

On Certain Large Random Hermitian Jacobi Matrices With Applications to Wireless Communications

Nathan Levy, *Student Member, IEEE*, Oren Somekh, *Member, IEEE*, Shlomo Shamai (Shitz), *Fellow, IEEE*, and Ofer Zeitouni, *Fellow, IEEE*

Abstract—In this paper we study the spectrum of certain large random Hermitian Jacobi matrices. These matrices are known to describe certain communication setups. In particular, we are interested in an uplink cellular channel which models mobile users experiencing a soft-handoff situation under joint multicell decoding. Considering rather general fading statistics we provide a closed-form expression for the per-cell sum-rate of this channel in high signal-to-noise ratio (SNR), when an intra-cell time-division multiple-access (TDMA) protocol is employed. Since the matrices of interest are *tridiagonal*, their eigenvectors can be considered as sequences with second-order linear recurrence. Therefore, the problem is reduced to the study of the exponential growth of products of two-by-two matrices. For the case where K users are simultaneously active in each cell, we obtain a series of lower and upper bound on the high-SNR power offset of the per-cell sum-rate, which are considerably tighter than previously known bounds.

Index Terms—Distributed antenna array, fading channels, high-signal-to-noise-ratio (SNR) characterization, multiuser detection, random matrices, sum-rate capacity, Wyner cellular uplink.

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. Cellular systems are of major interest as the most common method for providing continuous services to mobile users, in both indoor and outdoor environments. 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 multiple-cell processing (MCP), which allows the base stations (BSs) 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 multiple-cell processing).

Most of the works on the uplink channel of cellular systems deal with a single-cell setup. References that consider multiple-cell scenarios tend to adopt complex multiple-cell system models which render analytical treatment extremely hard (if not, impossible). Indeed, most of the results reported in these works are derived via intensive numerical calculations which provide little insight into the behavior of the system performance as a function of various key parameters (e.g., [3]–[8]).

Motivated by the fact that mobiles users in a cellular system “see” only a small number of BSs, and by the desire to provide analytical results, an attractive analytically tractable model for a multiple-cell system was suggested by Wyner in [9] (see also [4] for 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 nonfading channels and a “wideband” (WB) transmission scheme, where all bandwidth is available for coding (as opposed to *random* spreading), the throughput obtained with optimum and linear minimum mean-square error (MMSE) *joint* processing of the received signals from *all* cell sites are derived. Since it was first presented in [9], “Wyner-like” models have provided a framework for many works analyzing various transmission schemes in both the uplink and downlink channels (see [2] and references therein).

In this work we consider a simple “Wyner-like” cellular setup presented in [10] (see also [11]). According to this setup, the cells are arranged on a circle (or a line), and the mobile users “see” only the two BSs which are located on their cell’s boundaries. All the BSs are assumed to be connected through an ideal backhaul network to a central MCP, that can *jointly* process the uplink received signals of all cell sites, as well as preprocess the signals to be transmitted by all cell sites in the downlink channel. The users are hence in what is referred to as a “soft-handoff” situation, which is very common in practical real-life cellular systems, and is therefore of real practical as well as theoretical interest (see, for example, [12] for a recent survey on handoff schemes). With simplicity and analytical tractability in mind, and in a similar manner to previous work, the model provides perhaps the simplest framework for a soft-handoff setting in a

Manuscript received October 11, 2007; revised November 25, 2008. Current version published March 18, 2009. The material in this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT), Toronto, ON, July 2008. This work was supported by the fund for promotion of research at the Technion, a Marie Curie Outgoing International Fellowship and the NEWCOM++ network of excellence both within the 6th and 7th European Community Framework Programmes, the REMON Consortium, the U.S. National Science Foundation under Grants DMS-0503775 and DMS-0804133, and a grant from the Israel Science Foundation. This work was done while N. Levy was with the Département de Mathématiques et Applications, Ecole Normale Supérieure, Paris 75005, France. Part of this work was done while O. Zeitouni was with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa.

N. Levy is with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa 32000, Israel (e-mail: nlevy@tx.technion.ac.il).

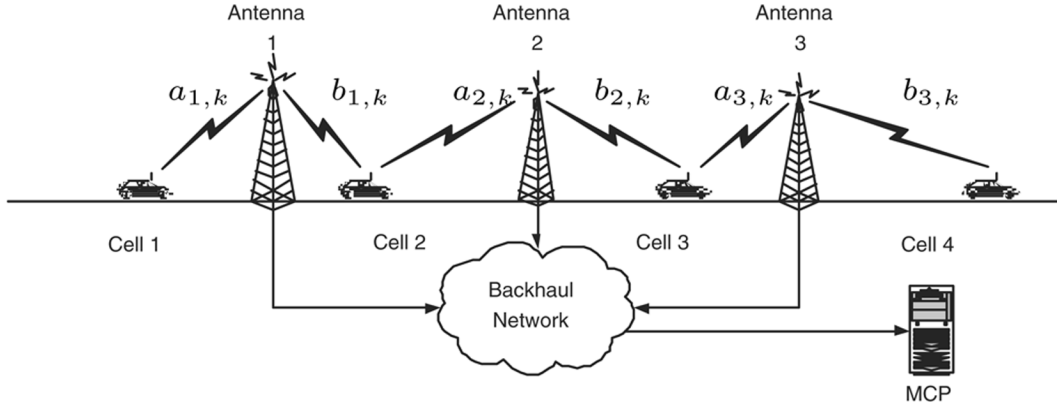
O. Somekh is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: orens@princeton.edu).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

O. Zeitouni is with the School of Mathematics, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: zeitouni@math.umn.edu).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2009.2013046

Fig. 1. Soft-handoff setup ($M = 3$).

cellular system, that still represents real-life phenomena such as intercell interference and fading.

Unfortunately, the analysis of “Wyner-like” models in general and the “soft-handoff” setup in particular presents some analytical difficulties (see Section II-B) when fading is present. These difficulties render conventional analysis methods such as large random matrix theory impractical. Indeed, the per-cell sum-rate rates supported by MCP in the uplink channel of the “soft-handoff” setups are known only for limited scenarios such as nonfading channels, phase-fading channels, fading channels but with large number of users per cell, and Rayleigh-fading channels with single user active per cell [10], [11], [13]. The latter result is due to a remarkable early work by Narula [14] dealing with the capacity of a two-tap time variant intersymbol interference (ISI) channel. Calculating the per-cell sum-rate capacity supported by the uplink channel of the “soft-handoff” setup in the presence of *general* fading channels (not necessarily Rayleigh-fading channels), when *finite* number of users are active simultaneously in each cell remains an open problem (see [11], [15] for bounds on this rate). As will be shown in the sequel, this problem is closely related to calculating the spectrum of certain large random Hermitian Jacobi matrices. The high-signal-to-noise-ratio (SNR) characterization of the sum-rate capacity, previously unknown, is the main focus of this work.

In particular, we calculate the high-SNR slope and power offset of the rate with a single user active per-cell (intra-cell time-division multiple access (TDMA)) under a rather generic fading distribution. We also prove the following results for any given number of active users per cell. We prove the existence of a limiting sum-rate capacity when the number of cells goes to infinity and calculate the high-SNR slope in Theorem 2. Moreover, we give bounds on the high-SNR power offset in Proposition 3. In particular, we give a sequence of explicit upper and lower bounds; the gap between the lower and the upper bounds is decreasing with the bounds’ order and complexity.

The rest of the paper is organized as follows. In Section II, we present the problem statement and main results. Section III includes a comprehensive review of previous works. Several applications of the main result are discussed in Section IV. Concluding remarks are included in Section V. Various derivations and proofs are deferred to the Appendices.

II. PROBLEM STATEMENT AND MAIN RESULTS

A. System Model

In this paper, we consider a linear version of the cellular “soft-handoff” setup introduced in [10], [11], according to which $M+1$ cells with K single-antenna users per cell are arranged on a line, where the M single antenna BSs are located on the boundaries of the cells (see Fig. 1 for the special case of $M = 3$). 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 BSs is given for an arbitrary time index by

$$\mathbf{y} = \mathbf{H}_M \mathbf{x} + \mathbf{n}. \quad (1)$$

The $M \times K(M+1)$ channel transfer matrix \mathbf{H}_M is a two block diagonal matrix defined by

$$\mathbf{H}_M = \begin{pmatrix} \mathbf{a}_1 & \mathbf{b}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{a}_M & \mathbf{b}_M \end{pmatrix} \quad (2)$$

where \mathbf{a}_m and \mathbf{b}_m are $1 \times K$ row vectors denoting the channel complex fading coefficients, experienced by the K users of the m th and $(m+1)$ th cells, respectively, when received by the m th BS antenna. \mathbf{n} represents the $M \times M$ zero mean circularly symmetric Gaussian noise vector $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$.

We assume throughout that the fading processes are independent and identically distributed (i.i.d.) among different users and BSs, and can be viewed for each user as ergodic processes with respect to the time index. We denote by π_a (resp., π_b) the distribution of $a_{m,k}$ (resp., $b_{m,k}$). Moreover, we denote by \mathbb{P} the probability associated with those random sequences and by \mathbb{E} the associated expectation. We will be working throughout with a subset of the following assumptions. We denote by \mathbb{R}^+ the interval $[0, \infty)$.

- (H1) $\mathbb{E}_{\pi_a}(\log |x|)^2 < \infty^1$ and $\mathbb{E}_{\pi_b}(\log |x|)^2 < \infty$.
- (H2) π_a and π_b are absolutely continuous with respect to Lebesgue measure on \mathbb{C} .
- (H3) There exists a real \mathcal{M} such that if x is distributed according to π_a (resp., π_b) then the density of $|x|^2$ is strictly positive on the interval $[\mathcal{M}; \infty)$.
- (H3') There exist $m_a < \mathcal{M}_a \in \mathbb{R}^+ \cup \{\infty\}$ (resp., $m_b < \mathcal{M}_b \in \mathbb{R}^+ \cup \{\infty\}$) such that if x is distributed according to π_a (resp., π_b) then the density of $|x|^2$ and the Lebesgue measure on $[m_a; \mathcal{M}_a]$ (resp., $[m_b; \mathcal{M}_b]$) are mutually absolutely continuous.
- (H4) There exists a ball in \mathbb{C} such that the Lebesgue measure outside that ball is absolutely continuous with respect to π_a and π_b .

