quad -\frac{\zeta_{i-1,i}^\dagger}{\zeta_{i,i}^\dagger} \end{array}\right)$$

and define $\xi_i = \bigwedge^d \delta_i$. Note that $(\Delta_i)_{i\in\mathbb{N}^*}$, $(\Xi_i)_{i\in\mathbb{N}^*}$, $(\delta_i)_{i\in\mathbb{N}^*}$ and $(\xi_i)_{i\in\mathbb{N}^*}$ depend on λ.

Proposition IV.2. *Assume (H1), (H2) and (H3).*

1. *For every $\lambda \geq 0$, $\gamma(\Xi) = \gamma(N)$, hence,*

$$Cap(\rho) = \log\rho + \mathbb{E}_{\pi_0,\pi_d} \log\left|\zeta_0\zeta_d^\dagger\right| + \frac{1}{d}\gamma\left(\Xi\right).$$

2. *Assume $K = 1$. Then,*

$$\Delta_i = \delta_{id}\cdots\delta_{(i-1)d+1}.$$

Therefore, for every $\lambda \geq 0$, $\gamma(N) = d\gamma(\xi)$, hence,

$$Cap(\rho) = \log\rho + \mathbb{E}_{\pi_0,\pi_d} \log\left|\zeta_0\zeta_d^\dagger\right| + \gamma\left(\xi\right).$$

Remark IV.3. Note that for $K = 1$, for all $i \in \mathbb{N}^*$,

$$(C_iD_i)^{-1\dagger} = C_i^{-1\dagger}D_i^{-1\dagger},$$

and therefore,

$$\Delta_i = \left(\begin{array}{c|c} -C_iD_i^{-1\dagger} - \lambda(C_iD_i)^{-1\dagger} & \lambda C_i^{-1\dagger}D_i \\ \hline (C_iD_i)^{-1\dagger} & -C_i^{-1\dagger}D_i \end{array}\right).$$

Proof: Let us start by proving point 1. We define for $i \in \mathbb{N}^*$,

$$P_1(i) = \begin{pmatrix} -C_iD_i & -C_iC_i^\dagger - \lambda \operatorname{Id}_d \\ 0_d & \operatorname{Id}_d \end{pmatrix}$$

and

$$P_2(i) = \begin{pmatrix} 0_d & \operatorname{Id}_d \\ (C_iD_i)^{-1\dagger} & -(C_iD_i)^{-1\dagger} D_i^\dagger D_i \end{pmatrix},$$

so that for all $i \in \mathbb{N}^*$, $M_i = P_2(i+1)P_1(i)$. For $i \in \mathbb{N}^*$, Δ_i is defined so that $\Delta_i = P_1(i)P_2(i)$. Then, for all $n \in \mathbb{N}^*$,

$$\begin{aligned} M_n \dots M_1 &= P_2(n+1)P_1(n)P_2(n)P_1(n-1)\cdots P_2(2)P_1(1) \\ &= P_2(n+1)\Delta_n \cdots \Delta_2 P_1(1). \end{aligned}$$

and

$$\begin{aligned} \|N_n \dots N_1\| &= \left\| \bigwedge^d P_2(n+1)\Xi_n \cdots \Xi_2 \bigwedge^d P_1(1) \right\| \\ &\leq \left\| \bigwedge^d P_2(n+1) \right\| \|\Xi_n \cdots \Xi_2\| \left\| \bigwedge^d P_1(1) \right\|. \end{aligned}$$

Therefore, $\gamma(N) \leq \gamma(\Xi)$. Since $P_1(1)$ and $P_2(n+1)$ are invertible, we get the opposite inequality and point 1 is proved.

The proof of point 2 is postponed to Appendix B2. ■

B. Results in high-SNR regime

Proposition IV.4. *Assume (H1), (H2) and (H3). Assume moreover that $K = 1$.*

1. *For $i \in \mathbb{N}^*$, we set $\Psi_i^1 = C_iD_i^{-1\dagger}$ and $\Psi_i^2 = C_i^{-1\dagger}D_i$. Then*

$$\mathcal{L}_\infty = \frac{-1}{\log 2}\left[\mathbb{E}_\pi \log\left|\zeta_0\zeta_d^\dagger\right| + \frac{1}{d}\max_{0\leq i\leq d}\left(\gamma\left(\bigwedge^i \Psi^1\right) + \gamma\left(\bigwedge^{d-i} \Psi^2\right)\right)\right].$$

2. *For $i \geq d+1$, we set*

$$\psi_i^1 = \left(\begin{array}{c|c} -\frac{\zeta_{i-d+1,i}}{\zeta_{i-d,i}} & \mathrm{Id}_{d-1} \\ \vdots & \\ -\frac{\zeta_{i,i}}{\zeta_{i-d,i}} & \hline 0_{1,d-1} \end{array}\right) \text{ and } \psi_i^2 = \left(\begin{array}{c|c} -\frac{\zeta_{i-1,i}}{\zeta_{i,i}} & \mathrm{Id}_{d-1} \\ \vdots & \\ -\frac{\zeta_{i-d,i}}{\zeta_{i,i}} & \hline 0_{1,d-1} \end{array}\right)^{\dagger}.$$

Then,

$$\mathcal{L}_\infty = \frac{-1}{\log 2}\left[\mathbb{E}_\pi \log\left|\zeta_0 \zeta_d^\dagger\right| + \max_{0\leq i\leq d}\left(\gamma\left(\bigwedge^i \psi^1\right) + \gamma\left(\bigwedge^{d-i}\psi^2\right)\right)\right].$$

Remark IV.5. 1. Recall that for a stationary ergodic sequence of complex random matrices $(X_i)_{i\in\mathbb{N}^*}$ of size d,

$$\gamma\left(\bigwedge^0 X\right) = 0 \text{ and } \gamma\left(\bigwedge^d X\right) = \mathbb{E}\log|\det X_1|.$$

2. Note that if for $0 \leq i \leq d$, $\pi_i = \pi_{d-i}$ and the vectors $(\zeta_i)_{i\in\mathbb{N}^*}$ are i.i.d, then $(\psi_i^1)_{i\geq d+1}$ and $((\psi_i^2)^\dagger)_{i\geq d+1}$ have the same distribution and $(\psi_i^2)_{i\geq d+1}$ and $((\psi_i^2)^\dagger)_{i\geq d+1}$ have the same Lyapunov exponents. Therefore, as ρ goes to infinity,

$$\mathcal{L}_\infty = \frac{-1}{\log 2}\left[\mathbb{E}_\pi \log\left|\zeta_0\zeta_d^\dagger\right| + \gamma\left(\bigwedge^{\lceil d/2\rceil}\psi^1\right) + \gamma\left(\bigwedge^{\lfloor d/2\rfloor}\psi^1\right)\right].$$

Proof: Using Corollary III.2 and Proposition IV.2, in order to prove point 1, we only have to prove that

$$\gamma(\Xi(\lambda=0)) = \max_{0\leq i\leq d}\left(\gamma\left(\bigwedge^i \Psi^1\right) + \gamma\left(\bigwedge^{d-i}\Psi^2\right)\right). \tag{IV.6}$$

Recall that $\gamma(\Xi) = \gamma_1(\Delta) + \cdots + \gamma_d(\Delta)$ and that

$$\Delta_i(\lambda=0) = \left(\begin{array}{c|c} -\Psi_i^1 & 0_{d,d} \\ \hline (C_iD_i)^{-1\dagger} & -\Psi_i^2 \end{array}\right).$$

By Proposition A.5, the sequence $\gamma_1(\Delta(\lambda=0)),\ldots,\gamma_{2d}(\Delta(\lambda=0))$ is equal up to the order to the sequence

$$\gamma_1(\Psi^1),\ldots,\gamma_d(\Psi^1),\gamma_1(\Psi^2),\ldots,\gamma_d(\Psi^2).$$

Therefore, (IV.6) is a direct consequence of $\gamma(\Xi(\lambda=0)) = \gamma_1(\Delta(\lambda=0)) + \cdots + \gamma_d(\Delta(\lambda=0))$ and $\gamma\left(\bigwedge^i \Psi^{1,2}\right) = \gamma_1(\Psi^{1,2}) + \cdots + \gamma_i(\Psi^{1,2})$.

The proof of point 2 goes along the same lines using the fact that

$$\delta_i(\lambda=0)=\left(\begin{array}{c|c} -\psi_i^1 & 0_{d,d} \\ \hline \psi_i^3 & -\widetilde{\psi}_i^2 \end{array}\right),$$

where

$$\widetilde{\psi}_i^2=\begin{pmatrix} 0 & 0 & 1 \\ 0 & & 0 \\ 1 & 0 & 0 \end{pmatrix}\psi_i^2\begin{pmatrix} 0 & 0 & 1 \\ 0 & & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ and } \tilde{\psi}_i^3=\left(\begin{array}{c|c} 0_{d-1,1} & 0_{d-1,d-1} \\ \hline \frac{1}{\zeta_{i-d,i}\zeta_{i,i}^\dagger} & 0_{1,d-1} \end{array}\right),$$

therefore, the Lyapunov exponents of $\widetilde{\psi}^2$ and ψ^2 are the same. ■

V. Bounds on the capacity

A. Bounds on the top Lyapunov exponent

We use the Fröbenius norm on the matrices, it is a sub-multiplicative norm, therefore, we can apply (A.3) to the different formulations of the capacity to get the following proposition.

Proposition V.1. *Assume (H1), (H2) and (H3).*

1. *For $\lambda=1/\rho$ and $p\geq 1$,*

$$Cap(\rho)\leq \log\rho+\mathbb{E}_{\pi_0,\pi_d}\log\left|\zeta_0\zeta_d^\dagger\right|+\frac{1}{dp}\mathbb{E}\log\left\|N_p(\lambda)\cdots N_1(\lambda)\right\|.$$

2. *For $\lambda=1/\rho$ and $p\geq 1$,*

$$Cap(\rho)\leq \log\rho+\mathbb{E}_{\pi_0,\pi_d}\log\left|\zeta_0\zeta_d^\dagger\right|+\frac{1}{p}\mathbb{E}\log\left\|\mathfrak{n}_p(\lambda)\cdots\mathfrak{n}_1(\lambda)\right\|.$$

3. *For $\lambda=1/\rho$ and $p\geq 1$,*

$$Cap(\rho)\leq \log\rho+\mathbb{E}_{\pi_0,\pi_d}\log\left|\zeta_0\zeta_d^\dagger\right|+\frac{1}{dp}\mathbb{E}\log\left\|\Xi_p(\lambda)\cdots\Xi_1(\lambda)\right\|.$$

4. *Assume $K=1$. For $\lambda=1/\rho$ and $p\geq 1$,*

$$Cap(\rho)\leq \log\rho+\mathbb{E}_{\pi_0,\pi_d}\log\left|\zeta_0\zeta_d^\dagger\right|+\frac{1}{p}\mathbb{E}\log\left\|\xi_p(\lambda)\cdots\xi_1(\lambda)\right\|.$$

Moreover, the bounds are tight as p goes to infinity.

The bound of point 2 with $p=1$ can be reformulated as follows.

Corollary V.2. *For $\lambda=1/\rho$,*

$$Cap(\rho)\leq \log\rho+$$

$$\frac{1}{2}\mathbb{E}\log\left[\binom{2d-2}{d-1}\left(\sum_{l=2}^{2d}\left|\sum_{t=(l-d-1)\vee 0}^{(l-1)\wedge d}\zeta_{d+1,d+t+1}\zeta_{l,d+t+1}^{\dagger}+\lambda\mathbf{1}_{(l=d+1)}\right|^2\right)+\right.$$
$$\left.\binom{2d-1}{d}\left(\left|\zeta_{d+1,d+1}\zeta_{1,d+1}^{\dagger}\right|^2+\left|\zeta_{d+1,2d+1}\zeta_{2d+1,2d+1}^{\dagger}\right|^2\right)\right].$$

The proof is postponed to Appendix B3.

B. Other bounds

Proposition V.3. *Assume (H1) and (H2). For $\lambda = 1/\rho$,*

$$\max\left(\mathbb{E}_{\pi_0}\log\left(\lambda+|\zeta_0|^2\right),\mathbb{E}_{\pi_d}\log\left(\lambda+|\zeta_d|^2\right)\right)\leq Cap(\rho)-\log\rho\leq$$
$$\mathbb{E}\log\left(\lambda+|\zeta_{1,1}|^2+\cdots+|\zeta_{1,d+1}|^2\right).$$

Proof: The upper bound is a consequence of Hadamard's inequality for semi-positive definite Hermitian matrices. Indeed,

$$\frac{1}{m}\log\det G_m(\lambda)\leq\frac{1}{m}\sum_{i=1}^{m}\log\left(\lambda+|\zeta_{i,i}|^2+\cdots+|\zeta_{i,i+d}|^2\right).$$

Let us show the lower bound of point V.3 using the tools of [9].

$$\begin{aligned}
Cap_m(\rho)&=\frac{1}{m}I\left(\boldsymbol{x},\boldsymbol{y}|(\zeta_{i,i})_{1\leq i\leq m},\ldots,(\zeta_{i,i+d})_{1\leq i\leq m}\right)\\
&=\frac{1}{m}\sum_{j=1}^{m}I\left(x_j,\boldsymbol{y}|(x_i)_{1\leq i<j},(\zeta_{i,i})_{1\leq i\leq m},\ldots,(\zeta_{i,i+d})_{1\leq i\leq m}\right)\\
&\geq\frac{1}{m}\sum_{j=1}^{m}I\left(x_j,y_{j-d}|(x_i)_{1\leq i<j},(\zeta_{i,i})_{1\leq i\leq m},\ldots,(\zeta_{i,i+d})_{1\leq i\leq m}\right)\\
&=\frac{1}{m}\sum_{j=1}^{m}I\left(x_j,\zeta_{j-d,j}x_j+n_{j-d}|\zeta_{j-d,j}\right),
\end{aligned}$$

which is the per-cell sum-rate capacity of a single user fading channel. Therefore, the lower bound is $\mathbb{E}_{\pi_d}\log\left(1+\rho\left|\zeta_d\right|^2\right)$.