Let us explain the rational behind those assumptions. (H1) is a weak condition of integrability on π_a and π_b . Note that the first moment of the logarithm of the fading coefficients will play a role in Theorem 1, which makes reasonable an assumption on the moments of the logarithm of the fading coefficients.

(H2), (H3), and (H3') are conditions on π_a and π_b . In the Proof of Theorem 1, we will use the theory of Harris Markov chains (see [16] for the relevant background). (H2) and [(H3) or (H3')] will be necessary to prove the property of Harris for Markov chains which will appear in the proofs.

(H4) is a stronger condition than (H3) on π_a and π_b . It will be used in the Proof of Theorem 2 to prove the property of Harris for relevant Markov chains. Together with (H2), it means that the density of π_a and π_b is positive outside a large box (except on a zero Lebesgue-measure set).

We further assume that the channel state information (CSI) is available to the MCP only, while the transmitters know only the channel statistics, and cannot cooperate their transmissions in any way. Therefore, independent zero mean circularly symmetric Gaussian codebooks conform with the capacity achieving statistics, where \mathbf{x} denotes the $(M+1)K \times 1$ transmit vector $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \rho \mathbf{I}_{MK})$, and ρ is the average transmit power of each user² (ρ is thus equal to the transmit SNR of the users).

With the above assumptions, the system (1) is a multiple-access channel (MAC). We are interested in the per-cell sum-rate capacity

$$C_M(P) = \frac{1}{M} \mathbb{E}(\log \det \mathbf{G}_M) \quad [\text{nats/channel use}] \quad (3)$$

where $P \triangleq K\rho$ is the per-cell transmitted average power, and

$$\mathbf{G}_M \triangleq \mathbf{I}_M + \rho \mathbf{H}_M \mathbf{H}_M^\dagger \quad (4)$$

and the expectation is taken over the channel transfer matrix entries. (Here and in the sequel, for a scalar $z \in \mathbb{C}$, z^\dagger denotes

¹A natural base logarithm is used throughout this work unless explicitly denoted otherwise.

²Note that since the channel transfer matrix \mathbf{H}_M is a *column-regular* gain matrix (see definition in [17]) when $M \rightarrow \infty$, the capacity achieving statistics remains the same in this case, even if we allow the users to cooperate as long as they are unaware of the CSI.

the complex conjugate, while for a matrix \mathbf{A} , \mathbf{A}^\dagger denotes the matrix with $\mathbf{A}^\dagger(i, j) = \mathbf{A}(j, i)^\dagger$.) The nonzero entries of the *Hermitian Jacobi* matrix \mathbf{G}_M are equal to

$$\begin{aligned} [\mathbf{G}_M]_{m,m-1} &= \rho \langle \mathbf{b}_{m-1}; \mathbf{a}_m \rangle \\ [\mathbf{G}_M]_{m,m} &= 1 + \rho(|\mathbf{a}_m|^2 + |\mathbf{b}_m|^2) \\ [\mathbf{G}_M]_{m,m+1} &= \rho \langle \mathbf{a}_{m+1}; \mathbf{b}_m \rangle \end{aligned} \quad (5)$$

where out-of-range indices should be ignored, and for any two arbitrary L -length vectors \mathbf{a}, \mathbf{b} we define $\langle \mathbf{a}; \mathbf{b} \rangle \triangleq \sum_{l=1}^L a_l^\dagger b_l$, and $|\mathbf{a}|^2 \triangleq \langle \mathbf{a}; \mathbf{a} \rangle$.

Since we shall focus on the asymptotes of infinite number of cells $M \rightarrow \infty$, boundary effects can be neglected and symmetry implies that the rate (3) equals the maximum equal rate (or symmetric capacity) supported by the channel [18].

The above description relates to the WB protocol where all users transmit simultaneously. According to the intra-cell TDMA protocol only one user is simultaneously active per cell, transmitting $1/K$ of the time using the *total* cell transmit power P . In this case, it is easily verified that with no loss of generality, we can consider a single user per cell in terms of the per-cell sum-rate, setting $K = 1$ in (1) and (2).

B. Analysis Difficulty

Many recent studies have analyzed the rates of various channels using results from (large) random matrix theory (see [19] for a recent review). In those cases, the number of random variables involved is of the order of the number of elements in the matrix \mathbf{G}_M (or \mathbf{H}_M), and self-averaging is strong enough to ensure convergence of the empirical measure of eigenvalues, and to derive equations for the limit (or its Stieltjes transform). In particular, this is the case if the normalized continuous power profile of \mathbf{H}_M , which is defined as

$$\mathcal{P}_M(r, t) \triangleq \mathbb{E}(|[\mathbf{H}_M]_{i,j}|^2) \quad \frac{i}{M} \leq r < \frac{i+1}{M}, \frac{j}{(M+1)K} \leq t < \frac{j+1}{(M+1)K} \quad (6)$$

converges uniformly to a bounded, piecewise continuous function as $M \rightarrow \infty$, see, e.g., [19, Theorem 2.50] and [20] for fluctuation results. In the case under consideration here, it is easy to verify that for K fixed, $\mathcal{P}_M(r, t)$ does *not* converge uniformly, and other techniques are required.

C. Extreme SNR Regime Characterization

As mentioned earlier, the per-cell sum-rate capacity of the “soft-handoff” setup is known only for certain limited cases to be elaborated in the next section, and in general analytical results are hard to derive. As an alternative to deriving exact analytical results, we focus here on extracting parameters which characterize the channel rate under extreme SNR scenarios. The reader is referred to [21]–[23] for an elaboration on the extreme SNR characterization.

a) The Low-SNR Regime: This regime is usually the operating regime for WB systems [22].

The average per-cell spectral efficiency in bits per second per hertz (bits/s/Hz), expressed as a function of the system average

transmit SNR, E_b/N_0 , is evaluated by solving the implicit equation obtained by substituting

$$P = C_M \left(\frac{E_b}{N_0} \right) \frac{E_b}{N_0} \quad (7)$$

in (3), where $C_M(E_b/N_0) = C_M(P)/\log 2$ stands for the uplink spectral efficiency measured in [bits/s/Hz]. The low-SNR regime is characterized through the minimum transmit E_b/N_0 that enables reliable communications

$$\frac{E_b}{N_{0\min}} \triangleq \frac{\log 2}{\dot{C}_M(0)} \quad (8)$$

and the low-SNR spectral efficiency slope

$$S_0 \triangleq \frac{2[\dot{C}_M(0)]^2}{-\ddot{C}_M(0)} \quad (9)$$

yielding the following low-SNR affine approximation:

$$C_M \left(\frac{E_b}{N_0} \right) \approx \frac{S_0}{3|_{\text{dB}}} \left(\frac{E_b}{N_0} \Big|_{\text{dB}} - \frac{E_b}{N_{0\min}} \Big|_{\text{dB}} \right) \text{ [bits/s/Hz]}. \quad (10)$$

In the above definitions $3|_{\text{dB}} = 10 \log_{10} 2$, and $\dot{C}_M(0)$ and $\ddot{C}_M(0)$ are the first and second derivatives (whenever exist) with respect to P of the per-cell sum-rate capacity, respectively, evaluated at $P = 0$. Focusing on Gaussian channels with receiver CSI only, it can be shown [22] that there is no need to calculate the two derivatives of the rate in $P = 0$, and that the low-SNR parameters are simply given by

$$\begin{aligned} \frac{E_b}{N_{0\min}} &= \frac{MK \log 2}{\text{tr}(\mathbb{E} \mathbf{H}_M^\dagger \mathbf{H}_M)} \\ S_0 &= \frac{2 (\text{tr}(\mathbb{E} \mathbf{H}_M^\dagger \mathbf{H}_M))^2}{M \text{tr}(\mathbb{E} (\mathbf{H}_M^\dagger \mathbf{H}_M)^2)}. \end{aligned} \quad (11)$$

b) The High-SNR Regime: This is usually the operating regime for high-data rate (high spectral efficiency) systems (that is the case actually in all 2.5/3 G standards).

The high-SNR regime is characterized through the high-SNR slope (also referred to as the “multiplexing gain,” or “pre-log”)

$$S_\infty \triangleq \lim_{P \rightarrow \infty} \frac{C_M(P)}{\log P} = \lim_{P \rightarrow \infty} P \dot{C}_M(P) \quad (12)$$

and the high-SNR power offset

$$\mathcal{L}_\infty \triangleq \lim_{P \rightarrow \infty} \frac{1}{\log 2} \left(\log P - \frac{C_M(P)}{S_\infty} \right) \quad (13)$$

yielding the following affine capacity approximation:

$$C_M(P) \approx \frac{S_\infty \log 2}{3|_{\text{dB}}} (P|_{\text{dB}} - 3|_{\text{dB}} \mathcal{L}_\infty). \quad (14)$$

Note that the high-SNR approximation reference channel here is that of a single isolated cell, with no fading, and total average transmit power P .

The high-SNR characterization of the per-cell sum-rate supported by the “soft-handoff” uplink channel is known only in certain limited scenarios (see Section III) and is the main focus of this work.

D. Main Results

Recall the definition of $C_M(P)$, cf. (3). Starting with intra-cell TDMA scheme where only one user is active per cell transmitting with power P we have the following.

Theorem 1: [intra-cell TDMA scheme $K = 1$, high-SNR characterization] Assume (H1) and (H2).

- a) For every $P > 0$, $C_M(P)$ converges as M goes to infinity. We call the limit $C(P)$.
- b) We get the following bounds on $C(P)$:

$$\begin{aligned} \max(\mathbb{E}_{\pi_a} \log(1 + P|x|^2), \mathbb{E}_{\pi_b} \log(1 + P|y|^2)) \\ \leq C(P) \leq \mathbb{E}_{\pi_a, \pi_b} \log(1 + P(|x|^2 + |y|^2)). \end{aligned}$$

- c) Further assume [(H3) or (H3')]. As P goes to infinity

$$C(P) = \log P + 2 \max(\mathbb{E}_{\pi_a} \log |x|, \mathbb{E}_{\pi_b} \log |x|) + o(1).$$

In particular

$$\begin{aligned} \mathcal{S}_\infty &= 1 \\ \mathcal{L}_\infty &= -2 \max(\mathbb{E}_{\pi_a} \log_2 |x|, \mathbb{E}_{\pi_b} \log_2 |x|). \end{aligned}$$

Note that point c) shows that the lower bound of point b) is tight in the high-SNR regime.