The role of the distributions π_0 and π_d can be exchanged by a right-left reflection, namely the transformation $\zeta'_{i,j}=\zeta_{m-i+1,m+d-j+1}$. Thus, we get the lower bound. ■

In the end of this section, we slightly modify the setting, by considering m cells with K single antenna users per cell and $m+d$ single antenna BSts. The communication is

characterized by the following $(m+d)\times Km$ channel transfer matrix H_m, which is a $d+1$ block diagonal matrix defined by,

$$H_m = \begin{pmatrix} \zeta_{1,1} & 0 & \cdots & 0 \\ \zeta_{2,1} & \zeta_{2,2} & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \zeta_{d+1,1} & \zeta_{d+1,2} & & \zeta_{m,m} \\ 0 & \zeta_{d+2,2} & & \zeta_{m+1,m} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \zeta_{d+m,m} \end{pmatrix},$$

where $\zeta_{i,j}$ are $1\times K$ row vectors.

We consider the per-cell sum-rate capacity $Cap_m(\rho)$ that is given by (II.2). We denote by $Cap(\rho)$ the limit of $Cap_m(\rho)$ as m tends to infinity.

Note that in the limit, this setting is equivalent to the setting we define in Section II. In particular, the normalization by $1/m$ or $1/(m+d)$ is equivalent.

Proposition V.4. *For all $n\in\mathbb{N}^*$,*

$$\frac{n}{n+d}Cap_n\left(\frac{n+d}{n}\rho\right)\leq Cap(\rho)\leq Cap_n(\rho).$$

Moreover, the bounds are tight as n goes to infinity.

Note that taking the upper bound for $n=1$, one gets the upper bound of Proposition V.3.

Proof: Number the cells from 1 to m and the antennas from 1 to $m+d$.

Upper bound. Take $1\leq n\leq m$. Define the matrix H_m^1 (resp. H_m^2) of size $(m+d)\times Kn$ (resp. $(m+d)\times K(m-n)$) to be the matrix whose columns are the Kn first columns of H_m (resp. the $K(m-n)$ last columns of H_m), such that $H_m=\left(\begin{array}{c|c} H_m^1 & H_m^2 \end{array}\right)$. Note that H_m^1 and H_m^2 are distributed like

$$\left(\frac{H_n}{0_{(m-n)\times Kn}}\right) \qquad \text{and} \qquad \left(\frac{0_{n\times K(m-n)}}{H_n}\right),$$

respectively.

Denote by $\boldsymbol{x}^1$ (resp. $\boldsymbol{x}^2$) the vector of size Kn (resp. $K(m-n)$) of symbols sent by the users of cells 1 to n (resp. the cells $n+1$ to m). Note that $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ are independent. Then (omitting the time index), (II.1) is equivalent to

$$\boldsymbol{y} = H_m^1 \boldsymbol{x}^1 + H_m^2 \boldsymbol{x}^2 + \boldsymbol{z},$$

where $\boldsymbol{y}$ and $\boldsymbol{z}$ are vectors of size $m+d$.

$$\begin{aligned} mCap_m(\rho) &= I(\boldsymbol{x}^1, \boldsymbol{x}^2; \boldsymbol{y}) \\ &= I(\boldsymbol{x}^1; \boldsymbol{y}) + I(\boldsymbol{x}^2; \boldsymbol{y}|\boldsymbol{x}^1) \\ &\leq I(\boldsymbol{x}^1; \boldsymbol{y}, \boldsymbol{x}^2) + I(\boldsymbol{x}^2; \boldsymbol{y}|\boldsymbol{x}^1) \\ &= I(\boldsymbol{x}^1; \boldsymbol{y}|\boldsymbol{x}^2) + I(\boldsymbol{x}^2; \boldsymbol{y}|\boldsymbol{x}^1), \end{aligned}$$

where in the second equality we used the chain rule for mutual information, in the inequality we used the fact that adding $\boldsymbol{x}^2$ increases the mutual information, and the last equality is due to the independence of $\boldsymbol{x}^1$, $\boldsymbol{x}^2$ (and therefore, $I(\boldsymbol{x}^1; \boldsymbol{x}^2) = 0$). Denote by $\boldsymbol{z}^1$ (resp. $\boldsymbol{z}^2$) the $n+d$ first coordinates (resp. the $m-n$ last coordinates) of $\boldsymbol{z}$. Using the fact that $\boldsymbol{x}^1$, $\boldsymbol{x}^2$, $\boldsymbol{z}^1$, and $\boldsymbol{z}^2$ are independent, one gets,

$$\begin{aligned} I(\boldsymbol{x}^1; \boldsymbol{y}|\boldsymbol{x}^2) &= I(\boldsymbol{x}^1; (H_m^1 \boldsymbol{x}^1 + \boldsymbol{z})|\boldsymbol{x}^2) \\ &= I(\boldsymbol{x}^1; (H_m^1 \boldsymbol{x}^1 + \boldsymbol{z})) \\ &= h(H_m^1 \boldsymbol{x}^1 + \boldsymbol{z}) - h(\boldsymbol{z}) \\ &= h(H_n \boldsymbol{x}^1 + \boldsymbol{z}^1) + h(\boldsymbol{z}^2) - \big(h(\boldsymbol{z}^1) + h(\boldsymbol{z}^2)\big) \\ &= nCap_n(\rho). \end{aligned}$$

Therefore,

$$mCap_m(\rho) \leq nCap_n(\rho) + (m-n)Cap_{m-n}(\rho).$$

Take $k \in \mathbb{N}^*$, by induction,

$$(nk)Cap_{nk}(\rho) \leq (nk)Cap_n(\rho).$$

Dividing by nk and taking k to infinity in the LHS gives the upper bound.

Lower bound. Take $k \in \mathbb{N}^*$ and consider $k(n+d)$ users and their corresponding antennas. For every group of $n+d$ users, silence the d last users and redistribute their power to the

n first users so that their SNR becomes $(n+d/n)\rho$ (the average SNR is still ρ). See Figure 1 for an illustration. Since asymptotically, the equal power distribution among the users is optimal [10, Appendix C], we get that the new system has a lower capacity. Therefore, for k going to infinity,

$$(nk)Cap_n\left(\frac{n+d}{n}\rho\right) \leq k(n+d)Cap_{k(n+d)}(\rho).$$

Dividing by $k(n+d)$ and taking k to infinity in the LHS gives the lower bound. ■

C. Numerical comparison of the bounds

We first compare the bound of Corollary V.2 and the bounds of Proposition V.3. Note that by the ergodic theorem, the upper-bound of Proposition V.3 grows like $\log d$, whereas in the bound of Corollary V.2, the part $\log\binom{2d-1}{d}\left|\zeta_{d+1,d+1}\zeta_{1,d+1}^{\dagger}\right|^2$ alone already grows like d. Nevertheless, it is not necessarily true that the upper-bound of Proposition V.3 is better that the one of Corollary V.2 for all d and all fading distributions.

In Figure 2, we present the bounds of Corollary V.2 and Proposition V.3 in the special case of Rayleigh fading (real and imaginary parts are independent Gaussian random variables with zero mean and variance $1/\sqrt{2}$). The curves are produced by Monte Carlo simulation with 10^6 samples. We see that in this case, even for d small, the upper-bound of Proposition V.3 is better than the one of Corollary V.2.

In Figures 3 and 4, we compare the bounds of Proposition V.1, point 1 and 3, Proposition V.3 and Proposition V.4 in the special case of Rayleigh fading (real and imaginary parts are independent Gaussian random variables with zero mean and variance $1/\sqrt{2}$). The curves are produced by Monte Carlo simulations with 10^5 samples.

Note that in the case $d=2$, for $K=1$, the bounds of Proposition V.1, point 3 are better than those of point 1, whereas for $K>1$, it is the opposite.

We see that in the case $d=2$, for $K=4$ and $K=10$, the upper-bound of Proposition V.3 is very close to the capacity and the upper-bounds of Proposition V.1.1 are getting tight very rapidly.

In the case $d=2$, we want to compare the random-fading channel with the non-fading channel. See Appendix D for the per-cell sum-rate capacity of the non-fading channel. The comparison is done in Figure 4; in the eight cases that we consider, the random-fading channel is better than the non random one.

VI. Results for particular cases in the high-SNR regime

A. Case $d = 1$

As a direct application of Proposition IV.4 in the case $d = 1$ and $K = 1$, we get the following result.

Proposition VI.1. *Assume (H1), (H2) and (H3). Then*

$$\mathcal{L}_\infty = \frac{-1}{\log 2} \left[2 \max \left(\mathbb{E}_{\pi_0} \log |\zeta_0| \; ; \mathbb{E}_{\pi_1} \log |\zeta_1| \right) \right].$$

Note that a similar result was already proved by other techniques in [5] under much stronger hypothesis, in particular, independence of the fading coefficients was assumed there. In contrary, our result depends only on the marginal distributions of the fading coefficients and is valid for a larger class of joint distributions.

We want to compare the per-cell sum-rate capacity of the random-fading and non-fading channels. For a random variable ζ, by Jensen's inequality,

$$\mathbb{E} \log |\zeta|^2 \leq \log \mathbb{E} |\zeta|^2 .$$

Therefore, under the constraints $\mathbb{E}_{\pi_0} |\zeta_0|^2 \leq 1$ and $\mathbb{E}_{\pi_1} |\zeta_1|^2 \leq 1$, the non-fading channel achieves the best per-cell sum-rate capacity in the high SNR regime.

B. Case $d = 2$

We now assume that $d = 2$ and $K = 1$ and that the fading coefficients have the following form; for $i \in \mathbb{N}^*$,

$$\zeta_{i-2,i} = \alpha a_i \ , \quad \zeta_{i-1,i} = \beta b_i \text{ and } \zeta_{i,i} = c_i,$$

where a_i, b_i and c_i are random variable distributed according to π_a, π_b and π_c respectively and α and β are parameters such that $\alpha > 0$ and $\beta \geq 0$. Moreover, take the following normalization that can always be achieved by modifying α and β.

$$\mathbb{E}_{\pi_a} \log |a_1| = \mathbb{E}_{\pi_b} \log |b_1| = \mathbb{E}_{\pi_c} \log |c_1| .$$

We use the notation of Proposition IV.4.

Proposition VI.2. *Assume that $(a_i, b_i, c_i)_{i\in\mathbb{N}^*}$ is a stationary ergodic sequence such that for all $i \in \mathbb{N}^*$, almost surely, a_i and c_i are non zero and that their exist $\varepsilon > 0$ such that $\mathbb{E}_{\pi_a}(\log|a_1|)^{1+\varepsilon}$, $\mathbb{E}_{\pi_b}(\log|b_1|)^{1+\varepsilon}$ and $\mathbb{E}_{\pi_c}(\log|c_1|)^{1+\varepsilon}$ are finite.*

Then, there exist a domain $\mathcal{D} \subset (0,1]\times[0,1]$ such that for all $(x,y) \in \mathcal{D}$, $(0,x)\times[0,y) \subset \mathcal{D}$ and for all $(\alpha,\beta) \in \mathcal{D}$, as ρ goes to infinity,

$$\mathcal{L}_\infty = \frac{-2}{\log 2}\mathbb{E}_{\pi_a}\log|a|\,. \tag{VI.3}$$

The proof is postponed to Appendix B4.

Remark VI.4.

1. The set $\mathcal{D}$ is not maximal in the sense that (VI.3) may hold for couples $(\alpha,\beta) \notin \mathcal{D}$.
2. Note that for $(\alpha,\beta) \in \mathcal{D}$, in the high-SNR regime, the lower bound of Proposition V.3 is tight.
3. The proof will yield an *effective* construction of $\mathcal{D}$, which allows us to find many points in $\mathcal{D}$. Indeed, we construct $(f_p)_{p\in\mathbb{N}^*}$ a family of functions on $(0,1]\times[0,1]$ with the following property: if there exists $p \in \mathbb{N}^*$ such that $f_p(\alpha,\beta) \leq 0$, then (VI.3) holds.
4. Note that (VI.3) does not hold when $\alpha > 1$, indeed, as it will appear in the course of the proof, as ρ goes to infinity,

$$Cap(\rho) \geq \log\rho + 2\mathbb{E}_{\pi_a}\log|a| + \log\alpha.$$

We conjecture that (VI.3) does not hold when $\beta > 1$ either.

Let us apply Proposition VI.2 to the case where $(a_i, b_i, c_i)_{i\in\mathbb{N}^*}$ are independent Rayleigh distributed coefficients. In Figure 5, we plot points for which f_{20} is less or equal to -0.05 (Monte Carlo simulations realized with 10^5 samples). Therefore, (VI.3) holds for (α,β) in the stripped region and in particular for $\alpha,\beta \leq 0.4$. Note that in this case, the power offset is

$$\mathcal{L}_\infty = \frac{\gamma}{\log 2},$$

where γ is the Euler constant.

C. Artificial fading

In the frame of non-fading channels, we consider artificial fading, that is, every user uses a pseudo-random fading and multiplies its signal by this artificial fading. The fading coefficients

then have the following form, for $i \in \mathbb{N}^*$ and $0 \leq s \leq d$,

$$\zeta_{i,i+s} = \alpha_s P_{i+s},$$

where $\alpha_0, \dots \alpha_d$ are non random positive numbers and $P_i\,, i \in \mathbb{N}^*$ are stationary ergodic pseudo-random complex row vectors of size K distributed according to a law denoted by π_P. We moreover assume that for all $i \in \mathbb{N}^*$, almost surely, the coefficients of P_i are non zero and that $\mathbb{E}_{\pi_P} \|P_1\|^2 = 1$.

In [11], it is proved that in the case $d = 2$, the per-cell sum-rate capacity is smaller with artificial fading. Indeed, had such a procedure helped, then it would be used in non-fading situations to enhance capacity. It is evident then that it is deleterious, as the expression in Proposition VI.5 exhibits.

We consider the high-SNR regime and derive the explicit influence of the artificial fading.

Proposition VI.5. *Denote by $\mathcal{L}_\infty^0$ the power off-set without artificial fading (that is, $P_i = (1, \dots, 1)$ almost surely) and by $\mathcal{L}_\infty$ the power off-set with artificial fading. Then,*

$$\mathcal{L}_\infty = \mathcal{L}_\infty^0 - \frac{1}{\log 2} \mathbb{E}_{\pi_P} \log \|P_1\|^2 .$$

Remark VI.6. By Jensen's inequality, we get that $\mathcal{L}_\infty \geq \mathcal{L}_\infty^0$, therefore, in the high-SNR regime, the per-cell sum-rate capacity is smaller with artificial fading.

Proof: We set until the end of the proof $\lambda = 0$. Using Corollary III.2 and Proposition IV.2,

$$\mathcal{L}_\infty = \frac{1}{\log 2} \left[\log K - \mathbb{E}_\pi \log \left| \alpha_0 \alpha_d^\dagger \right| - \mathbb{E}_{\pi_P} \log \|P_1\|^2 - \frac{1}{d} \gamma(\Delta) \right],$$

whereas

$$\mathcal{L}_\infty^0 = \frac{1}{\log 2} \left[\log K - \mathbb{E}_\pi \log \left| \alpha_0 \alpha_d^\dagger \right| - \frac{1}{d} \gamma\left(\widetilde{\Delta}\right) \right],$$

where $\left(\widetilde{\Delta}_i\right)_{i \in \mathbb{N}^*}$ denote the matrices without artificial fading. Therefore, we only have to prove that for $i \in \mathbb{N}^*$, Δ_i does not depend on $(P_i)_{i \in \mathbb{N}^*}$. In the case $K = 1$, $\Delta_i = \delta_{id} \cdots \delta_{(i-1)d+1}$ and for $i \geq d+1$, δ_i does not depend on $(P_i)_{i \in \mathbb{N}^*}$, therefore, Δ_i does not depend on $(P_i)_{i \in \mathbb{N}^*}$.

Let us assume $K > 1$. Using Proposition IV.2, we only have to prove that for $i \in \mathbb{N}^*$, Δ_i does not depend on $(P_i)_{i \in \mathbb{N}^*}$.

$$C_i = \begin{pmatrix} \alpha_0 P_{d(i-1)+1} & \alpha_1 P_{d(i-1)+2} & \cdots & \alpha_{d-1} P_{di} \\ 0 & \alpha_0 P_{d(i-1)+2} & \cdots & \alpha_{d-2} P_{di} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \alpha_0 P_{di} \end{pmatrix}$$

and

$$D_i = \begin{pmatrix} \alpha_d {P_{d(i-1)+1}}^\dagger & \alpha_{d-1} {P_{d(i-1)+1}}^\dagger & \cdots & \alpha_1 {P_{d(i-1)+1}}^\dagger \\ 0 & \alpha_d {P_{d(i-1)+2}}^\dagger & \cdots & \alpha_2 {P_{d(i-1)+2}}^\dagger \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \alpha_d {P_{di}}^\dagger \end{pmatrix}.$$

Let us define another channel transfer matrix $\widetilde{H}_m$ by $\widetilde{K} = 1$ and for $i \in \mathbb{N}^*$ and $0 \leq s \leq d$,

$$\widetilde{\zeta}_{i,i+s} = \alpha_s \left\| P_{i+s} \right\|.$$

In the same manner, we define $\widetilde{C}_i$, $\widetilde{D}_i$ and $\widetilde{\Delta_i}$. A straight forward verification shows that for $i \in \mathbb{N}^*$

$$C_i C_i^\dagger = \widetilde{C}_i \widetilde{C}_i^\dagger,\ C_i D_i = \widetilde{C}_i \widetilde{D}_i \text{ and } D_i^\dagger D_i = \widetilde{D}_i^\dagger \widetilde{D}_i.$$

Moreover, since Δ_i is a function of $C_i C_i^\dagger$, $C_i D_i$ and $D_i^\dagger D_i$, $\Delta_i = \widetilde{\Delta}_i$. However, since $\widetilde{K} = 1$, we have already proved that $\widetilde{\Delta}_i$ does not depend on $(P_i)_{i \in \mathbb{N}^*}$, therefore, Δ_i does not depend on $(P_i)_{i \in \mathbb{N}^*}$. ■

VII. Numerical simulations

A. Influence of the correlation

We assume that the fading coefficients are Rayleigh distributed (real and imaginary parts are independent Gaussian random variables with zero mean and variance $1/\sqrt{2}$) and independent for different users. We are interested in the following question, which of the non-fading channel and the Rayleigh fading channel gives a higher per-cell sum-rate capacity.

In the case $d = 2$, $\lambda = 0.1, 1$ and $K = 1, 2, 4, 10$, with all fading coefficients independent, it is shown in Subsection V-C that the Rayleigh fading is beneficial.

In the case $d = 1$ if we assume independence between the $\zeta_{i,j}$, it is known that Rayleigh fading is beneficial over non-fading channels in the high-SNR region already for $K = 2$ ([5]).

If we assume that for $i \in \mathbb{N}^*$, $\zeta_{i-1,i} = \zeta_{i,i}$, then, the sum-rate per-cell capacity is less than the one of a non-fading channel (see Subsection VI-C and [11]). We investigate the following question: what is the maximal level of correlation between $\zeta_{i-1,i}$ and $\zeta_{i,i}$ that still provides benefit over the non-fading channel. See Appendix D for the derivation of the capacity of the non-fading channels. We denote by c the correlation between the real (resp. imaginary) part of $\zeta_{i,i}$ and the real (resp. imaginary) part of $\zeta_{i-1,i}$

In Figure 6 we present the bounds of Proposition V.1.1 and Proposition V.4 in the special case of Rayleigh fading. In Figure 7 we present the bounds of Proposition V.1.1 and Proposition V.4 in the following special case: $\zeta_{i,i}$ is Rayleigh distributed, $\zeta_{i-1,i}$ is $\alpha \in [0,1]$ times a Rayleigh distributed random variable. In both cases, the curves are produced by Monte Carlo simulation with 10^5 samples.

We see that even with a correlation close to 1, fading still provides an advantage over non-fading channel. Moreover, note that K large, high SNR and α close to 1 are conditions in which the advantage of the fading is larger.

B. The asymmetric Wyner model

With the following specification, the model studied is the Rayleigh-fading Wyner model ([3]). We take $d = 2$ and the $\zeta_{i,j}$ independent with the following distributions. For $i \in \mathbb{N}^*$, $\zeta_{i,i+1}$ is Rayleigh distributed (real and imaginary parts are independent Gaussian random variables with zero mean and variance $1/\sqrt{2}$) and $\zeta_{i,i}$ (resp. $\zeta_{i,i+2}$) is $\alpha \in [0,1]$ times a Rayleigh distributed random variable. The asymmetric (Rayleigh-fading) Wyner model is similar to Rayleigh-fading Wyner with a slight modification. For $i \in \mathbb{N}^*$, $\zeta_{i,i}$ is Rayleigh distributed and $\zeta_{i,i+1}$ (resp. $\zeta_{i,i+2}$) is α times a Rayleigh distributed random variable. Note that in Subsection VI-B we prove that in the asymmetric case, the power offset for $\alpha \leq 0.4$ is $\gamma/\log 2$.

The two models are very similar and yet, in the non-fading case, the per-cell sum-rate capacity is notably different (see Appendix D for the derivation of the capacity of the non-fading channels). In Figure 8 we present the capacity of the two models without fading and the bounds of Proposition V.4 for the two models with Rayleigh fading. We study one case in moderate SNR ($\lambda = 1$) and one case in high SNR ($\lambda = 10^{-4}$). The curves are produced by Monte Carlo simulation with 10^5 samples.

Note that in the high-SNR region, for the non-fading channel, the per-cell sum-rate capacity is very different for symmetric and the asymmetric models, whereas the per-cell sum-rate capacities for the symmetric and asymmetric Rayleigh-fading models are very close (but not equal as shown in Figure 9 for $\lambda = 10^{-4}$ and $\alpha = 0.5$).

To understand better the influence of fading on the difference between the two models, we present in Figure 10 the bounds of Proposition V.4 for the capacity of the two models (symmetric and asymmetric) with the following fading: the modulus is uniformly distributed between $1-\varepsilon$ and $1+\varepsilon$ and the phase is uniformly distributed between 0 and $2\varepsilon\pi$, where ε is a parameter between 0 and 1. Note that for $\varepsilon = 0$, there is no fading and for $\varepsilon = 1$, the fading is uniformly distributed on the disc of center 0 and of radius 2. The curves are produced by Monte Carlo simulation with 10^5 samples. We notice that the difference between the two models decreases between $\varepsilon = 0$ and $\varepsilon = 0.5$ and that in high-SNR, it increases slightly between $\varepsilon = 0.5$ and $\varepsilon = 1$.

VIII. Concluding Remarks

In this paper, we study the per-cell sum-rate capacity of a channel communication with multiple cell processing. The main tools is a version of the Thouless formula for the strip which we prove in the article. It allows us to prove that the per-cell sum-rate capacity converges as the number of cells and antennas goes to infinity. We give several expressions of the limiting capacity in terms of Lyapunov exponents and several bounds on the per-cell sum-rate capacity.

We apply those results to several examples of communication channels and get insight on the evolution of the capacity as a function of the key parameters of the problem. In particular, in the high-SNR regime, some explicit formulas are derived.

Note that the model here applies verbatim to randomly varying intersymbol interference channels.

Some of the tools of this article can be used to derive CLT-type results on the capacity in order to study the outage-probability. Details will appear elsewhere [12].

Acknowledgments

We thank Oren Somekh for his help with the derivation of the capacity of the non-fading channels.

This research was partially supported by Technion Research Funds, the REMON Consortium, a grant from the Israel Science Foundation and NSF grant DMS-0503775.

Appendix

A. Random Schrödinger operators techniques

1) Lyapunov exponents theory: We use the theory of product of random matrices. For a general introduction to the aspects of the theory we use here, the reader may consult [8], [13], [14], [15], [16] or [17]. See appendix C for the relevant background on exterior products.

Theorem A.1 (Furstenberg H., Kesten H. (1960))**.** *Consider a stationary ergodic sequence of complex random matrices $(X_i)_{i\in\mathbb{N}^*}$ of size p and any norm on the matrices. Assume moreover that*

$$\mathbb{E}\log^+\|X_1\| < \infty,$$

then a.s, $n^{-1}\log\|X_n\cdots X_1\|$ converges to a constant:

$$\lim_{n\to\infty}\frac{1}{n}\log\|X_n\cdots X_1\| \triangleq \gamma(X).$$

We define p constants $\gamma_1(X),\ldots,\gamma_p(X)$ such that for $1\le i\le p$,

$$\gamma\left(\bigwedge^i X\right) = \gamma_1(X)+\cdots+\gamma_i(X).$$

Proposition A.2.

$$\gamma_1(X)\ge\cdots\ge\gamma_p(X).$$

The constants $\gamma_1(X)\ge\cdots\ge\gamma_p(X)$ are called the Lyapunov exponents and $\gamma(X)=\gamma_1(X)$ is called the top Lyapunov exponent.

We will also use the three following properties:

1. For any sub-multiplicative norm, for $p\in\mathbb{N}^*$

$$\gamma(X)\le\frac{1}{p}\mathbb{E}\log\|X_p\cdots X_1\|, \tag{A.3}$$

   and the limit of the RHS as p goes to infinity is $\gamma(X)$.

2.

$$\frac{1}{p}\mathbb{E}\log|\det X_1|\le\gamma(X). \tag{A.4}$$

3. Assume that the matrices $(X_i)_{i\in\mathbb{N}^*}$ are i.i.d, then for all $1 \leq i \leq p$, $\gamma_i(X) = \gamma_i(X^\dagger)$.

Finally, we quote the following proposition [18, Proposition 1].

Proposition A.5. *Consider a stationary ergodic sequence of complex random matrices $(X_i)_{i\in\mathbb{N}^*}$ of size p and any norm on the matrices. Assume moreover that*

$$\mathbb{E}\log^+ \|X_1\| < \infty.$$

Finally, assume that there exist three sequences of random matrices $(X_i^1)_{i\in\mathbb{N}^*}$, $(X_i^2)_{i\in\mathbb{N}^*}$, $(X_i^3)_{i\in\mathbb{N}^*}$, of respective sizes $k\times k$, $(p-k)\times k$ and $(p-k)\times(p-k)$, for $1 \leq k \leq p-1$, such that almost surely, for all i

$$X_i = \left(\begin{array}{c|c} X_i^1 & 0_{k,p-k} \\ \hline X_i^2 & X_i^3 \end{array}\right).$$

Then, $\gamma_1(X), \ldots, \gamma_p(X)$ is equal up to the order to the sequence

$$\gamma_1(X^1), \ldots, \gamma_k(X^1), \gamma_1(X^3), \ldots, \gamma_{p-k}(X^3).$$

2) Proof of Theorem III.1.1: In order to prove point 1 of Theorem III.1, we first prove a slightly more general lemma.