Proof: The proof of points a) and c) follows from Proposition 5 of Appendix A, where we prove that the variable $\mathcal{C}_M(P) \triangleq 1/M \log \det \mathbf{G}_M$ converges almost surely (a.s.). Note however that

$$0 \leq \frac{1}{M} \log \det \mathbf{G}_M \leq \frac{1}{M} \sum_{m=1}^M \log(1 + \rho(|a_m|^2 + |b_m|^2)) \quad (15)$$

where the second inequality is due to Hadamard’s inequality for semi-positive definite (SPD) Hermitian matrices. With (H1), it follows that $\mathcal{C}_M(P)$ is uniformly integrable, and hence the a.s. convergence implies convergence in expectation. Recalling that $C_M(P) = \mathbb{E} \mathcal{C}_M(P)$ completes the proof of point a) and c).

Let us show point b) using the tools of [24]. We first show the lower bound. We consider \mathbf{n}, \mathbf{x} , and \mathbf{y} as in (1).

$$\begin{aligned} C_M(P) &= \frac{1}{M} I(\mathbf{x}; \mathbf{y} | (a_i)_{1 \leq i \leq M}, (b_i)_{1 \leq i \leq M}) \\ &= \frac{1}{M} \sum_{j=1}^M I(x_j; \mathbf{y} | (x_i)_{1 \leq i < j}, (a_i)_{1 \leq i \leq M}, (b_i)_{1 \leq i \leq M}) \\ &\geq \frac{1}{M} \sum_{j=1}^M I(x_j; y_{j-1} | (x_i)_{1 \leq i < j}, (a_i)_{1 \leq i \leq M}, (b_i)_{1 \leq i \leq M}) \\ &= \frac{1}{M} \sum_{j=1}^M I(x_j; b_{j-1} x_j + n_{j-1} | b_{j-1}) \end{aligned}$$

which is the per-cell sum-rate capacity of a single-user fading channel. Therefore, the lower bound is [25] $\mathbb{E}_{\pi_b} \log(1 + P|y|^2)$. As argued in the Proof of Proposition 5 in Appendix A, we can exchange the role of π_a and π_b , thereby getting the claimed lower bound. Finally, the upper bound of b) follows immediately from Hadamard’s inequality for SPD Hermitian matrices. \square

In the Proof of Proposition 5 (intra-cell TDMA scheme), we use ideas from the theory of product of random matrices. Note that $C_M(P) = 1/M \sum_{m=1}^M \log(1 + P\lambda_m)$ where $\{\lambda_m\}_{m=1}^M$ are the eigenvalues of $\mathbf{H}_M \mathbf{H}_M^\dagger$, and the analysis of capacity hinges upon the study of spectral properties of $\mathbf{H}_M \mathbf{H}_M^\dagger$. The main idea is to link the spectral properties of the latter matrix with the exponential growth of the elements of its eigenvectors. Since $\mathbf{H}_M \mathbf{H}_M^\dagger$ is a *Hermitian Jacobi* matrix, hence tridiagonal, its eigenvectors can be considered as sequences with second-order linear recurrence. Therefore, the problem boils down to the study of the exponential growth of products of two-by-two matrices. This is closely related to the evaluation of the top Lyapunov exponent of the product. The explicit link between $C_M(P)$ and the top Lyapunov exponent is the Thouless formula (see [26] or [27] for the background on Hermitian Jacobi matrices and Lyapunov exponents), a version of which we prove in Appendix D. We emphasize however that we do not use the Thouless formula or Lyapunov exponents explicitly in the Proof of Proposition 5.

Like in the result of Narula [14] described below in Section III, our approach uses the analysis of a certain Markov chain. Unlike [14], we are not able to explicitly evaluate the invariant measure of this chain. Instead, we use the theory of Harris chains to both prove convergence and continuity results for the chain. The appropriate definitions are introduced in the course of proving Proposition 5.

We remark that Theorem 1 continues to hold in a real setup, that is if instead of (H2), we assume

$$(H2') \quad \pi_a \text{ and } \pi_b \text{ are supported on } \mathbb{R} \text{ and are absolutely continuous with respect to Lebesgue measure on } \mathbb{R}.$$

Since the argument is identical, we do not discuss this case further. It is also noted that unlike the nonfading case, where intra-cell TDMA scheme is optimal (see [9]), it is proved to be suboptimal for $K > 1$ in the presence of fading [28], yet TDMA it is one of the most common access protocols in cellular systems.

Turning to the WB scheme (which is the capacity-achieving scheme [28]), where all the bandwidth is used for coding, and all K users are transmitting simultaneously with average power ρ (and total cell average power $P = K\rho$), we have the following less explicit high-SNR characterization.

Theorem 2: [WB scheme $K > 1$, high-SNR characterization] Assume (H1), (H2), and (H4), and $K > 1$.

- a) For every $P > 0$, $C_M(P)$ converges as M goes to infinity. We call the limit $C(P)$.
- b) We get the following bounds on $C(P)$:

$$\begin{aligned} \max(\mathbb{E} \log(1 + P|\mathbf{a}|^2/K), \mathbb{E} \log(1 + P|\mathbf{b}|^2/K)) \\ \leq C(P) \leq \mathbb{E} \log(1 + P(|\mathbf{a}|^2 + |\mathbf{b}|^2)/K) \end{aligned}$$

where the expectation is taken in the following way: the random variables \mathbf{a} and \mathbf{b} are independent, and \mathbf{a} (resp., \mathbf{b}) is a complex K -vector whose coefficients are independent and distributed according to π_a (resp., π_b).

- c) As P goes to infinity

$$C(P) = \log P + \mathbb{E} \log \left(\frac{e + |\mathbf{b}|^2}{K} \right) + o(1) \quad (16)$$

where the expectation is taken in the following way: the random variables e and \mathbf{b} are independent, and \mathbf{b} is a complex K -vector whose coefficients are independent and distributed according to π_b . The law of e is m_0 , which is the unique invariant probability of the Markov chain defined by

$$e_{n+1} = |\mathbf{a}_n|^2 \left(\frac{e_n + |\mathbf{b}_{n-1}|^2 \sin^2(\mathbf{a}_n, \mathbf{b}_{n-1})}{e_n + |\mathbf{b}_{n-1}|^2} \right) \quad (17)$$

where the coefficients of $(\mathbf{a}_n)_{n \geq 0}$ (resp., $(\mathbf{b}_n)_{n \geq -1}$) are i.i.d with distribution π_a (resp., π_b). Moreover, for any two arbitrary equal length vectors \mathbf{a}, \mathbf{b} ,

$$\sin^2(\mathbf{a}, \mathbf{b}) \triangleq 1 - \frac{|\langle \mathbf{a}, \mathbf{b} \rangle|^2}{|\mathbf{a}|^2 |\mathbf{b}|^2}. \quad (18)$$

In particular, $S_\infty = 1$ and $\mathcal{L}_\infty = -\mathbb{E} \log_2(\frac{e + |\mathbf{b}|^2}{K})$.

As with the case $K = 1$, point a) and c) of Theorem 2 follow from the a.s. convergence stated in Proposition 19 of Appendix C, using (H1) and (15). As with Proposition 5, we do not use the Thouless formula or Lyapunov exponents explicitly in the Proof of Proposition 19. The proof of point b) is the same as the proof of Theorem 1.b). It is worth mentioning that in contrast to Theorem 1, the nonasymptotic lower bound b) is not tight in general for large SNR. This is since it is an increasing function of K and converges to a rate of a single-user Gaussian scalar channel, which is smaller than the asymptotic rate of (23).

Note that although the roles of the sequences $\{\mathbf{a}_n\}$ and $\{\mathbf{b}_n\}$ in (17) are not symmetric, the expression (16) is symmetric in π_a and π_b , as is the case for $K = 1$.

We conclude this section by noting that while Theorem 2 (WB scheme $K > 1$) does not give explicit expressions for the high-SNR power offset as Theorem 1, its proof leads immediately to easily computable bounds. In the following, the notation is as in Theorem 2, and we let $e_n(a)$ denote the Markov chain (17), with initial condition $e_0(a) = a$. Note that in the case $e_n(\infty)$, we take the initial condition $e_1(\infty) = |\mathbf{a}_0|^2$.

Proposition 3: Assume (H1), (H2), and (H4), and $K > 1$. Then

$$\begin{aligned} \mathbb{E} \log \left(\frac{e_n(0) + |\mathbf{b}|^2}{K} \right) &\leq \lim_{P \rightarrow \infty} [C(P) - \log P] \\ &\leq \mathbb{E} \log \left(\frac{e_n(\infty) + |\mathbf{b}|^2}{K} \right) \end{aligned}$$

where the expectation is taken in the following way. $e_n(0)$ (resp., $e_n(\infty)$) and \mathbf{b} are independent. \mathbf{b} is a complex K -vector whose coefficients are independent and distributed according to π_b . $e_n(0)$ (resp., $e_n(\infty)$) is the n th step of the Markov chain defined by (17) with initial condition $e_0(0) = 0$ (resp., $e_0(\infty) = \infty$).

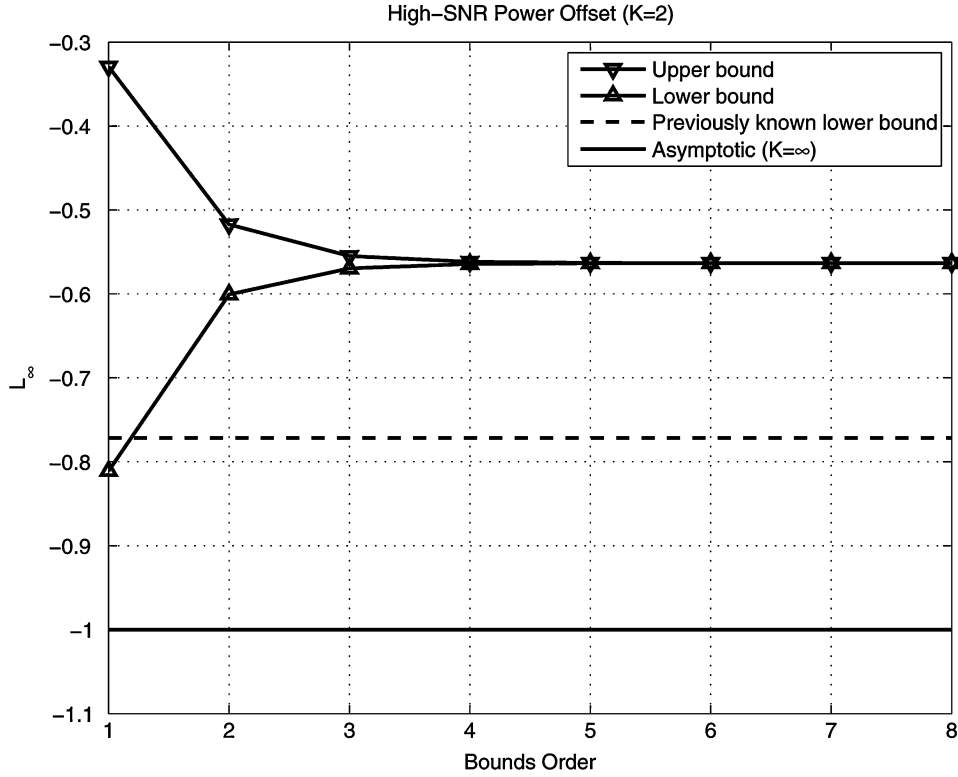


Fig. 2. High-SNR power offset bounds for Rayleigh fading, $K = 2$, and bounds order $n = 1, 2, \dots, 8$.