Lemma A.6. *Assume (H1), (H2) and (H3). For all $\lambda \in \mathbb{C}$ such that $\lambda \notin \mathbb{R}^-$, almost surely,*

$$\frac{1}{dn}\log|\det G_{dn}| \xrightarrow[n\to\infty]{} \frac{1}{d}\mathbb{E}\log|\det(C_2D_2)| + \frac{1}{d}\gamma(N).$$

Proof: For $i \in \mathbb{N}^*$, set $B_i = C_iC_i^\dagger + D_{i+1}^\dagger D_{i+1} + \lambda\,\mathrm{Id}_d$ and $A_i = C_iD_i$. Note that the eigenvalues of G_{dn} are bounded away from zero. To compute $\log|\det G_{dn}|$, we write the following decomposition: $G_{dn}U_{dn} = L_{dn}$, where U_{dn} is the upper triangular by block matrix

$$U_{dn} = \left(\begin{array}{c|c|c|c} X_1 & X_1 & \cdots & X_1 \\ \hline 0_d & X_2 & \cdots & X_2 \\ \hline \vdots & \ddots & \ddots & \vdots \\ \hline 0_d & \cdots & 0_d & X_n \end{array}\right),$$

the X_i are $d\times d$ matrices such that $X_0 = 0_d$, $X_1 = \mathrm{Id}_d$, and for $i \geq 1$,

$$A_iX_{i-1} + B_iX_i + A_{i+1}^\dagger X_{i+1} = 0_d. \tag{A.7}$$

L_{dn} is the lower triangular by block matrix

$$\left(\begin{array}{c|c|c|c} -A_2^\dagger X_2 & 0_d & \cdots & 0_d \\ \hline A_2X_1 & -A_3^\dagger X_3 & \ddots & \vdots \\ \hline 0_d & \ddots & \ddots & 0_d \\ \hline 0_d & 0_d & A_nX_{n-1} & -A_{n+1}^\dagger X_{n+1} \end{array}\right),$$

That decomposition allows us to write $\log|\det G_{dn}|$ as a determinant by block,

$$\begin{aligned}\log|\det G_{dn}| &= \sum_{i=1}^n \log|\det A_{i+1}| + \sum_{i=1}^n \log|\det X_{i+1}| - \sum_{i=1}^n \log|\det X_i| \\ &= \sum_{i=1}^n \log|\det A_{i+1}| + \log|\det X_{n+1}|\,.\end{aligned}$$

Therefore

$$\frac{1}{dn}\log|\det G_{dn}| = \frac{1}{dn}\sum_{i=1}^n \log|\det A_{i+1}| + \frac{1}{dn}\log|\det X_{n+1}|\,. \tag{A.8}$$

$\frac{1}{n}\sum_{i=1}^n \log|\det A_{i+1}|$ converges by ergodicity toward $\mathbb{E}\log|\det A_2|$. Note that the choice of A_{n+1} is arbitrary, indeed, if we take another value, say $\widetilde{A}_{n+1}$, then $\widetilde{A}_{n+1}^\dagger \widetilde{X}_{n+1} = A_{n+1}^\dagger X_{n+1}$ and (A.8) stays unchanged.

We emphasize that the derivation of (A.8) is inspired by Narula's thesis ([19]).

The X_i are defined by (A.7). We can reformulate it in the following way. Set $V_i = \left(\begin{array}{c} X_{i-1} \\ \hline X_i \end{array}\right)$, then (A.7) is equivalent to $V_{i+1} = M_iV_i$ and moreover,

$$X_{n+1} = \left(\begin{array}{c|c} 0_d & \mathrm{Id}_d \end{array}\right) M_n\cdots M_1 \left(\begin{array}{c} 0_d \\ \hline \mathrm{Id}_d \end{array}\right).$$

Denote $f = \bigwedge^d \left(\begin{array}{c} 0_d \\ \hline \mathrm{Id}_d \end{array}\right)$. For the relevant background on exterior products, see Appendix C. We get

$$N_nN_{n-1}\ldots N_1\left(v_1^1\wedge\cdots\wedge v_1^d\right).$$

However, $v_1^1\wedge\cdots\wedge v_1^d = f$. Therefore,

$$\begin{aligned}\det X_{n+1} = \bigwedge^d X_{n+1} &= \bigwedge^d \left(\left(\begin{array}{c|c} 0_d & \mathrm{Id}_d \end{array}\right) M_n\cdots M_1 \left(\begin{array}{c} 0_d \\ \hline \mathrm{Id}_d \end{array}\right)\right) \\ &= f^\dagger N_nN_{n-1}\ldots N_1 f.\end{aligned} \tag{A.9}$$

Taking the canonical basis of $\bigwedge^d \mathbb{C}^{2d}$, f is the last vector of the basis and $\det X_{n+1}$ grows like the bottom-right coefficient of the product of the N_i, therefore, its growth rate is bounded above by the Lyapunov exponent of the N_i.

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\log|\det X_{n+1}| \leq \gamma(N). \tag{A.10}$$

Using (A.8), it is enough to prove the opposite inequality to conclude the proof. The end of the proof is inspired by [20].

Lemma A.11. *If there exist a basis of $\bigwedge^d \mathbb{C}^{2d}$, say $(g_i)_{i\in I}$, such that for all $i,j \in I$, almost surely,*

$$\liminf_{n\to\infty} \frac{1}{n}\log\left|g_j^\dagger N_n \cdots N_1 g_i\right| \leq \liminf_{n\to\infty} \frac{1}{n}\log|\det X_{n+1}|\,, \tag{A.12}$$

then, almost surely,

$$\gamma(N) \leq \liminf_{n\to\infty} \frac{1}{n}\log|\det X_{n+1}|\,.$$

Let us first prove the lemma.

Proof: For any finite basis S_1 and S_2 in a vector space, we have for all A

$$\sup_{\alpha\in S_1,\beta\in S_2} \left|\alpha^\dagger A\beta\right| \geq c\,\|A\|$$

for some universal c. Thus, (A.12) shows that, almost surely,

$$\gamma(N) \leq \liminf_{n\to\infty} \frac{1}{n}\log|\det X_{n+1}|\,.$$

■

To finish the proof of Lemma A.6, we denote by $\{e_1,\ldots,e_{2d}\}$ the canonical basis of $\mathbb{C}^{2d}$ and we apply the lemma with the following spanning system of $\bigwedge^d \mathbb{C}^{2d}$

$$S \triangleq \{(e_1+\widetilde{e}_1)\wedge\cdots\wedge(e_d+\widetilde{e}_d)\ ,\ \widetilde{e}_1,\ldots,\widetilde{e}_d \in \mathrm{vect}(e_{d+1},\ldots,e_{2d})\}.$$

For a choice of $e_1^\#,\ldots,e_d^\#$, such that for $1\leq j\leq d$, $e_j^\# = \sum_{i=1}^d \alpha_{i,j}e_{d+i}$, we define $E^\#$ the $d\times d$ matrix of the $\alpha_{i,j}$. We get

$$g_1 \triangleq (e_1+e_1^\#)\wedge\cdots\wedge(e_d+e_d^\#) = \bigwedge^d \begin{pmatrix} Id_d \\ E^\# \end{pmatrix}.$$

In the same way, for a choice of $e_1^b \ldots, e_d^b \in \text{vect}(e_{d+1}, \ldots, e_{2d})$, we define E^b, a $d \times d$ matrix, such that

$$g_2 \triangleq (e_1 + e_1^b) \wedge \cdots \wedge (e_d + e_d^b) = \bigwedge^d \begin{pmatrix} Id_d \\ \hline E^b \end{pmatrix}.$$

We define two new sequences $(\widetilde{A}_i)$ and $(\widetilde{B}_i)$ such that

- For $2 \leq i \leq n-1$, $\widetilde{B}_i = B_i$,
- For $1 \leq i \leq n$, $\widetilde{A}_i = A_i$,
- $\widetilde{B}_1 = -A_2^\dagger E^\#$,
- $\widetilde{B}_n = A_n \left(E^b\right)^\dagger$,
- $\widetilde{A}_{n+1} = -A_n^\dagger$.

We also define $\widetilde{G}_{dn}$, $\widetilde{M}_i$, $\widetilde{N}_i$ and $\widetilde{X}_i$ using $\left(\widetilde{A}_i\right)_{i\in\mathbb{N}^*}$ and $\left(\widetilde{B}_i\right)_{i\in\mathbb{N}^*}$. Then,

$$\begin{aligned} f^\dagger \widetilde{N}_n \widetilde{N}_{n-1} \ldots \widetilde{N}_1 f &= f^\dagger \bigwedge^d \left(\begin{array}{c|c} 0_d & \mathrm{Id}_d \\ \hline \mathrm{Id}_d & \left(E^b\right)^\dagger \end{array}\right) \widetilde{N}_{n-1} \ldots \widetilde{N}_2 \bigwedge^d \left(\begin{array}{c|c} 0_d & \mathrm{Id}_d \\ \hline -T_1 T_2^{-1\dagger} & E^\# \end{array}\right) f \\ &= g_2^\dagger \widetilde{N}_{n-1} \ldots \widetilde{N}_2 g_1. \end{aligned}$$

Therefore, to prove the condition (A.12), it is enough to prove that, almost surely,

$$\limsup_{n\to\infty} \left(\frac{1}{n} \log \left| \det \widetilde{G}_{dn} \right| - \frac{1}{n} \log \left| \det G_{dn} \right| \right) \leq 0.$$

We now use perturbation theory techniques. Indeed, we denote by ρ the spectral radius of a matrix, i.e. its largest eigenvalue in absolute value. Recall that for a matrix S, $\rho(S) \leq \sqrt{\rho(SS^\dagger)}$ and that $\sqrt{\rho(SS^\dagger)}$ is a sub-multiplicative norm. As a consequence, for positive Hermitian matrices, the spectral radius is sub-multiplicative. Moreover, we denote by $\|\cdot\|_F$ the Fröbenius norm. Recall that $\sqrt{\rho(SS^\dagger)} \leq \|S\|_F$. We will also use the fact that the eigenvalues of G_{dn} are bounded away from 0 by $\mu = \lambda$ if $\lambda > 0$ or $\mu = |\Im\lambda|$ if $\lambda \notin \mathbb{R}$. Moreover, we define $U_{dn} = \widetilde{G}_{dn} - G_{dn}$, which has rank less than or equal to $2d$.

$$\begin{aligned} \frac{1}{n} \log \left| \det \widetilde{G}_{dn} \right| - \frac{1}{n} \log \left| \det G_{dn} \right| &= \frac{1}{n} \log \left| \det(G_{dn} + U_{dn}) \right| - \frac{1}{n} \log \left| \det G_{dn} \right| \\ &= \frac{1}{n} \log \left| \det(\mathrm{Id}_{dn} + G_{dn}^{-1} U_{dn}) \right|. \end{aligned}$$

$G_{dn}^{-1} U_{dn}$ has rank at most 2d, therefore,

$$\frac{1}{n} \log \left| \det \widetilde{G}_{dn} \right| - \frac{1}{n} \log \left| \det G_{dn} \right| \leq \frac{2d}{n} \log \left| 1 + \rho(G_{dn}^{-1} U_{dn}) \right|$$

$$\leq \frac{2d}{n}\log\left|1+\sqrt{\rho(G_{dn}^{-1}U_{dn}U_{dn}^{\dagger}G_{dn}^{-1\dagger})}\right|$$
$$\leq \frac{2d}{n}\log\left|1+\sqrt{\rho(G_{dn}^{-1\dagger}G_{dn}^{-1})}\sqrt{\rho(U_{dn}U_{dn}^{\dagger})}\right|$$
$$\leq \frac{2d}{n}\log\left|1+\frac{1}{\mu}\sqrt{\rho(U_{dn}U_{dn}^{\dagger})}\right|$$
$$\leq \frac{2d}{n}\log\left|1+\frac{1}{\mu}\left\|U_{dn}\right\|_F\right|.$$

Moreover,

$$\|U_{dn}\|_F^2 = \left\|S_1+\lambda\,\mathrm{Id}_d+T_2^{\dagger}E^{\#}\right\|_F^2+\left\|S_n+\lambda\,\mathrm{Id}_d+T_n^{\dagger}E^{b}\right\|_F^2,$$

hence, with the integrability condition, $\sup_n \mathbb{E}\log\left|1+\frac{1}{\mu}\left\|U_{dn}\right\|_F\right|^{1+\varepsilon} < \infty$. By Tchebicheff inequality, for a given $\eta>0$,

$$\mathbb{P}\left(\frac{1}{n}\log\left|1+\frac{1}{\mu}\left\|U_{dn}\right\|_F\right|>\eta\right)\leq\frac{\sup_n\mathbb{E}\log\left|1+\frac{1}{\mu}\left\|U_{dn}\right\|_F\right|^{1+\varepsilon}}{(\eta n)^{1+\varepsilon}}.$$

The RHS is a summable series, therefore, by Borel-Cantelli Lemma, almost surely,

$$\limsup_{n\to\infty}\left(\frac{1}{n}\log\left|\det\widetilde{G}_{dn}\right|-\frac{1}{n}\log\left|\det G_{dn}\right|\right)\leq 0.$$

This finishes the proof of Lemma A.6. ■

Remark A.13. By the same kind of perturbation theory techniques, we can show that in order to study the limit in m of $Cap_m(\rho)$, it is enough to study the sequence every d steps.

For a hermitian matrix h whose ordered eigenvalues are $\alpha_1,\ldots,\alpha_n$, we denote by the *spectral distribution* of h, the measure

$$\frac{1}{n}\sum_{i=1}^{n}\delta_{\alpha_i},$$

where δ_x is a Dirac measure at x.

The following technical lemma will be used several times to prove domination properties.

Lemma A.14. *Denote by μ_n the spectral distribution of $H_{dn}H_{dn}^{\dagger}$. Consider the following diagonal by blocks matrix:*

$$F_{dn}\triangleq\left(\begin{array}{c|c|c|c} 2B_1 & 0_d & \cdots & 0_d \\ \hline 0_d & 2B_2 & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & 0_d \\ \hline 0_d & \cdots & 0_d & 2B_n \end{array}\right),$$

and denote by $\tilde{\mu}_n$ its spectral distribution. Then, for any non-decreasing function f,

$$\int f d\mu_n \leq \int f d\tilde{\mu}_n.$$

Proof: Denote

$$\tilde{H}_{dn} = \left(\begin{array}{c|c|c|c|c} C_1 & -D_2^\dagger & 0_{d,dK} & \cdots & 0_{d,dK} \\ \hline 0_{d,dK} & C_2 & -D_3^\dagger & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \ddots & 0_{d,dK} \\ \hline 0_{d,dK} & \cdots & 0_{d,dK} & C_n & -D_{n+1}^\dagger \end{array}\right),$$

then $F_{dn} = H_{dn}H_{dn}^\dagger + \tilde{H}_{dn}\tilde{H}_{dn}^\dagger$. Since $\tilde{H}_{dn}\tilde{H}_{dn}^\dagger$ is a non-negative Hermitian matrix, by Weyl's inequalities, for all $1 \leq i \leq dn$, the i-th eigenvalue of $H_{dn}H_{dn}^\dagger$ is less or equal than the i-th eigenvalue of F_{dn}. ■

First note that $(1/d)\mathbb{E}\log|\det(C_2D_2)| = \mathbb{E}_\pi \log\left|\zeta_0\zeta_d^\dagger\right|$. From Lemma A.14, we deduce that for $\lambda > 0$,

$$\frac{1}{dn}\log|\det G_{dn}| \leq \log 2 + \frac{1}{n}\sum_{i=1}^{n}\log|\det B_i|\,.$$

Therefore by (H2) and Hadamard's inequality, $(1/dn)\log|\det G_{dn}|$ is a uniformly integrable sequence and the almost sure convergence of Lemma A.6 implies point 1 of Theorem III.1.

3) Proof of Theorem III.1.2: We begin by a few notations. For $\lambda \notin \mathbb{R}$, set

$$f(\lambda) = \lim_{n\to\infty}\frac{1}{dn}\log\left|\det\left(H_{dn}H_{dn}^\dagger + \lambda\,\mathrm{Id}_{dn}\right)\right|,$$

which exists by Lemma A.6. The existence of the weak limit of μ_n and the fact that it is non random is a classical fact of the random Schrödinger operators theory, see for example [16, Theorem 4.4]. For $\lambda \in \mathbb{C}$, we set (if it exists)

$$g(\lambda) = \int \log|x+\lambda|\,d\mu(x).$$

We emphasize that since log is not a bounded function, we cannot directly deduce from Lemma A.6 and the weak convergence of the μ_n to μ that for $\lambda \notin \mathbb{R}$, $f(\lambda) = g(\lambda)$.

Finally, for $\lambda \in \mathbb{C}$, define

$$h(\lambda) = \mathbb{E}_\pi \log\left|\zeta_0\zeta_d^\dagger\right| + \frac{1}{d}\gamma\left(N\right).$$

The following lemma is a generalization of the Thouless formula for the strip proved in [20].

Lemma A.15. *Assume (H1), (H2) and (H3). For all $\lambda \in \mathbb{C}$, $g(\lambda) = h(\lambda)$.*

The proof of this result is done in the frame of channel transfer matrices but one does not need to assume that the A_i and the B_i are upper triangular by blocks, one just need instead of (H3) the hypothesis that almost surely, $A_i B_i$ is invertible.

Proof: The proof goes along the following lines, we first prove that for $\lambda \notin \mathbb{R}$, $g(\lambda)$ exists and equals to $h(\lambda)$, then, following [21] we argue that g and h are two subharmonic functions equal everywhere except a set of 0 measure, therefore they are equal everywhere.

Step 1: Let us first prove that for $\lambda \notin \mathbb{R}$, $g(\lambda)$ is well defined. $\log|x+\lambda|$ is bounded away from $-\infty$, therefore, $g(\lambda)$ exists although it may be ∞. For $R \geq 0$, let us denote by $\log_R$ the function $t \to \log(t) \wedge R$. By monotone convergence, it is enough to prove that $\int \log_R|x+\lambda|\, d\mu(x)$ is bounded uniformly in R. Since $x \to \log_R|x+\lambda|$ is a bounded continuous function,

$$\int \log_R|x+\lambda|\, d\mu(x) = \lim_{n\to\infty} \int \log_R|x+\lambda|\, d\mu_n(x).$$

By Lemma A.14, and using that $\int \log_R|x+\lambda|\, d\mu_n(x) \leq \int \log|x+\lambda|\, d\mu_n(x)$,

$$\lim_{n\to\infty} \int \log_R|x+\lambda|\, d\mu_n(x) \leq \lim_{n\to\infty} \int \log|x+\lambda|\, d\tilde{\mu}_n(x) = \mathbb{E}\log|\det B_1| < \infty,$$

where the last inequality comes from (H2) and Hadamard's inequality. Finally, we get that for $\lambda \notin \mathbb{R}$,

$$g(\lambda) \leq \mathbb{E}\log|\det B_1| < \infty.$$

Step 2: Let us prove that for $\lambda \notin \mathbb{R}$, $f(\lambda) = g(\lambda)$. Applying Lemma A.14 one shows that for $\lambda \notin \mathbb{C}$, the sequence $\left(\int \log|x+\lambda|\, d\mu_n(x)\right)_{n\in\mathbb{N}^*}$ is uniformly integrable and therefore,

$$f(\lambda) = \lim_{n\to\infty} \mathbb{E}\int \log|x+\lambda|\, d\mu_n(x).$$

By Lemma A.14, for $R \geq 0$,

$$\begin{aligned}\mathbb{E}\int_{x\geq R} \log|x+\lambda|\, d\mu_n(x) &\leq \mathbb{E}\int_{x\geq R} \log|x+\lambda|\, d\tilde{\mu}_n(x) \\ &= \mathbb{E}\int_{x\geq R} \log|x+\lambda|\, d\tilde{\mu}_1(x).\end{aligned}$$

Therefore, for $n \in \mathbb{N}^*$ and $R \geq 0$,

$$\left|\mathbb{E}\int \log|x+\lambda|\, d\mu_n(x) - g(\lambda)\right| \leq \left|\mathbb{E}\int_{x\geq R} \log|x+\lambda|\, d\tilde{\mu}_1(x)\right|$$

$$+\left|\mathbb{E}\int \log_R|x+\lambda|\,d\mu_n(x)-\mathbb{E}\int \log_R|x+\lambda|\,d\mu(x)\right|$$
$$+\left|\mathbb{E}\int_{x\geq R}\log|x+\lambda|\,d\mu(x)\right|.$$

We first fix $R\geq 0$ such that the first and the third terms are arbitrary small and then, by weak convergence, the second term goes to 0 as n goes to infinity. Therefore, for $\lambda\notin\mathbb{R}$, $f(\lambda)=g(\lambda)$ and by Lemma A.6, $g(\lambda)=h(\lambda)$.

Step 3: Let us prove that g and h are subharmonic on $\mathbb{C}$. See [21] for the relevant definitions. Since for $i\in\mathbb{N}^*$, $N_i(\lambda)$ is an entire function of λ, h is subharmonic ([21]).

Let us prove that g is subharmonic. For $R\geq 0$, set

$$g_R(\lambda)\triangleq\int(\log|x+\lambda|\vee -R)\,d\mu(x).$$

By Lemma A.14, g_R a continuous function. As R goes to infinity, g_R is a decreasing sequence of functions converging point wise to g, therefore, g is subharmonic.

The functions g and h are subharmonic on $\mathbb{C}$ and equal on $\mathbb{C}-\mathbb{R}$, therefore, g and h are equal on $\mathbb{C}$. ■

To finish the proof of point 2 of Theorem III.1, let us prove that $h(\lambda)$ converges to $h(0)$ when λ goes to 0 in $\mathbb{R}^+$. Note that $\mathbb{E}\log|\det M_1|=0$, therefore, $\mathbb{E}\log|\det N_1|=0$. By (A.4), $h(0)\geq\mathbb{E}_\pi\log\left|\zeta_0\zeta_d^\dagger\right|$, therefore, using Lemma A.14 and the fact that for $\lambda,x\in\mathbb{R}^+$, $\log|x+\lambda|\geq\log|x|$, we get the desired result.

B. Other proofs

1) Proof of Proposition IV.1: We use the notation of Subsection A2. We define x_j^i for $i\in\mathbb{N}$ and $1\leq j\leq d$ such that the the element at the position (s,t) of X_i is $x_t^{(i-1)d+s}$. Recall that $G_{dn}U_{dn}=L_{dn}$. Therefore, for a given j such that $1\leq j\leq d$, we get the following characterization of the sequence $(x_j^i)_i$. $x_j^i=0$ for $-d+1\leq i\leq 0$, $x_j^i=\delta_{i,j}$ for $1\leq i\leq d$ and for $i\geq d+1$,

$$\sum_{l=i-d}^{i+d}\widetilde{\zeta}_{i,l}x_j^l=0. \tag{A.16}$$

Therefore,

$$\begin{pmatrix}x_j^{i-d+1}\\ \vdots\\ x_j^{i+d}\end{pmatrix}=\mathfrak{m}_i\begin{pmatrix}x_j^{i-d}\\ \vdots\\ x_j^{i+d-1}\end{pmatrix}.$$

Moreover,

$$\begin{aligned} V_{i+1} &= \begin{pmatrix} x_1^{(i-1)d+1} & \cdots & x_d^{(i-1)d+1} \\ \vdots & & \vdots \\ x_1^{(i+1)d} & \cdots & x_d^{(i+1)d} \end{pmatrix} \\ &= \mathfrak{m}_{id} \cdots \mathfrak{m}_{(i-1)d+1} \begin{pmatrix} x_1^{(i-2)d+1} & \cdots & x_d^{(i-2)d+1} \\ \vdots & & \vdots \\ x_1^{id} & \cdots & x_d^{id} \end{pmatrix} \\ &= \mathfrak{m}_{id} \cdots \mathfrak{m}_{(i-1)d+1} V_i. \end{aligned}$$

Therefore, together with A.7, it proves Proposition IV.1.

2) Proof of Proposition IV.2.2: In order to prove point 2 of Proposition IV.2, we first prove the following lemma:

Lemma A.17. *For all $i \geq d+1$, there exist matrices $p_1^s(i)$, $p_2^s(i)$ for $1 \leq s \leq d$ and $\delta^s(i)$ for $1 \leq s \leq d+1$ such that $\delta^1(i) = \mu(i)$, $\delta^s(i) = \underline{\delta}^s(\zeta^i, \ldots, \zeta^{i+d-s+1})$, $p_1^s(i) = \underline{p}_1^s(\zeta^i, \ldots, \zeta^{i+d-s})$ and $p_2^s(i) = \underline{p}_2^s(\zeta^{i+d-s})$, where $\underline{\delta}^s$, $\underline{p}_1^s$ and $\underline{p}_2^s$ are deterministic functions. We have moreover the two relationships*

$$\delta^s(i) = p_2^s(i+1) p_1^s(i). \tag{A.18}$$

$$\delta^{s+1}(i) = p_1^s(i) p_2^s(i). \tag{A.19}$$

Finally, for $i \geq d+1$, $\delta_i = \delta^{d+1}(i)$

Proof: For $i \geq d+1$ and $1 \leq s \leq d$, define

- for $s \leq l \leq 2d$,
$$a_l^{i,s} = -\lambda \mathbf{1}_{(l=d+s)} - \sum_{t=(l-s-d)\vee 0}^{(l-s)\wedge(d-s)} \zeta_{i,i+t} \zeta^\dagger_{i+l-d-s,i+t},$$
- for $1 \leq l \leq d$, $\alpha_l^{i-s} = -\zeta^\dagger_{i-s+l-1,i+d-s} / \zeta^\dagger_{i+d-s,i+d-s}$,
- for $1 \leq l \leq s$, $b_l^{i-s} = -\zeta_{i-s+l,i+d-s} / \zeta_{i-s,i+d-s}$,
- $\beta^{i-s=} 1/\zeta_{i-s,i+d-s} \zeta^\dagger_{i+d-s,i+d-s}$.

Then

$$p_1^s(i) = \left(\begin{array}{c|cccc} \mathrm{Id}_{s-1} & 0_{s-1,2d-s+1} \\ \hline 0_{1,s-1} & a_s^{i,s} & \cdots & a_{2d}^{i,s} \\ \hline 0_{2d-s,s} & \mathrm{Id}_{2d-s} \end{array}\right)$$

and

$$p_2^s(i) = \left(\begin{array}{c|ccc} b_1^{i-s} & & & \\ \vdots & & & \\ b_{s-1}^{i-s} & & & \\ \hline & & \mathrm{Id}_{2d-1} & \\ 0_{2d-s,1} & & & \\ & & & \\ \hline \beta^{i-s} & 0_{1,d-1} & \alpha_1^{i-s} & \cdots \quad \alpha_d^{i-s} \end{array}\right).$$

Finally, for $2 \le s \le d$,

$$\delta^s(i) = \left(\begin{array}{c|c|cccc} b_1^{i-s+1} & \mathrm{Id}_{s-2} & 0_{s-2,2d-s+1} \\ \vdots & & \\ b_{s-1}^{i-s+1} & & a_s^{i,s} & \cdots & \cdots & a_{2d}^{i,s} \\ \hline & & 0_{2d-s,1} & \multicolumn{3}{c}{\multirow{2}{*}{Id_{2d-s}}} \\ 0_{2d-s,1} & 0_{2d-s+2,s-2} & & & & \\ \hline \beta^{i-s+1} & & 0_{1,d-s+1} & \alpha_1^{i-s+1} & \cdots & \alpha_d^{i-s+1} \end{array}\right).$$

A (straight forward yet tedious) verification shows that (A.18) and (A.19) are satisfied.

■

Note that in the proof, we make a choice of particular $\underline{p}_1^s$, $\underline{p}_2^s$ and $\underline{\delta}^s$. Point 2 of Proposition IV.2 is a direct consequence of the following lemma

Lemma A.20. *For all $i \in \mathbb{N}^*$,*

$$\Delta_i = \delta_{id} \cdots \delta_{(i-1)d+1}.$$

Therefore,

$$\Xi_i = \xi_{id} \cdots \xi_{(i-1)d+1}.$$

Proof: With the matrices of Lemma A.17, we can transform the product of the μ_i using alternatively (A.18) and (A.19).

$$\begin{aligned} M_i &= P_2(i+1)P_1(i) \\ &= P_2(i+1)\Delta(i)\left(P_2(i)\right)^{-1}, \end{aligned}$$

$$\begin{aligned} &\mu_{id} \cdots \mu_{(i-1)d+1} \\ &= \delta^1(id) \cdots \delta^1((i-1)d+1) \\ &= p_2^1(id+1)p_1^1(id)p_2^1(id)p_1^1(id-1) \cdots p_2^1((i-1)d+2)p_1^1((i-1)d+1) \\ &= p_2^1(id+1)\delta^2(id) \cdots \delta^2((i-1)d+1)\left(p_2^1((i-1)d+1)\right)^{-1} \\ &= p_2^1(id+1) \cdots p_2^d(id+1)\delta^{d+1}(id) \cdots \delta^{d+1}((i-1)d+1)\left[p_2^1((i-1)d+1) \cdots p_2^d((i-1)d+1)\right]^{-1}, \end{aligned}$$

where the last equality is proved by induction.