Indeed, since the expression (17) for e_{n+1} is monotone increasing in e_n , the law of e in Theorem 2 is stochastically dominated below by the law of e_n with initial condition 0, and stochastically dominated above by the law of e_n with initial condition ∞ . That same monotonicity also shows that the sequences of laws of $e_n(0)$ (resp., $e_n(\infty)$) are monotone increasing (resp., decreasing) with respect to stochastic order.

As a direct consequence of Proposition 3 with $n = 1$ and (13), we get the following bounds on the high-SNR power offset:

$$-\mathbb{E} \log_2 \left(\frac{|\mathbf{a}|^2 + |\mathbf{b}|^2}{K} \right) \leq \mathcal{L}_\infty$$

$$\leq -\mathbb{E} \log_2 \left(\frac{|\mathbf{a}|^2 \sin^2(\mathbf{a}, \mathbf{b}) + |\mathbf{b}'|^2}{K} \right) \quad (19)$$

where the expectation is taken in the following way: \mathbf{a}, \mathbf{b} , and \mathbf{b}' are independent, and \mathbf{a} (resp., \mathbf{b}, \mathbf{b}') is a complex K -vector whose coefficients are independent and distributed according to π_a (resp., π_b). Note that for K going to infinity, if we assume $\pi_a = \pi_b$ and zero mean, then $\sin^2(\mathbf{a}, \mathbf{b})$ converges to 1, therefore, the ratio between the upper and lower bound of (19), converges to 1, which also agrees with the asymptotic result of (38).

1) *Numerical Results:* In Figs. 2 and 3, we present the high-SNR power offset bounds of Proposition 3 in the special case of Rayleigh fading (real and imaginary parts are independent Gaussian random variables with zero mean and variance $1/\sqrt{2}$), for $K = 2$ and $K = 10$ users per cell, respectively. The curves are produced by Monte Carlo simulation with 10^5 samples. The figures include also the lower bound of [11], see (39), and the asymptotic results (and lower bound) for large number of users per cell $\mathcal{L}_\infty = -1$ (achieved by taking K

to infinity in (39)). Examining the figures it is observed that the new bounds are getting tighter with their order n and that the new lower bound is tighter than (39) already for $n = 2$. Moreover, fixing the order n , the new bounds are getting tighter with the number of users per cell K . This observation is also evident from Fig. 4, where the bounds are plotted for a fixed order $n = 2$ versus the number of users per cell K . Finally, since the upper bound of Fig. 2 is negative, we conclude that the presence of Rayleigh fading is beneficial over nonfading channels in the high-SNR region already for $K = 2$. (See [11] for a similar conclusion in the low-SNR region.)

III. BACKGROUND, PREVIOUS RESULTS AND BOUNDS

In this section, we briefly summarize previous work on the “soft-handoff” uplink cellular model introduced in [10], [11]. For conciseness, we restrict the discussion to the case where $\pi_a = \pi_b$. Most of the results in the sequel can be extended to include the general case where $\pi_a \neq \pi_b$.

Starting with nonfading channels (i.e., when π_a and π_b are singletons at 1), the per-cell sum-rate capacity of the uplink channel is given for $M \rightarrow \infty$ by [11]

$$R_{\text{nf}} = \log \left(\frac{1 + 2P + \sqrt{1 + 4P}}{2} \right). \quad (20)$$

This rate is achieved by any symmetric intra-cell protocol with average transmit power of P (e.g., intra-cell TDMA, and WB protocols). It is noted that the same result holds also for phase fading processes [13].

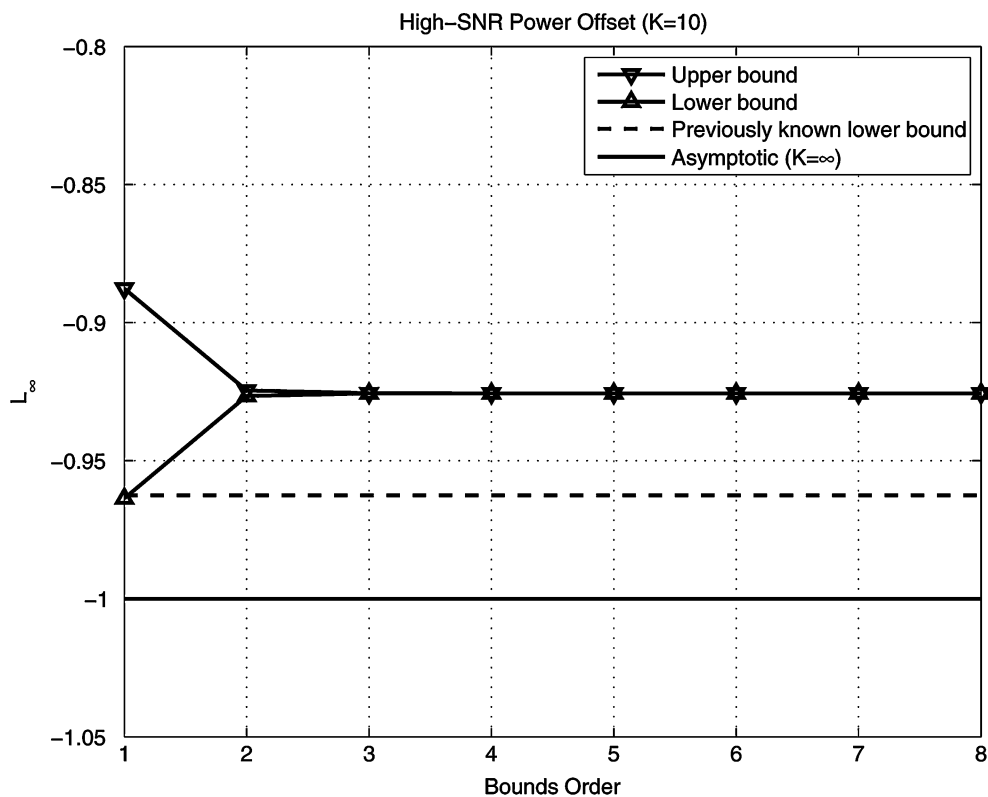


Fig. 3. High-SNR power offset bounds for Rayleigh fading, $K = 10$, and bounds order $n = 1, 2, \dots, 8$.

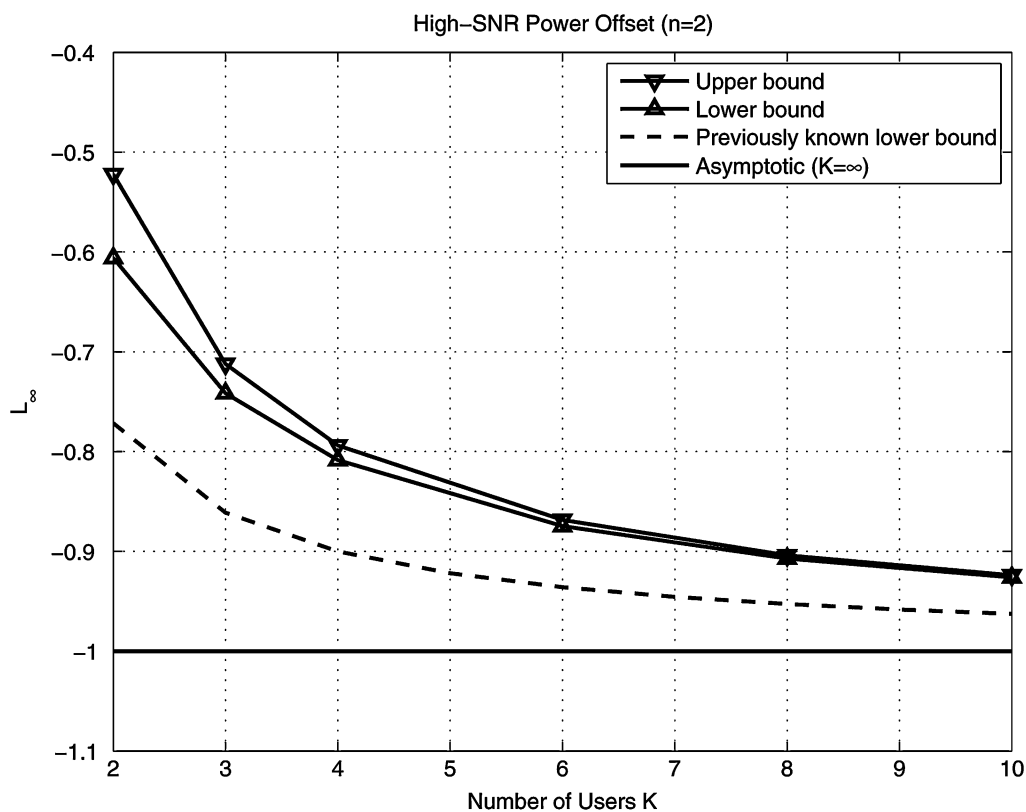


Fig. 4. High-SNR power offset bounds (order $n = 2$) for Rayleigh, and $K = 2, 3, 4, 6, 8, 10$. Note that for $K = 1$, $\mathcal{L}_\infty = \frac{\gamma}{\log 2} \approx 0.833$.