Therefore

$$\begin{aligned} &P_2(i+1)\Delta(i)\left(P_2(i)\right)^{-1} = \\ &p_2^1(id+1) \cdots p_2^d((id+1))\delta_{id} \cdots \delta_{(i-1)d+1}\left(p_2^1((i-1)d+1) \cdots p_2^d((i-1)d+1)\right)^{-1} \end{aligned}$$

and

$$\begin{aligned} &\left[p_2^1(id+1) \cdots p_2^d((id+1))\right]^{-1} P_2(i+1) = \\ &\delta_{id} \cdots \delta_{(i-1)d+1}\left[p_2^1((i-1)d+1) \cdots p_2^d((i-1)d+1)\right]^{-1} P_2(i)\left(\Delta(i)\right)^{-1} \end{aligned} \tag{A.21}$$

At this point, we emphasize that their exist a deterministic matrix valued function $\underline{\Delta}$ such that for all $i \in \mathbb{N}^*$, $\Delta_i = \underline{\Delta}\left(\zeta^{d(i-1)+1}, \ldots, \zeta^{di}\right)$. In the same way, we define $\underline{P}_1$ and $\underline{P}_2$. The

RHS of (A.21) is a function of $\zeta^{(i-1)d+1}, \ldots, \zeta^{id}$ whereas the LHS is a matrix valued function of $\zeta^{id+1}, \ldots, \zeta^{d(i+1)}$, thus, both functions are constant. Therefore, there exist a matrix $\mathcal{I}$ such that for all $i \in \mathbb{N}^*$

$$P_2(i+1) = p_2^1(id+1)\cdots p_2^d((id+1))\mathcal{I}. \tag{A.22}$$

Therefore

$$\Delta_i = \mathcal{I}^{-1}\delta_{id}\cdots\delta_{(i-1)d+1}\mathcal{I}.$$

Note that (A.22) can be rephrased in the following way. $\underline{P_2}$ and $\underline{p_2^1}\cdots\underline{p_2^d}$ are equal up to multiplication by a constant to $\mathcal{I}$. Therefore, to prove that $\mathcal{I} = \mathrm{Id}_{2d}$ for the choice for $\underline{p_1^s}$, $\underline{p_2^s}$ and $\underline{\delta^s}$ that we have made in Lemma A.17, it is enough to prove that for one given value of $\zeta^1, \ldots, \zeta^d$,

$$\underline{P_2}\left(\zeta^1, \ldots, \zeta^d\right) = \underline{p_2^1}(\zeta^d)\cdots\underline{p_2^d}(\zeta^1).$$

We will prove it for $\zeta^1 = \cdots = \zeta^d = (1, 0, \ldots, 0, 1)$. Indeed,

$$\underline{P_2}((1,0,\ldots,0,1),\ldots,(1,0,\ldots,0,1)) = \begin{pmatrix} 0_d & \mathrm{Id}_d \\ \mathrm{Id}_d & -\mathrm{Id}_d \end{pmatrix}.$$

For $1 \le s \le d$,

$$\underline{p_2^s}((1,0,\ldots,0,1)) = \left(\begin{array}{c|ccc} 0_{2d-1,1} & \multicolumn{3}{c}{\mathrm{Id}_{2d-1}} \\ \hline 1 & 0_{1,d-1} & -1 & 0_{1,d-1} \end{array}\right),$$

Hence, by induction on $1 \le t \le d$, $\underline{p_2^1}((1,0,\ldots,0,1))\cdots\underline{p_2^t}((1,0,\ldots,0,1)) =$

$$\left(\begin{array}{c|ccc} 0_{2d-t,t} & \multicolumn{3}{c}{\mathrm{Id}_{2d-t}} \\ \hline \mathrm{Id}_t & 0_{t,d-t} & -\mathrm{Id}_t & 0_{t,d-t} \end{array}\right).$$

Therefore, $\mathcal{I} = \mathrm{Id}_{2d}$. ■

3) Proof of Corollary V.2: Let us compute $\mathbb{E}\log\|\mathfrak{n}_{d+1}\|$. To that extent, we define $(e_1,\dots,e_{2d})$ the canonical basis of $\mathbb{C}^{2d}$ and we take $(e_{i_1}\wedge\cdots\wedge e_{i_d}|1\leq i_1<\cdots<i_d\leq 2d)$ as a basis of $\mathbb{C}^{\binom{2d}{d}}$. For given $1\leq i_1<\cdots<i_d\leq 2d$ and $1\leq j_1<\cdots<j_d\leq 2d$ the coefficient of $\mathfrak{n}_{d+1}(e_{i_1}\wedge\cdots\wedge e_{i_d})$ in $e_{j_1}\wedge\cdots\wedge e_{j_d}$ (we denote by a its absolute value) is the determinant of the $d\times d$ sub-matrix of μ_{d+1} obtained by taking the lines $1\leq j_1<\cdots<j_d\leq 2d$ and the columns $1\leq i_1<\cdots<i_d\leq 2d$; we denote the latter sub-matrix by $\mathcal{D}$. Denote by $\widetilde{\zeta}_{i,l}$ the coefficient at position (i,l) of G_{dn}.

- If $1\leq j_1<\cdots<j_d\leq 2d-1$,
  - if for all $1\leq s\leq d$, $i_s=j_s+1$, then $a=1$;
  - otherwise, there exists a line of zeros in $\mathcal{D}$, therefore, $a=0$.
- If $1\leq j_1<\cdots<j_{d-1}\leq 2d-1$, and $j_d=2d$,
  - if there exists $1\leq s_0\leq d-1$ such that for all $1\leq s<s_0$, $i_s=j_s+1$, for all $s_0<s\leq d$, $i_s=j_{s-1}+1$ and $j_{s_0}=l\notin\{j_1,\dots,j_{d-1}\}$, then $a=\left|\widetilde{\zeta}_{d+1,l}/\widetilde{\zeta}_{d+1,2d+1}\right|$;
  - otherwise, there exists a line of zeros in $\mathcal{D}$, therefore, $a=0$.

We now count how many times each value appears as the absolute value of a coefficient of $\mathfrak{n}_{d+1}$.

- To pick 1, one needs to pick d lines among the first $2d-1$ lines of μ_{d+1} and then, one has no choice for the columns: $\binom{2d-1}{d}$ choices.
- To pick $\left|\widetilde{\zeta}_{d+1,1}/\widetilde{\zeta}_{d+1,2d+1}\right|$, one needs to pick $d-1$ lines among the first $2d-1$ lines of μ_{d+1} and then, one has no choice for the remaining line and the columns: $\binom{2d-1}{d-1}=\binom{2d-1}{d}$ choices.
- To pick $\left|\widetilde{\zeta}_{d+1,l}/\widetilde{\zeta}_{d+1,2d+1}\right|$ for a given $2\leq l\leq 2d$, one needs to pick $d-1$ lines among the first $2d-1$ lines of μ_{d+1} and one cannot pick the $(k-1)$-th line. Then one has no choice for the remaining line and the columns: $\binom{2d-2}{d-1}$ choices.

We factorize the term $1/\left|\widetilde{\zeta}_{d+1,2d+1}\right|$, whose log-expectation cancels out with $\mathbb{E}_{\pi_0,\pi_d}\log\left|\zeta_0\zeta_d^\dagger\right|$ and get the claimed bound.

4) Proof of Proposition VI.2: According to Proposition IV.4,

$$\mathcal{L}_\infty=\frac{-1}{\log 2}\Big[\log\rho+\mathbb{E}\log\left|\zeta_0\zeta_d^\dagger\right|+\max\left(\mathbb{E}\log\left|\det\psi_1^1\right|\;;\;\mathbb{E}\log\left|\det\psi_1^2\right|\;;\;\gamma(\psi^1)+\gamma(\psi^2)\right)\Big],$$

where

$$\psi_i^1 = \begin{pmatrix} -\frac{\beta b_i}{\alpha a_i} & 1 \\ -\frac{c_i}{\alpha a_i} & 0 \end{pmatrix} \text{ and } \psi_i^2 = \begin{pmatrix} -\frac{\beta b_i^\dagger}{c_i^\dagger} & -\frac{\alpha a_i^\dagger}{c_i^\dagger} \\ 1 & 0 \end{pmatrix}.$$

Therefore,

$$\mathcal{L}_\infty = \frac{-1}{\log 2} \max \big(2\mathbb{E} \log |a_1| \; ; \; 2\mathbb{E} \log |a_1| + 2 \log \alpha \, ; \\ 2\mathbb{E} \log |a_1| + \log \alpha + \gamma(\psi^1) + \gamma(\psi^2) \big).$$

Since $\alpha \leq 1$, $\log \alpha \leq 0$, therefore

$$\mathcal{L}_\infty = \frac{-1}{\log 2} \max \left(2\mathbb{E} \log |a_1| \; ; \; 2\mathbb{E} \log |a_1| + \gamma \left(\widetilde{\psi}^1 \right) + \gamma(\psi^2) \right), \tag{A.23}$$

where

$$\widetilde{\psi}_i^1 = \alpha \psi_i^1 = \begin{pmatrix} -\frac{\beta b_i}{a_i} & \alpha \\ -\frac{c_i}{a_i} & 0 \end{pmatrix}.$$

In order to finish the proof, we will construct of family of functions $(f_p)_{p \in \mathbb{N}^*}$ from $(0,1] \times [0,1]$ to $\mathbb{R}$ such that for all $p \in \mathbb{N}^*$, $f_p(\alpha, \beta)$ is non-decreasing in α and in β and such that for all $p \in \mathbb{N}^*$ and for all $(\alpha, \beta) \in (0,1] \times [0,1]$,

$$\gamma \left(\widetilde{\psi}^1 \right) + \gamma(\psi^2) \leq f_p(\alpha, \beta).$$

We define $\mathcal{D}$ in the following way:

$$\mathcal{D} \triangleq \bigcup_{p \in \mathbb{N}^*} \{ (\alpha, \beta) \in (0,1] \times [0,1] \; ; \; f_p(\alpha, \beta) \leq 0 \}.$$

Since for all $p \in \mathbb{N}^*$, $f_p(\alpha, \beta)$ is non-decreasing in α and in β, we get that for all $(x, y) \in \mathcal{D}$, $(0, x) \times [0, y) \subset \mathcal{D}$. Moreover, by (A.23), if $(\alpha, \beta) \in \mathcal{D}$, then (VI.3) is verified.

Fix $p \in \mathbb{N}^*$. First note that by (A.3),

$$\gamma \left(\widetilde{\psi}^1 \right) + \gamma(\psi^2) \leq \frac{1}{p} \left(\mathbb{E} \log \left\| \widetilde{\psi}_p^1 \cdots \widetilde{\psi}_1^1 \right\| + \mathbb{E} \log \left\| \psi_p^2 \cdots \psi_1^2 \right\| \right).$$

Recall that we use the Fröbenius norm on matrices. Denote $\phi^1(\alpha, \beta) = \widetilde{\psi}_p^1 \cdots \widetilde{\psi}_1^1$ and $\phi^2(\alpha, \beta) = \psi_p^2 \cdots \psi_1^2$. Note that the coefficients of $\phi^1(\alpha, \beta)$ and $\phi^2(\alpha, \beta)$ are polynomials in α and β. The function $1/p \left(\mathbb{E} \log \|\phi^1(\alpha, \beta)\| + \mathbb{E} \log \|\phi^2(\alpha, \beta)\| \right)$ would be a good candidate for f_p but it is not non-decreasing in α and β, therefore, we have to modify it slightly.

Consider P a polynomial in α and β,

$$P(\alpha,\beta)=\sum_{i,j=1}^{n}\theta_{i,j}\alpha^i\beta^j.$$

Define the polynomial $|P|$ in the following way

$$|P|(\alpha,\beta)=\sum_{i,j=1}^{n}|\theta_{i,j}|\,\alpha^i\beta^j.$$

By the triangle inequality, for all $(\alpha,\beta)\in(0,1]\times[0,1]$, $|P(\alpha,\beta)|\leq|P|(\alpha,\beta)$. Moreover, $|P|(\alpha,\beta)$ is non decreasing for $(\alpha,\beta)\in(0,1]\times[0,1]$.

Define the matrices $|\phi^1|(\alpha,\beta)$ and $|\phi^2|(\alpha,\beta)$ in the following way.
For $i,j,k=1,2$, set $\left|\phi^k\right|_{i,j}=\left|\phi^k_{i,j}\right|$. Then,

$$\left\|\phi^1(\alpha,\beta)\right\|\leq\left\|\left|\phi^1\right|(\alpha,\beta)\right\|\ \text{ and }\ \left\|\phi^2(\alpha,\beta)\right\|\leq\left\|\left|\phi^2\right|(\alpha,\beta)\right\|.$$

Moreover $\|\,|\phi^1|(\alpha,\beta)\|$ and $\|\,|\phi^2|(\alpha,\beta)\|$ are non decreasing for $(\alpha,\beta)\in(0,1]\times[0,1]$. Thus, we conclude the proof by defining

$$f_p=\frac{1}{p}\left(\mathbb{E}\log\left\|\left|\phi^1\right|(\alpha,\beta)\right\|+\mathbb{E}\log\left\|\left|\phi^2\right|(\alpha,\beta)\right\|\right).$$

Remark A.24. Note that if we define

$$\left|\widetilde{\psi}^1_i\right|=\begin{pmatrix}\frac{\beta|b_i|}{|a_i|} & \alpha\\ \frac{|c_i|}{|a_i|} & 0\end{pmatrix}\ \text{ and }\ \left|\psi^2_i\right|=\begin{pmatrix}\frac{\beta|b_i|}{|c_i|} & \frac{\alpha|a_i|}{|c_i|}\\ 1 & 0\end{pmatrix},$$

then

$$f_p=1/p\left(\mathbb{E}\log\left\|\left|\widetilde{\psi}^1_p\right|\cdots\left|\widetilde{\psi}^1_1\right|\right\|+\mathbb{E}\log\left\|\left|\psi^2_p\right|\cdots\left|\psi^2_1\right|\right\|\right).$$

We use that fact in the numerical computation of the functions f_p.

C. Exterior product

In this section we give the material on exterior products. We provide only the properties relevant to the article, see [22, Chapter XVI.6-7] and [13, Chapter A.III.5] for more details.

Proposition A.25. *For $0\leq k\leq p$, the exterior product of k vectors in $\mathbb{C}^p$, $v_1,\ldots,v_k$ is denoted by $v_1\wedge\cdots\wedge v_k$. Is is a vector of the exterior product of degree k of $\mathbb{C}^p$ that we denote by $\bigwedge^k\mathbb{C}^p$. $\bigwedge^k\mathbb{C}^p$ is a $\mathbb{C}$-vector space of dimension $\binom{k}{p}$.*

The exterior product $v_1, \ldots, v_k$ is a multi-linear (i.e. linear in every v_i, $1 \leq i \leq k$) and anti-symmetric (i.e. $v_{\sigma(1)} \wedge \cdots v_{\sigma(k)} = \varepsilon(\sigma)$ for σ permutation of $\{1, \ldots, k\}$ and $\varepsilon(\sigma)$ its signature) function.

If $e_1, \ldots, e_p$ is a basis of $\mathbb{C}^p$, then $(e_{i_1} \wedge \cdots \wedge e_{i_k} | 1 \leq i_1 < \cdots < i_k \leq p)$ is a basis of $\bigwedge^k \mathbb{C}^p$. The later is called the canonical basis *of $\bigwedge^k \mathbb{C}^p$ if $e_1, \ldots, e_p$ is the canonical basis of $\mathbb{C}^p$.*

If M is a matrix of size $p \times q$, the exterior product of M that we denote by $\bigwedge^k M$ is a map from $\bigwedge^k \mathbb{C}^q$ to $\bigwedge^k \mathbb{C}^p$ such that

$$\bigwedge^k M \left(v_1 \wedge \cdots \wedge v_k\right) = Mv_1 \wedge \cdots \wedge Mv_k.$$

Finally, for two matrices M and N, $\bigwedge^k (MN) = \bigwedge^k (M) \bigwedge^k (N)$.

Proposition A.26. *If X is a square matrix of size p, then*

$$\bigwedge^p X = \det X.$$

Moreover

$$\det \bigwedge^p X = (\det X)^p$$

D. Capacity of the non-fading channels

In this Section, we give expressions of the limiting sum-rate per-cell capacity for the Soft-Handoff model and the Wyner model (both symmetric and asymmetric) for the non-fading channels. Those expressions are consequences of results on Toeplitz matrices [23]. See [3] for an example of derivation.

1) The Soft-Handoff model: We assume that $d = 1$, and for $i \in \mathbb{N}^*$, $\zeta_{i,i+1} = \alpha \in [0, 1]$ and $\zeta_{i,i} = 1$. Then, the limiting per-cell sum-rate capacity is

$$Cap(\rho) = \log \left(\frac{1 + K\rho\left(1 + \alpha^2\right) + \sqrt{1 + 2K\rho\left(1 + \alpha^2\right) + K^2\rho^2\left(1 - \alpha^2\right)^2}}{2} \right).$$

2) The Wyner model:

The symmetric setting: We assume that $d = 2$, and for $i \in \mathbb{N}^*$, $\zeta_{i,i+2} = \zeta_{i,i} = \alpha \in [0, 1]$ and $\zeta_{i,i+1} = 1$. Then, the limiting per-cell sum-rate capacity is

$$Cap(\rho) = \int_0^1 \log\left(1 + K\rho\left(1 + 2\alpha\cos(2\pi f)\right)^2\right) df.$$

The asymmetric setting: We assume that $d = 2$, and for $i \in \mathbb{N}^*$, $\zeta_{i,i+2} = \zeta_{i,i+1} = \alpha \in [0,1]$ and $\zeta_{i,i} = 1$. Then, the limiting per-cell sum-rate capacity is

$$Cap(\rho) = \int_0^1 \log\left(1 + K\rho\left(1 + 2\alpha^2 + 2\alpha(1+\alpha)\cos(2\pi f) + 2\alpha\cos(4\pi f)\right)\right) df.$$

References

[1] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, U. Spagnolini, and S. Shamai (Shitz), "An information theoretic view of distributed antenna processing in cellular systems," in *Distributed Antenna Systems: Open Architecture for Future Wireless Communications* (H. Hu, Y. Zhang, and J. Luo, eds.), Auerbach Publications, CRC Press, May 2007.

[2] S. Shamai (Shitz), O. Somekh, and B. M. Zaidel, "Multi-cell communications: An information theoretic perspective," in *Proceedings of the Joint Workshop on Communications and Coding (JWCC'04)*, (Donnini, Florence, Italy), Oct.14–17, 2004.

[3] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Transactions on Information Theory*, vol. 40, pp. 1713–1727, Nov. 1994.

[4] S. V. Hanly and P. A. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommun. Syst.*, vol. 1, pp. 1–42, 1993.

[5] N., O. Somekh, S. Shamai (Shitz), and O. Zeitouni, "On certain large jacobi matrices with applications to wireless communications," *Submitted to the IEEE Transactions on Information Theory*, 2007.

[6] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–598, Nov. 1999.

[7] A. Lozano, A. Tulino, and S. Verdú, "High-snr power offset in multi-antenna communications," *IEEE Transactions on Information Theory*, vol. 51, pp. 4134–4151, Dec. 2005.

[8] R. Carmona and J. Lacroix, *Spectral theory on random Schrödinger operators*. Probability and its Applications, Boston, MA: Birkhäuser Boston Inc., 1990.

[9] S. Shamai (Shitz), L. H. Ozarow, and A. D. Wyner, "Information rates for a discrete-time gaussian channel with intersymbol interference and stationary inputs," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1527–1539, Nov. 1991.

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

[11] O. Somekh and S. Shamai (Shitz), "Shannon-theoretic approach to a Gaussian cellular multi-access channel with fading," *IEEE Transactions on Information Theory*, vol. 46, pp. 1401–1425, July 2000.

[12] N. Levy, *Ph.D. Thesis, forthcoming*. Technion.

[13] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger Operators*, vol. 8 of *Progress in Probability and Statistics*. Boston, MA: Birkhäuser Boston Inc., 1985.

[14] J. E. Cohen, H. Kesten, and C. M. Newman, "Oseledec's multiplicative ergodic theorem: a proof," in *Random matrices and their applications (Brunswick, Maine, 1984)*, vol. 50 of *Contemp. Math.*, pp. 23–30, Providence, RI: Amer. Math. Soc., 1986.

[15] F. Ledrappier, "Quelques propriétés des exposants caractéristiques," in *École d'été de probabilités de Saint-Flour, XII—1982*, vol. 1097 of *Lecture Notes in Math.*, pp. 305–396, Berlin: Springer, 1984.

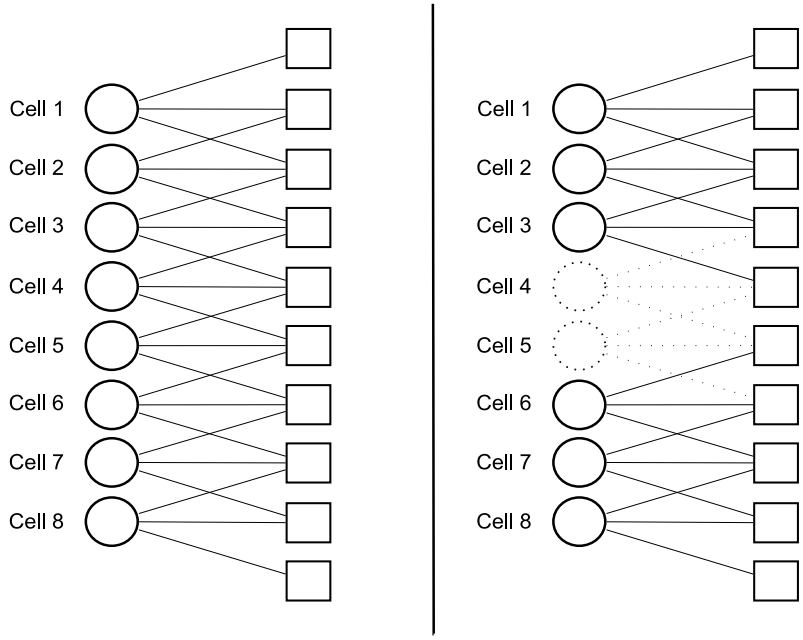

Fig. 1. Proof of Proposition V.4, lower bound.

[16] L. Pastur and A. Figotin, *Spectra of random and almost-periodic operators*, vol. 297 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Berlin: Springer-Verlag, 1992.

[17] J. C. Watkins, "Limit theorems for products of random matrices: a comparison of two points of view," in *Random matrices and their applications (Brunswick, Maine, 1984)*, vol. 50 of *Contemp. Math.*, pp. 5–22, Providence, RI: Amer. Math. Soc., 1986.

[18] H. Hennion, "Loi des grands nombres et perturbations pour des produits réductibles de matrices aléatoires indépendantes," *Z. Wahrsch. Verw. Gebiete*, vol. 67, no. 3, pp. 265–278, 1984.

[19] A. Narula, *Information Theoretic Analysis of Multiple-Antenna Transmission Diversity*. PhD thesis, Massachusetts Institute of Technology (MIT), Boston, MA, June 1997.

[20] W. Craig and B. Simon, "Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices," *Comm. Math. Phys.*, vol. 90, no. 2, pp. 207–218, 1983.

[21] W. Craig and B. Simon, "Subharmonicity of the Lyaponov index," *Duke Math. J.*, vol. 50, no. 2, pp. 551–560, 1983.

[22] S. Mac Lane and G. Birkhoff, *Algebra*. New York: The Macmillan Co., 1967.

[23] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Transactions on Information Theory*, vol. IT-18, pp. 725–730, Nov. 1972.

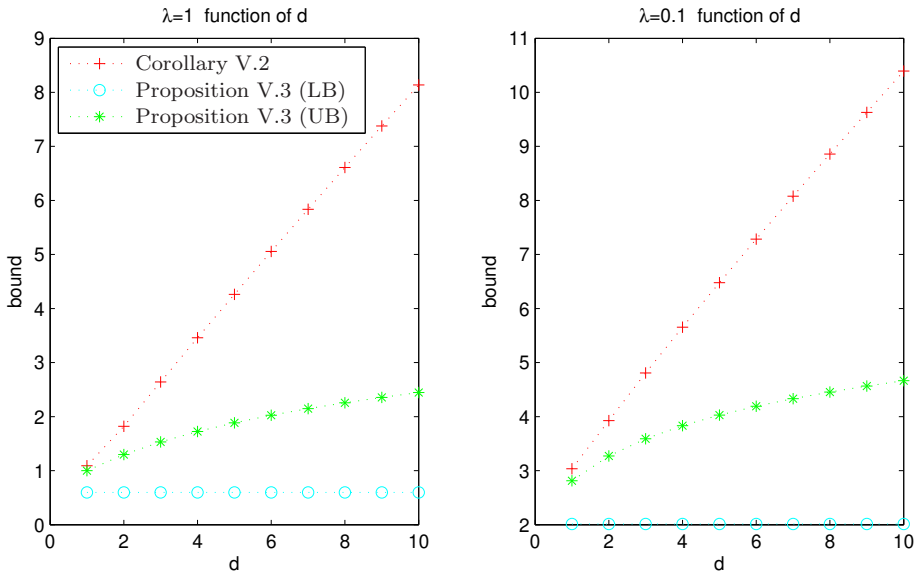

Fig. 2. Comparison of the bounds of Corollary V.2 and Proposition V.3 for $\lambda = 0.1, 1$, in function of d.

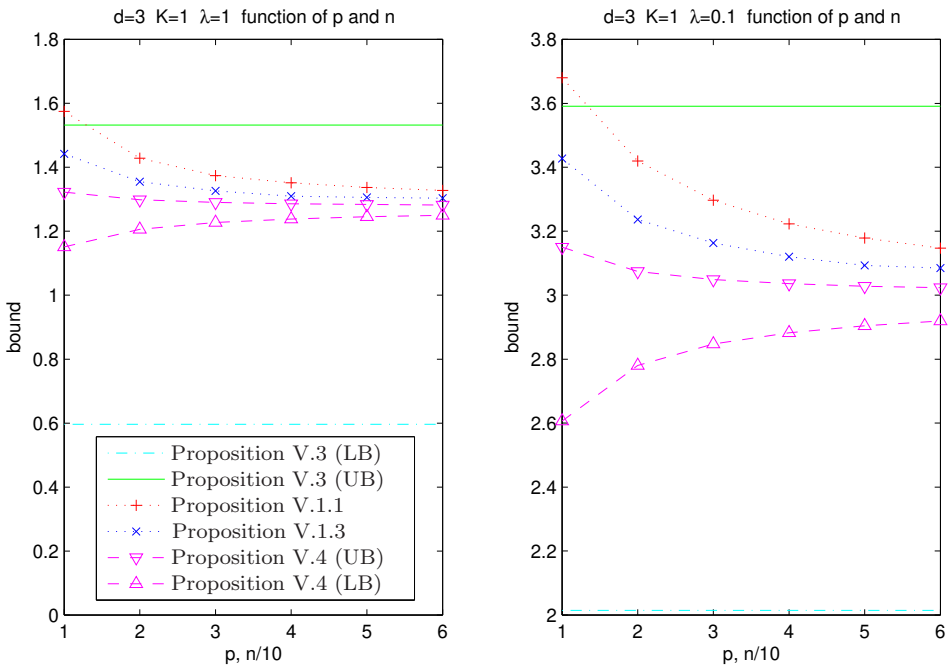

Fig. 3. Comparison of the different bounds for $\lambda = 0.1, 1$ and $d = 3$.