The extreme SNR characterization of (20) is summarized for the nonfading setup by

$$S_0 = \frac{4}{3}, \quad \frac{E_b}{N_{0 \min}} = \frac{\log 2}{2}, \quad S_\infty = 1, \quad \mathcal{L}_\infty = 0. \quad (21)$$

Returning to the flat-fading setup, the channel coefficients are taken as i.i.d. random variables, and

$$\begin{aligned} m_1 &\triangleq \mathbb{E}(a_{m,k}) = \mathbb{E}(b_{m,k}) \\ m_2 &\triangleq \mathbb{E}(|a_{m,k}|^2) = \mathbb{E}(|b_{m,k}|^2) \\ m_4 &\triangleq \mathbb{E}(|a_{m,k}|^4) = \mathbb{E}(|b_{m,k}|^4), \mathcal{K} \triangleq \frac{m_4}{m_2^2} \end{aligned} \quad (22)$$

denote the mean, second-power moment, fourth-power moment, and the kurtosis of an individual fading coefficient $\forall m, k$.

The per-cell sum-rate capacity of the WB scheme with fixed P and increasing number of users and cells $M, K \rightarrow \infty$, is given by [11].³

$$R_{\text{wb-f}} = \log \left(\frac{1 + 2Pm_2 + \sqrt{1 + 4Pm_2 + 4P^2(m_2^2 - |m_1|^4)}}{2} \right). \quad (23)$$

The rate is maximized for a zero-mean fading distribution and is given by

$$R_{\text{wb-f}} = \log(1 + 2m_2P). \quad (24)$$

Comparing (20) and (24) (with $m_2 = 1$), it follows that the presence of fading is beneficial in case the number of users is large. We note that (23) is also shown in [11] to upper-bound the respective rate for any finite number of users K .

Returning to the intra-cell TDMA ($K = 1$), for which standard random matrix theory is not suitable (see Section II-B), the powerful moment bounding technique employed in [28] for the Wyner model can be utilized to obtain lower and upper bounds on the per-cell sum-rate.

An alternative approach which replaces the role of the singular values with the diagonal elements of the *Cholesky* decomposition of the matrix \mathbf{G}_M , was presented by Narula [14] for a two-diagonal nonzero channel matrix \mathbf{H}_M whose entries are i.i.d. zero-mean complex Gaussian (Rayleigh fading). Originally, Narula had studied the capacity of a time-varying two taps ISI channel, where the channel coefficients are i.i.d. zero-mean complex Gaussian. With the above assumptions regarding the ISI channel coefficients it is easy to verify that the capacity of this model is equal to the per-cell sum-rate capacity of an uplink intra-cell TDMA scheme employed in the “soft-handoff” model.

Following [14], we use the *Cholesky* decomposition applied to the covariance matrix of the uplink intra-cell TDMA scheme output vector $\mathbf{G}_M = \mathbf{L}_M \mathbf{D}_M \mathbf{U}_M$, where \mathbf{L}_M (resp., \mathbf{U}_M) is a lower triangular (resp., upper triangular) matrix with 1 on the

diagonal. The diagonal entries of \mathbf{G}_M are given (with $K = 1$) by

$$d_m = 1 + P|a_m|^2 + P|b_m|^2 \left(1 - P \frac{|a_{m-1}|^2}{d_{m-1}} \right), \quad m = 2, \dots, M \quad (25)$$

where the initial condition of (26) is $d_1 = 1 + P|a_1|^2 + P|b_1|^2$. Thus, the diagonal entries $\{d_m\}$ form a discrete-time continuous-space Markov chain; Narula’s main observation was that this chain possesses a unique ergodic stationary distribution, given by

$$f_d(x) = \frac{\log(x)e^{-\frac{x}{P}}}{\text{Ei}(\frac{1}{P})P}, \quad x \geq 1 \quad (26)$$

where $\text{Ei}(x) = \int_x^\infty \frac{\exp(-t)}{t} dt$ is the exponential integral function. Further, as is proved in [14], the strong law of large numbers (SLLN) holds for the sequence $\{\log d_m\}$ as $M \rightarrow \infty$. Hence, the average per-cell sum-rate capacity of the intra-cell TDMA scheme ($K = 1$) can be expressed as

$$\begin{aligned} R_{\text{tdma-f}} &= \lim_{M \rightarrow \infty} \mathbb{E} \left(\frac{1}{M} \log \det \mathbf{G}_M \right) \\ &= \lim_{M \rightarrow \infty} \mathbb{E} \left(\frac{1}{M} \log \det (\mathbf{L}_M \mathbf{D}_M \mathbf{U}_M) \right) \\ &= \lim_{M \rightarrow \infty} \mathbb{E} \left(\frac{1}{M} \sum_{m=0}^M \log d_m \right) = \mathbb{E}_{\pi_d}(\log d) \end{aligned} \quad (27)$$

where the last expectation is taken with respect to $f_d(x)$, as defined in (26). In particular

$$R_{\text{tdma-f}} = \int_1^\infty \frac{(\log(x))^2 e^{-\frac{x}{P}}}{\text{Ei}(\frac{1}{P})P} dx. \quad (28)$$

Narula’s approach is based on an explicit calculation of the invariant distribution f_d , and is thus tied to Rayleigh fading. Modifications of key parameters (such as the entries’ probability density function (PDF), and the number of nonzero diagonals) lead to analytically intractable expressions.

Another result derived by following the footsteps of [14] is an upper bound on the per-cell sum-rate of the WB scheme with finite K and infinite number of cells $M \rightarrow \infty$, in the presence of a general fading distribution, given by

$$\begin{aligned} R_{\text{wbk-f}} &\leq \log \left(\frac{1 + 2Pm_2}{2} \right. \\ &\quad \left. + \frac{\sqrt{1 + 4Pm_2 + 4P^2(1 - \frac{1}{K})(m_2^2 - |m_1|^4)}}{2} \right) \end{aligned} \quad (29)$$

$$(x) \quad (30)$$

and in the special case of zero mean unit power ($m_1 = 0$, $m_2 = 1$) fading distribution (e.g., Rayleigh fading) the bound reduces to

$$R_{\text{wbk-f}} \leq \log \left(\frac{1 + 2P + \sqrt{(1 + 2P)^2 - (4P^2/K)}}{2} \right). \quad (31)$$

³Here, the number of users K is taken to infinity and then the number of cells M is taken to infinity.

This result which is proved in [14] for $K = 1$ (intra-cell TDMA protocol) and expanded to an arbitrary K in [15], is derived by noting that the average of the determinant of the received vector covariance matrix \mathbf{G}_M can be recursively expressed by

$$\mathbb{E}(\det \mathbf{G}_m) = A \mathbb{E}(\det \mathbf{G}_{m-1}) - B \mathbb{E}(\det \mathbf{G}_{m-2}), \quad m = 3, \dots, M \quad (32)$$

with initial conditions

$$\mathbb{E}(\det \mathbf{G}_1) = A, \quad \mathbb{E}(\det \mathbf{G}_2) = A^2 - B \quad (33)$$

where

$$A = 1 + 2Pm_2, \quad B = \frac{P^2}{K} (m_2^2 + (K-1)|m_1|^4). \quad (34)$$

See Appendix E for more details. The solution to (32) is given by

$$\mathbb{E}(\det \mathbf{G}_m) = \varphi r^m - \phi s^m \quad (35)$$

where

$$r = \frac{1}{2}(A + \sqrt{A^2 - 4B}), \quad s = \frac{1}{2}(A - \sqrt{A^2 - 4B}) \quad (36)$$

are real and positive, and φ, ϕ are determined by the initial conditions (33). Finally, (29) is derived by the following set of inequalities:

$$\begin{aligned} R_{\text{wbk-f}} &= \lim_{M \rightarrow \infty} \frac{1}{M} \mathbb{E}(\log \det \mathbf{G}_M) \\ &\leq \lim_{M \rightarrow \infty} \frac{1}{M} \log \mathbb{E}(\det \mathbf{G}_M) = \log r \end{aligned} \quad (37)$$

where the inequality is due to Jensen's inequality, and the last equality follows from the fact that $r > s$, and $M \rightarrow \infty$. In the case of $K = 1$, the upper bound of (31) coincides with the per-cell sum-rate capacity of the nonfading setup (20). Thus, the presence of Rayleigh fading decreases the rates of the intra-cell TDMA protocol supported by the "soft-handoff" model. Nevertheless, it is shown in [11] that already for $K = 2$ the presence of fading may be beneficial at least for low SNR values. The tightness of the bound is demonstrated by noting the for $K \rightarrow \infty$ it coincides with the asymptotic expression of (23).

The extreme SNR characterization of the WB rate for $M \rightarrow \infty$ in the presence of a general fading distribution is summarized by [11]

$$\begin{aligned} S_0 &= \frac{2}{\frac{\kappa}{2K} + \frac{|m_1|^4}{2m_2^2} + 1}, \quad \frac{E_b}{N_{0 \min}} = \frac{\log 2}{2m_2} \\ S_\infty &\leq 1, \\ &-\log_2 \left(m_2 + \sqrt{\left(1 - \frac{1}{K}\right) (m_2^2 - |m_1|^4)} \right) \leq \mathcal{L}_\infty. \end{aligned} \quad (38)$$

The bounds of the high-SNR parameters are tight for $K \gg 1$. For the special case of Rayleigh fading the extreme SNR characterization are given by [11]

$$\begin{aligned} S_0 &= \frac{2}{1 + \frac{1}{K}}, \quad \frac{E_b}{N_{0 \min}} = \frac{\log 2}{2} \\ S_\infty &= 1, \quad -\log_2 \left(1 + \sqrt{1 - \frac{1}{K}} \right) \leq \mathcal{L}_\infty \leq \frac{\gamma}{\log 2} \end{aligned} \quad (39)$$

where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant. It is noted that the right inequality of the high-SNR power offset is tight for $K = 1$, while the left inequality is tight for $K \gg 1$. The beneficial effects of Rayleigh fading and increasing number of users are evident when compared to the nonfading extreme-SNR parameters of the respective nonfading setup (21).

To conclude this section, we emphasize that calculating exact expressions for the high-SNR parameters of the WB protocol rate with finite number of users per cell and *general* fading distribution remains an open problem.

IV. APPLICATIONS

In this section, we present several applications of the main results presented in this work (see Section II-D).

a) Intra-Cell TDMA and Rayleigh Fading: Assuming that only one user is active per cell $K = 1$ and symmetric Rayleigh-fading channels (i.e., $\pi_{|a|^2}$ and $\pi_{|b|^2}$ are exponential distributions with parameter 1), the high-SNR power offset is given according to Theorem 1, by

$$\begin{aligned} \mathcal{L}_\infty &= -\max(\mathbb{E}(\log_2 |a|^2), \mathbb{E}(\log_2 |b|^2)) \\ &= \frac{-1}{\log 2} \int_0^\infty e^{-x} \log x \, dx = \frac{\gamma}{\log 2} \end{aligned} \quad (40)$$

where the last equality is due to [29, p. 567, formula 4.331.1]. Obviously, this result coincides with the high-SNR power offset derived by applying the definition of \mathcal{L}_∞ (see (13)) directly to the exact expression derived in [14] (see (28)).

Note that the same result holds if an attenuation factor is added to one of the fading paths, e.g., $\tilde{b}_m = \alpha b_m$ where $b_m \sim \mathcal{CN}(0, 1)$ and $\alpha \in [0, 1]$; this follows directly from Theorem 1, but not from [14], which requires symmetric fading paths (i.e., $\alpha = 1$).

b) Intra-Cell TDMA and General Fading Statistic: Consider the following single-user single-input single-output (SISO) flat-fading channel for an arbitrary time index

$$y = ax + n \quad (41)$$

where x is the input signal $x \sim \mathcal{CN}(0, P)$, and n is the additive circularly symmetric Gaussian noise $n \sim \mathcal{CN}(0, 1)$. In addition, a is the fading coefficient $a \sim \pi_a$ satisfying conditions (H1)–(H3) and known only to the receiver (receiver CSI). Assuming that the fading process is also ergodic in the time domain, the ergodic capacity of the channel is given by [25]

$$C = \mathbb{E}_{\pi_a} \log(1 + P|a|^2) \quad (42)$$

where the expectation is taken over the fading distribution π_a . Accordingly, under the mild conditions (H1)–(H3), the high-SNR regime of this channel is characterized by

$$S_\infty = 1, \quad \mathcal{L}_\infty = -\mathbb{E}_{\pi_a} \log_2 |a|^2. \quad (43)$$

Using Theorem 1, we can now establish the following analogy between the multiple-cell setup and the SISO channel at hand.

Corollary 4: The high-SNR characterization of the intra-cell TDMA per-cell sum-rate supported by the "soft-handoff" setup with fading distributions π_a, π_b such that

$\mathbb{E}_{\pi_a} \log_2 |a|^2 > \mathbb{E}_{\pi_b} \log_2 |b|^2$ coincides with those of a scalar single-user fading channel with fading distribution π_a .

This observation allows us to use the vast body of work done for the celebrated scalar flat-fading channel [25]. In particular, the high-SNR characterization of flat-fading channels with the following fading statistics have been considered in previous works: a) Rayleigh distribution, b) Rice distribution, c) log-normal distribution, and d) Nakagami distribution (see [25] and references therein).

c) *Intra-Cell TDMA and Opportunistic Scheduling:* Throughout this work we have assumed that the instantaneous channel state information is known to the MCP receiver only. Here, we further assume that some sort of ideal feedback channel is available between the MCP receiver and the K mobile users included in each cell. This feedback channel is used to schedule the “best” local user in each cell for transmission during the current time slot.⁴ In other words, in each cell the user with the strongest channel fade towards the BS located on the right boundary of each cell is scheduled for transmission⁵ with power P . Hence, the index of the selected user in the m th cell reads

$$\tilde{k}_m = \underset{k=1,2,\dots,K}{\operatorname{argmax}} |a_{m,k}|^2, \quad m = 1, 2, \dots, M. \quad (44)$$

The resulting $M \times (M+1)$ channel transfer matrix $\tilde{\mathbf{H}}_M$ of this scheduling scheme is a two-diagonal matrix with independent entries. The PDF of the main diagonal i.i.d. entries’ amplitudes is given by

$$d\pi_{K,|a|^2} = K\pi_{|a|^2}^{K-1} d\pi_{|a|^2} \quad (45)$$

following the maximum order statistics [31]. On the other hand, the i.i.d. entries of the second nonzero diagonal are distributed according to the original fading statistics π_b .

Assuming that $\pi_{K,|a|^2}$ and π_b satisfy conditions (H1)–(H3), we can apply Theorem 1 in order to derive the high-SNR characteristics of the per-cell sum-rate achievable by this opportunistic scheduling

$$\mathcal{S}_\infty = 1 \\ \mathcal{L}_\infty = -\max \left(\mathbb{E}_{\pi_{K,|a|^2}} (\log_2 y), \mathbb{E}_{\pi_b} (\log_2 |b|^2) \right). \quad (46)$$

For Rayleigh-fading channels and in the case where the number of users per cell is large $K \gg 1$, we can use the well-known fact that the square of the maximum of the K amplitudes behaves like $\log K$ with high probability (see [32]). Hence, the rate high-SNR power offset of this scheme is

$$\mathcal{L}_\infty \approx -\log_2 \log K \quad (47)$$

revealing a multiuser diversity gain of $\log \log K$. It is noted that allowing additional power control to this scheme will yield better performances. However, we are unable to apply Theorem 1 for this situation. Finally, choosing the BS located on the right

⁴See [30] for a similar scheduling deployed in the Wyner cellular uplink channel.

⁵Since the right-most cell indexed $(M+1)$ has no BS on its right boundary it randomly schedules a user for transmission.

boundary of the cell is arbitrary; taken the BS located on the left boundary of the cell yields the same results.

V. CONCLUDING REMARKS

In this paper, we study the high-SNR characterization of the per-cell sum-rate capacity of the “soft-handoff” uplink cellular channel with multiple-cell processing. Taking advantage of the special topology induced by the setup, the problem reduces to the study of the spectrum of certain large random Hermitian Jacobi matrices. For the intra-cell TDMA protocol where only one user is active simultaneously per cell, we provide an exact closed-form expression for the per-cell sum-rate high-SNR power offset for rather general fading distribution. Examining the result, it is concluded that in the high-SNR regime, the rate of the cellular setup at hand is equivalent to the one of a single-user SISO channel with similar fading statistics.

Turning to the capacity-achieving WB protocol, where all K users are active simultaneously in each cell, we derive a series of lower and upper bounds to the rate. These bounds are shown (via Monte Carlo simulations) to be tighter than previously known bounds.

Note that in Theorem 2 points a) and c) and in Proposition 3, we take the fading coefficients relative to the users of one cell to be independent. Those results continue to be true if we assume correlation between the fading coefficients relative to the users of the same cell (but independence between cells). The proof is identical to the proof given in the paper.

Moreover, note that bounds like the ones of Proposition 3 can be derived for finite P using the random sequence e_n defined by (65).

Some of the analysis reported here can be extended to include the case where \mathbf{G}_M is $(2p-1)$ -diagonal for some $p > 2$ (e.g., $p = 3$ for the channel matrix of the Wyner model), using an adaptation of the “Thouless formula for the strip” derived originally in [33]. Using this approach, bounds similar to those of Proposition 3 may be provided on the rate. Details will appear elsewhere [34].

APPENDIX

A. Proof of Theorem 1

In order to streamline the proof we somewhat modify notation. We consider two random sequences of complex numbers (a_n) and (b_n) . The (a_n) (resp., (b_n)) are i.i.d. of law π_a (resp., π_b) and the (a_n) are independent of the (b_n) . We set $\Omega \triangleq ((a_n), (b_n))$. We denote by \mathbb{P} the probability associated with those random sequences and by \mathbb{E} the associated expectation. For a given integer n , we consider a channel transfer matrix \mathbf{H}_M of size $M \times (M+1)$.

$$\mathbf{H}_M = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_M & b_M \end{pmatrix}.$$

We consider the following random variable:

$$\mathcal{C}_M(P) = \frac{1}{M} \operatorname{tr} \{ \log(I + P\mathbf{H}_M\mathbf{H}_M^\dagger) \}$$

where $\mathbf{H}_M \mathbf{H}_M^\dagger$ is explicitly given by (48), shown at the bottom of the page.

With this notation, as explained in Section II-D, Theorem 1 follows from the following.

Proposition 5: $[K = 1]$ Assume (H1) and (H2).

- a) For every $\rho > 0$, $\mathcal{C}_M(P)$ converges \mathbb{P} -a.s. as M goes to infinity. We call the limit $\mathcal{C}(P)$.
- b) Further assume [(H3) or (H3')]. As ρ goes to infinity

$$\mathcal{C}(P) = \log P + 2 \max(\mathbb{E}_{\pi_a} \log |x|; \mathbb{E}_{\pi_b} \log |x|) + o(1).$$

Proof of Proposition 5: Without loss of generality, in the proof we can assume

$$(H5) \quad \mathbb{E}_{\pi_a} \log |x| \leq \mathbb{E}_{\pi_b} \log |x|.$$

Indeed, we may exchange the role of entries a_i and b_i for $1 \leq i \leq M$ by a right-left reflection, namely, the transformation $\hat{a}_j = b_{M-j+1}$, $\hat{b}_j = a_{M-j+1}$, $1 \leq j \leq M$.

For part a), only (H1) and (H2) are needed. Since part a) is a consequence of general facts concerning products of random matrices and does not use much of the special structure in the problem, we bring it in Appendix D.

Part b) uses the theory of Markov chains and is specific to the particular matrix \mathbf{H}_M . We note that as a by-product of this approach, we obtain a second proof of part a), however, under the additional assumption [(H3) or (H3')]. We provide a Proof of Proposition 5 under the assumptions (H1), (H2), and [(H3) or (H3')] in Appendices A and B.

The structure of the Proof of Proposition 5 is as follows. We first introduce an auxiliary sequence which allows us to reformulate the problem in terms of a special Markov chain. The study of the latter, which forms the bulk of the Proof of Proposition 5, is carried out in Appendix B.