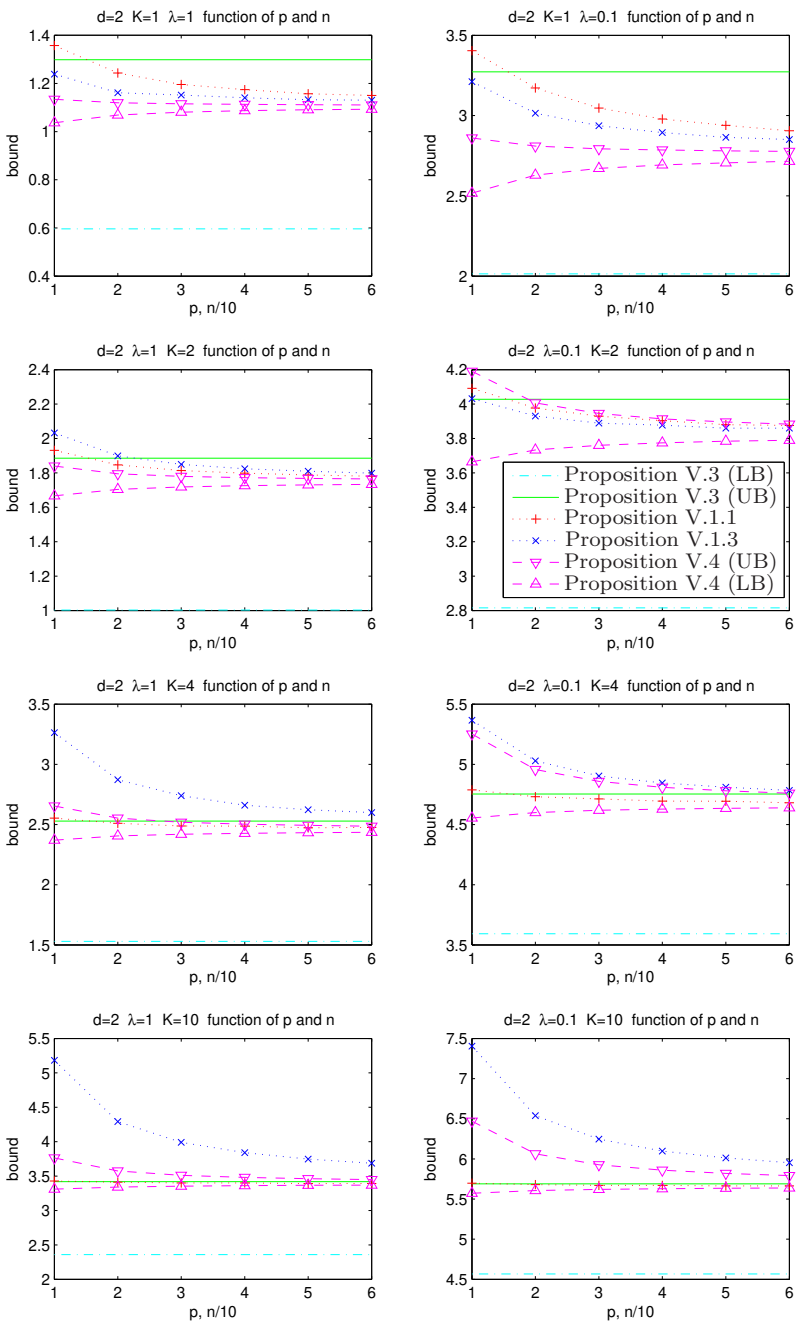

Fig. 4. Comparison of the different bounds for $\lambda = 0.1, 1$ and $K = 1, 2, 4, 10$. The values for the non random channel with $\lambda = 1$ (resp. $\lambda = 0.1$) and $K = 1, 2, 4, 10$ are respectively 1.06, 1.47, 1.95, 2.66 (resp. 2, 66, 3.25, 3.87, 4.72).

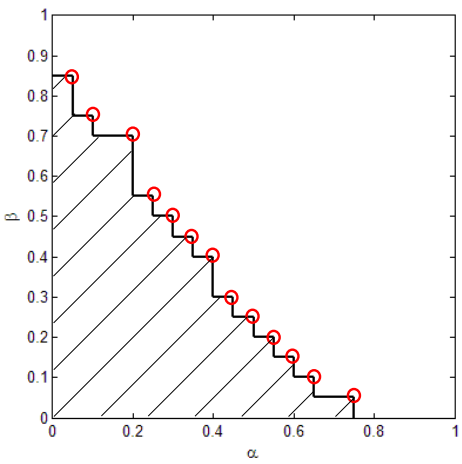

Fig. 5. Region where (VI.3) holds for Rayleigh fading.

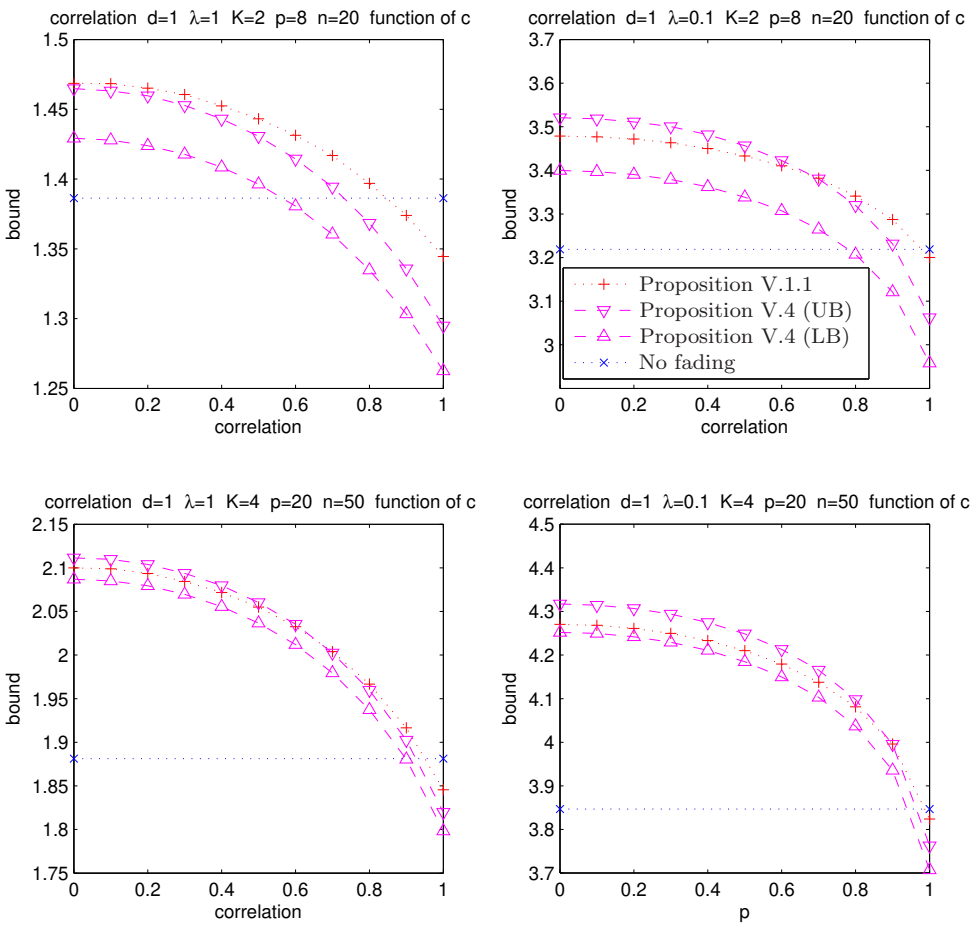

Fig. 6. Influence of the correlation on the capacity in function of the SNR and K. Bounds of Proposition V.4.

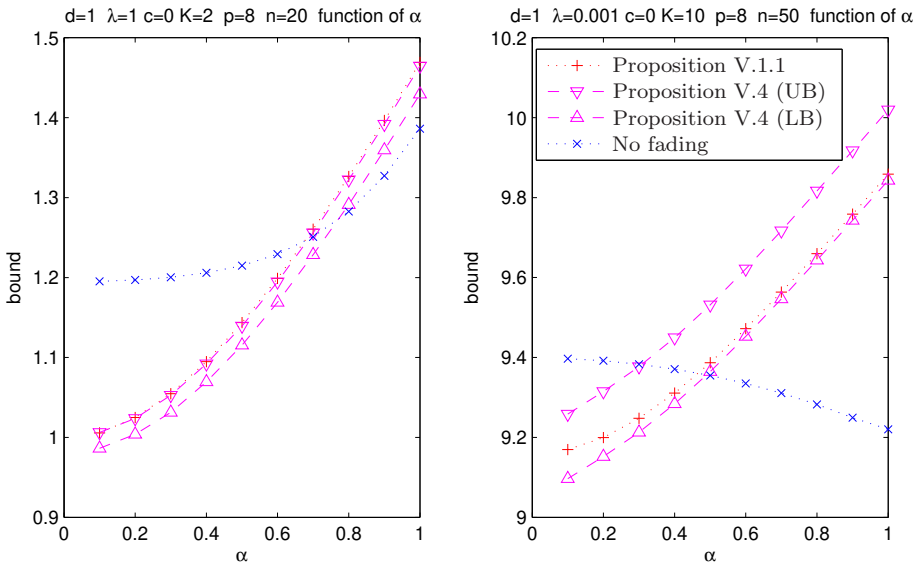

Fig. 7. Influence of the correlation on the capacity in function of α. Bounds of Proposition V.4.

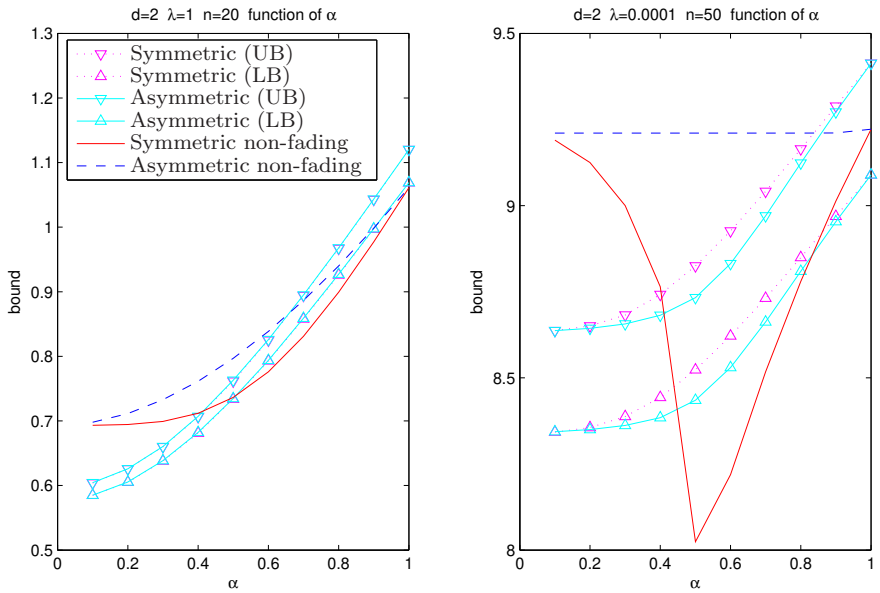

Fig. 8. Comparison of the symmetric and asymmetric channels with and without Rayleigh fading for $\lambda = 1, 10^{-4}$ in function of α. Bounds of Proposition V.4.

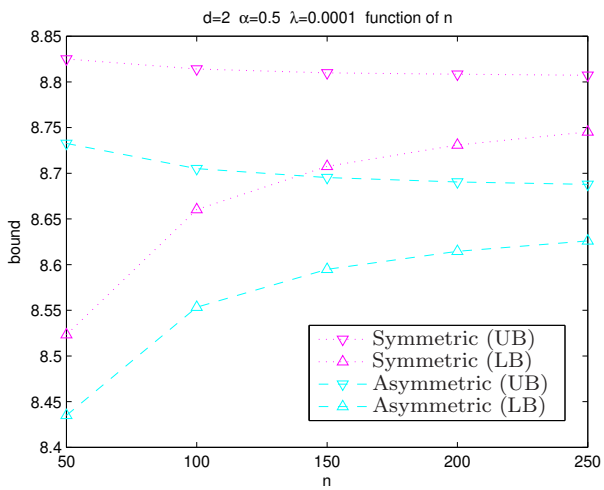

Fig. 9. Comparison of the symmetric and asymmetric channels with Rayleigh fading for $\lambda = 10^{-4}$ and $\alpha = 0.5$. Bounds of Proposition V.4.

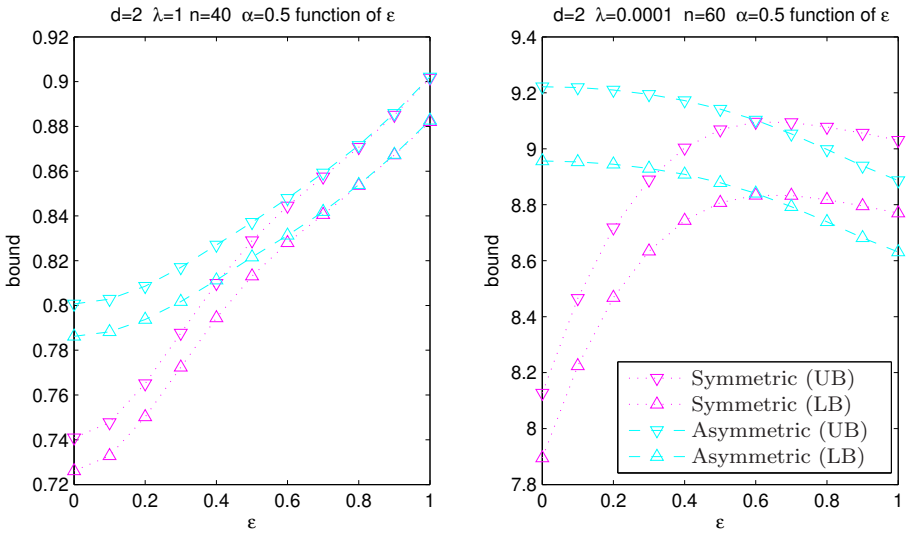

Fig. 10. Comparison of the symmetric and asymmetric channels with uniform fading for $\lambda = 1, 10^{-4}$ and $\alpha = 0.5$ in function of ε. Bounds of Proposition V.4.