1) *Auxiliary Sequence:* By (H2), the elements of $\mathbf{H}_M \mathbf{H}_M^\dagger$ above the diagonal are nonzero a.s. Therefore, by [35, Lemma 7.7.1], the eigenvalues of $\mathbf{H}_M \mathbf{H}_M^\dagger$ are distinct. In the sequel, we denote by $\lambda_1 \geq \dots \geq \lambda_M$ the ordered eigenvalues of $\mathbf{H}_M \mathbf{H}_M^\dagger$. For a given λ , we consider the following sequence (indexed by n) of complex numbers (the dependence in λ will only be mentioned when it is relevant): $x_0 = 0, x_1 = 1$, and for $n \geq 1$

$$a_n b_{n-1}^\dagger x_{n-1} + (|a_n|^2 + |b_n|^2) x_n + a_{n+1}^\dagger b_n x_{n+1} = \lambda x_n$$

that is

$$x_{n+1} = \frac{\lambda - |a_n|^2 - |b_n|^2}{a_{n+1}^\dagger b_n} x_n - \frac{a_n b_{n-1}^\dagger}{a_{n+1}^\dagger b_n} x_{n-1}. \quad (49)$$

Note that $x_{M+1}(\lambda) = 0$ if and only if λ is an eigenvalue of $\mathbf{H}_M \mathbf{H}_M^\dagger$. Moreover, x_{n+1} is a polynomial in λ of degree n with highest coefficient $1 / \prod_{i=1}^n (a_{i+1}^\dagger b_i)$. One can thus write

$$x_{n+1}(\lambda) = \prod_{i=1}^n (a_{i+1}^\dagger b_i)^{-1} \prod_{i=1}^n (\lambda - \lambda_i), \quad \mathbb{P}\text{-a.s.}$$

Hence, for $\lambda = -1/P$

$$\begin{aligned} \mathcal{C}_M(P) &= \log(P) + \frac{1}{M} \log |x_{M+1}(\lambda)| \\ &\quad + \frac{1}{M} \sum_{i=1}^M \log |a_{i+1} b_i|, \quad \mathbb{P}\text{-a.s.} \end{aligned} \quad (50)$$

By the Law of Large Numbers (LLN),

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M \log |a_{i+1} b_i| = \mathbb{E}_{\pi_a} \log |x| + \mathbb{E}_{\pi_b} \log |x|, \quad \mathbb{P}\text{-a.s.}$$

Because of (50), to prove Proposition 5, we only need to show the following lemma.

Lemma 6: Assume (H1), (H2), and [(H3) or (H3')]

- a) For every $\lambda < 0$, $\frac{1}{n} \log |x_{n+1}(\lambda)|$ converges \mathbb{P} -a.s. as n goes to infinity. The limit is $\gamma(\lambda)$, the Lyapunov exponent defined by (69).
- b) Assume further (H5). Then $\gamma(\lambda)$ converges to $\mathbb{E}_{\pi_b} \log |x| - \mathbb{E}_{\pi_a} \log |x|$ as λ goes to 0.

2) *Reduction to a Markov Chain:* To prove Lemma 6, we take $c_n \triangleq x_n / x_{n-1}$, for $n \geq 2$. Note that by (49) and (H2), \mathbb{P} -a.s., $x_n \neq 0$, hence, c_n is well defined and nonzero. By (49), we get

$$c_{n+1} = \frac{\lambda - |a_n|^2 - |b_n|^2}{a_{n+1}^\dagger b_n} - \frac{a_n b_{n-1}^\dagger}{c_n a_{n+1}^\dagger b_n}.$$

Let $d_n = c_n a_n^\dagger b_{n-1}$. Then

$$\begin{aligned} d_{n+1} &= \lambda - |a_n|^2 - |b_n|^2 - \frac{|a_n|^2 |b_{n-1}|^2}{d_n} \\ &= \lambda - |b_n|^2 - |a_n|^2 \left(1 + \frac{|b_{n-1}|^2}{d_n} \right). \end{aligned} \quad (51)$$

Let $e_n = (1 + \frac{|b_{n-1}|^2}{d_n})$. Then $d_{n+1} = \lambda - |b_n|^2 - |a_n|^2 e_n$, and

$$e_n = \frac{-\lambda + |a_{n-1}|^2 e_{n-1}}{-\lambda + |b_{n-1}|^2 + |a_{n-1}|^2 e_{n-1}} \quad (52)$$

$$\mathbf{H}_M \mathbf{H}_M^\dagger = \begin{pmatrix} |a_1|^2 + |b_1|^2 & a_2^\dagger b_1 & 0 & \dots & 0 \\ a_2 b_1^\dagger & |a_2|^2 + |b_2|^2 & a_3^\dagger b_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & a_M^\dagger b_{M-1} \\ 0 & \dots & 0 & a_M b_{M-1}^\dagger & |a_M|^2 + |b_M|^2 \end{pmatrix} \quad (48)$$

with the initial conditions

$$c_2 = \frac{\lambda - |a_1|^2 - |b_1|^2}{a_2^\dagger b_1}$$

$$d_2 = \lambda - |b_1|^2 - |a_1|^2.$$

$d_2 \in \mathbb{R}$ and $d_2 < -|b_1|^2$, hence, $0 < e_2 < 1$. From (52) we conclude that for all $n, e_n \in \mathbb{R}$ and $0 < e_n < 1$. Now, for all n ,

$$c_n = \frac{d_n}{a_n^\dagger b_{n-1}} = \frac{b_{n-1}^\dagger}{a_n^\dagger} \frac{1}{e_n - 1}.$$

Then

$$\begin{aligned} \frac{1}{n} \log |x_{n+1}| &= \frac{1}{n} \sum_{i=2}^{n+1} \log |c_i| \\ &= \frac{1}{n} \sum_{i=2}^{n+1} \left(\log \left| \frac{b_{i-1}}{a_i} \right| - \log(1 - e_i) \right) \end{aligned} \quad (53)$$

where $\frac{1}{n} \sum_{i=2}^{n+1} \log \left| \frac{b_{i-1}}{a_i} \right|$ converges to $\mathbb{E}_{\pi_b} \log |x| - \mathbb{E}_{\pi_a} \log |x|$ by the LLN. We now study in details the Markov chain e_n .

B. Study of the Markov Chain e_n and Proof of Lemma 6

1) Proof of Point a): The proof is based on the Harris property of the Markov chain e_n . We first recall some definitions from the theory of Harris chains (Definitions 7–10). Then we state three lemmas (11–13), which we will need to prove that e_n is Harris with the right properties. From the Harris character of e_n , we deduce ergodic properties (Proposition 14), which will give the Proof of Lemma 6a).

For simplicity, we write $\delta \triangleq -\lambda$ and we re-index the chain so that it starts from e_0 . As in (52)

$$e_n = \frac{\delta + |a_{n-1}|^2 e_{n-1}}{\delta + |b_{n-1}|^2 + |a_{n-1}|^2 e_{n-1}}. \quad (54)$$

We denote by \mathbb{P}_{e_0} the law of the sequence starting from e_0 and by \mathbb{E}_{e_0} the associated expectation.

We first recall some definitions from the theory of Harris Markov chains, which will be used extensively in the proof. We refer the reader to [16] for the relevant background.

Definition 7: Denote by (r_n) a Markov chain on I an interval of \mathbb{R} . Set l a probability measure on I , it is an *irreducibility measure* if for all measurable set A such that $l(A) > 0$ and for all $r_0 \in I$

$$(\exists n) \mathbb{P}_{r_0}(r_n \in A) > 0.$$

l is a *maximal irreducibility measure* if it satisfies the following conditions.

- l is an irreducibility measure.
- For any other irreducibility measure l' , l' is absolutely continuous with respect to l .
- If $l(A) = 0$ then $l\{r_0 : (\exists n) \mathbb{P}_{r_0}(r_n \in A) > 0\} = 0$.
- For any irreducibility measure l' , l is equivalent to l'' defined by

$$l''(A) = \int_I l'(dr_0) \sum_{i=0}^{\infty} \frac{1}{2^i} \mathbb{P}_{r_0}(r_i \in A).$$

Definition 8: Denote by (r_n) a Markov chain on I an interval of \mathbb{R} . A set A is called *Harris recurrent* if for all $r_0 \in A$, \mathbb{P}_{r_0} -a.s., the chain r_n visits A an infinite number of times. The chain (r_n) is called *Harris recurrent* if given a maximal irreducibility measure l , every measurable set A such that $l(A) > 0$ is Harris recurrent.

Definition 9: Denote by (r_n) a Markov chain on I an interval of \mathbb{R} . Denote by l a maximal irreducibility measure. For every measurable set A such that $l(A) > 0$ we denote by τ_A the time when the chain (r_n) enters A . A measurable set B is called *regular* if for every measurable set A such that $l(A) > 0$

$$\sup_{r_0 \in B} \mathbb{E}_{r_0}(\tau_A) < \infty.$$

Definition 10: Denote by (r_n) a Markov chain on I an interval of \mathbb{R} . Denote by A and B two measurable sets. We say that B is *uniformly accessible* from A if there exists an $\varepsilon > 0$ such that

$$\inf_{r_0 \in A} \mathbb{P}_{r_0}((\exists n) r_n \in B) \geq \varepsilon.$$

We now prove three lemmas, which will be used in the proof.

Lemma 11: For $\alpha, \beta, \delta \in \mathbb{R}^+$, we define the function $\phi_{\alpha, \beta}$ (we suppress δ from the notation) such that for $e \in [0, 1]$

$$\phi_{\alpha, \beta}(e) = \frac{\delta + \alpha e}{\delta + \beta + \alpha e}.$$

For any given $e \in [0, 1]$, we define the sequence $(\theta_n(e))$ by $\theta_0 = e$ and for $n \geq 1$, $\theta_n(e) = \phi_{\alpha, \beta}(\theta_{n-1}(e))$. Then, $\phi_{\alpha, \beta}$ has exactly one fixed point in $[0, 1]$, say $\kappa_{\alpha, \beta}$, and $\theta_n(e)$ converges to $\kappa_{\alpha, \beta}$. Moreover, the convergence is uniform in the starting point in the following sense:

$$(\forall \varepsilon > 0) (\exists n_0 \in \mathbb{N}) (\forall e \in [0, 1]) (\forall n \geq n_0) (|\theta_n(e) - \kappa_{\alpha, \beta}| < \varepsilon). \quad (55)$$

Finally if $\alpha_1 < \alpha_2$ and $\beta_1 > \beta_2$, then $\kappa_{\alpha_1, \beta_1} < \kappa_{\alpha_2, \beta_2}$.

Proof: For $e \in [0, 1]$

$$\phi'_{\alpha, \beta}(e) = \frac{\alpha \beta}{(\delta + \beta + \alpha e)^2}.$$

$\phi'_{\alpha, \beta}$ is decreasing and $\phi'_{\alpha, \beta}(1) < 1$. If $\phi'_{\alpha, \beta}(0) < 1$, then $\phi_{\alpha, \beta}$ is contracting hence it admits a fixed point and its iteration on any starting point converges to the fixed point. Suppose $\phi'_{\alpha, \beta}(0) \geq 1$. Denote by \bar{e} the only point of $[0, 1]$ such that $\phi'_{\alpha, \beta}(\bar{e}) = 1$. Set $\tilde{\phi}_{\alpha, \beta}(e) = \phi(e)_{\alpha, \beta} - e$. Then $\tilde{\phi}_{\alpha, \beta}(0) > 0$, $\tilde{\phi}_{\alpha, \beta}(1) \leq 0$, and $\tilde{\phi}_{\alpha, \beta}$ is increasing on $[0, \bar{e}]$ and decreasing on $[\bar{e}, 1]$. Hence, $\tilde{\phi}_{\alpha, \beta}(\bar{e}) > 0$ and $\tilde{\phi}_{\alpha, \beta}$ is 0 on exactly one point which is a fixed point for $\phi_{\alpha, \beta}$. We denote that fixed point $\kappa_{\alpha, \beta}$. If $e \in [\kappa_{\alpha, \beta}, 1]$, since $\phi_{\alpha, \beta}$ is increasing, for all n , $\theta_n(e) \in [\kappa_{\alpha, \beta}, 1]$ and $\phi_{\alpha, \beta}$ is contracting on $[\kappa_{\alpha, \beta}, 1]$ hence $\theta_n(e)$ converges to $\kappa_{\alpha, \beta}$. If $e \in [0, \kappa_{\alpha, \beta}]$, for all n , $\theta_n(e) \in [0, \kappa_{\alpha, \beta}]$, and $\tilde{\phi}_{\alpha, \beta}$ is nonnegative on that interval, hence $\theta_n(e)$ is nondecreasing. Therefore, it converges and since $\phi_{\alpha, \beta}$ is continuous, the only possible limit is $\kappa_{\alpha, \beta}$. To prove the uniformity in the

starting point, we use the fact that $\phi_{\alpha,\beta}$ is increasing, hence for all $e \in [0, 1]$ and n

$$\theta_n(0) \leq \theta_n(e) \leq \theta_n(1).$$

That gives the uniformity. Finally, assume $\alpha_1 < \alpha_2$ and $\beta_1 > \beta_2$. $\phi_{\alpha,\beta}(e)$ is nondecreasing in α , decreasing in β , and non-decreasing in e hence by induction, $\phi_{\alpha_1,\beta_1}^n(0) \leq \phi_{\alpha_2,\beta_2}^n(0)$, where ϕ^n is the n th iteration of the function ϕ . Hence, $\kappa_{\alpha_1,\beta_1} \leq \kappa_{\alpha_2,\beta_2}$. If $\kappa_{\alpha_1,\beta_1} = \kappa_{\alpha_2,\beta_2}$, then

$$\begin{aligned} \kappa_{\alpha_1,\beta_1} &= \phi_{\alpha_1,\beta_1}(\kappa_{\alpha_1,\beta_1}) < \phi_{\alpha_2,\beta_2}(\kappa_{\alpha_1,\beta_1}) \\ &= \phi_{\alpha_2,\beta_2}(\kappa_{\alpha_2,\beta_2}) = \kappa_{\alpha_2,\beta_2} \end{aligned}$$

which gives a contradiction. \square

Lemma 12: Assume (H2) and [(H3) or (H3')].

- For $e_0 \in [0, 1]$, there exist two sequences $(\theta_n^1(e_0))$ and $(\theta_n^2(e_0))$ in $[0, 1]$ such that the law of e_n under \mathbb{P}_{e_0} and the Lebesgue measure on $[(\theta_n^1(e_0)), (\theta_n^2(e_0))]$ are mutually absolutely continuous.
- $(\theta_n^1(e_0))$ and $(\theta_n^2(e_0))$ converge to, say, Θ^1 and Θ^2 , respectively, Θ^1 and Θ^2 are independent of e_0 and $\Theta^1 < \Theta^2$. Finally, the convergence is uniform in the starting point in the sense of Lemma 11.
- If $e_0 \in [\Theta^1, \Theta^2]$, then for all n , the law of e_n under \mathbb{P}_{e_0} is absolutely continuous with respect to the Lebesgue measure on $[\Theta^1, \Theta^2]$.

Proof: Let us start assuming (H3').

a) We first assume that $\mathcal{M}_a, \mathcal{M}_b \in \mathbb{R}^+$. We use the notation of Lemma 11. For $e_0 \in [0, 1]$ and n , we define $\theta_n^1(e_0) = \phi_{\mathcal{M}_a, \mathcal{M}_b}^n(e_0)$ and $\theta_n^2(e_0) = \phi_{\mathcal{M}_a, \mathcal{M}_b}^n(e_0)$, where ϕ^n is the n th iteration of the function ϕ . Note that for $e_1 \leq e_2 \in [0, 1]$, $\alpha_1 < \alpha_2 \in \mathbb{R}^+$ and $\beta_1 < \beta_2 \in \mathbb{R}^+$

$$\begin{aligned} \psi : [e_1, e_2] \times [\alpha_1, \alpha_2] \times [\beta_1, \beta_2] &\longrightarrow [\phi_{\alpha_1,\beta_2}(e_1), \phi_{\alpha_2,\beta_1}(e_2)] \\ (x, \alpha, \beta) &\longmapsto \phi_{\alpha,\beta}(x) \end{aligned}$$

is well defined and onto and the inverse image of an interval which is not a singleton has positive Lebesgue measure. Therefore, by induction, the Lebesgue measure on $[\theta_n^1, \theta_n^2]$ is absolutely continuous with respect to the law of e_n under \mathbb{P}_{e_0} . Moreover, by (H2) and (54), the Lebesgue measure on $[\theta_n^1, \theta_n^2]$ and the law of e_n under \mathbb{P}_{e_0} are mutually absolutely continuous.

b) It is a direct consequence of Lemma 11 and we get $\Theta^1 = \kappa_{\mathcal{M}_a, \mathcal{M}_b}$ and $\Theta^2 = \kappa_{\mathcal{M}_a, \mathcal{M}_b}$. By Lemma 11 and (H3'), $\kappa_{\mathcal{M}_a, \mathcal{M}_b} < \kappa_{\mathcal{M}_a, \mathcal{M}_b}$, hence, $\Theta^1 < \Theta^2$.

c) $\phi_{\mathcal{M}_a, \mathcal{M}_b}$ is increasing and $\kappa_{\mathcal{M}_a, \mathcal{M}_b}$ a fixed point hence if $\kappa_{\mathcal{M}_a, \mathcal{M}_b} \leq e_0$, then for all n , $\kappa_{\mathcal{M}_a, \mathcal{M}_b} \leq \theta_n^1(e_0)$. In the same way, for all n , $\kappa_{\mathcal{M}_a, \mathcal{M}_b} \geq \theta_n^2(e_0)$.

If $\mathcal{M}_a = \infty$ (resp., $\mathcal{M}_b = \infty$), we take for all $n \geq 1$, $\theta_n^2 = 1$ (resp., $\theta_n^1 = 0$) and $\Theta^2 = 1$ (resp., $\Theta^1 = 0$) and the proof is the same.

Let us now assume (H3). The proof is the same with for all $n \geq 1$ and all $e_0 \in [0, 1]$, $\theta_n^1(e_0) = 0$, for all $n \geq 1$ and all $e_0 \in [0, 1]$ (except for $n = 1$ and $e_0 = 0$), $\theta_n^2(e_0) = 0$. We get $\Theta^1 = 0$ and $\Theta^2 = 1$. \square

Denote by l the Lebesgue measure on $[\Theta^1, \Theta^2]$ and by \mathcal{B}^+ the set of Lebesgue measurable subsets of $[0, 1]$ with positive l -measure.

Lemma 13: Assume (H2) and [(H3) or (H3')]. For all $B \in \mathcal{B}^+$, there exists $n_0 = n_0(B)$ such that for all $n \geq n_0$

$$p_n \triangleq \inf_{e_0 \in [0, 1]} \mathbb{P}_{e_0}(e_n \in B) > 0.$$

Proof: The Lebesgue measure on $[\Theta^1, \Theta^2]$ is regular hence there exists an $\varepsilon > 0$ such that $B \cap [\Theta^1 + \varepsilon, \Theta^2 - \varepsilon]$ has positive Lebesgue measure. By Lemma 12 a) and b), we can take n_0 such that for any given $n \geq n_0$ and any given starting point e_0 , $\mathbb{P}_{e_0}(e_n \in B) > 0$. Fix $n \geq n_0$. Set $\psi(e_0) = \mathbb{P}_{e_0}(e_n \in B)$. By (H2), ψ is a continuous function on $[0, 1]$. By compactness

$$\inf_{e_0 \in [0, 1]} \mathbb{P}_{e_0}(e_n \in B) > 0. \quad \square$$

Lemma 6a) will be a direct consequence of the following proposition.

Proposition 14: Assume (H2) and [(H3) or (H3')]. The Markov chain e_n has a unique stationary probability, say, μ_δ and for $s \in \mathbb{L}^1(\mu_\delta)$, for every starting point $e_0 \in [0, 1]$, \mathbb{P}_{e_0} -a.s.

$$\frac{1}{n} \sum_{i=0}^n s(e_i) \xrightarrow{n \rightarrow \infty} \int s d\mu_\delta.$$

Proof: By [16, Theorem 17.0.1], it is enough to prove that the Markov chain e_n is l -irreducible, positive Harris with invariant probability μ_δ .

Step 1: The Markov chain e_n is l -irreducible, Harris and admits an invariant measure unique up to a constant multiple. By Lemma 13, for $e_0 \in [0, 1]$ and $B \in \mathcal{B}^+$, the chain has a positive probability to reach B in n_0 steps starting from e_0 . Therefore, the Markov chain e_n is l -irreducible and by Lemma 12 c), l is a maximal irreducibility measure for the chain e_n . For a given $B \in \mathcal{B}^+$, by Lemma 13, the chain e_n has a probability at least p_{n_0} to reach B in n_0 steps, hence the chain will eventually reach B and hence come back to B an infinite number of times, therefore, B is Harris-recurrent and the Markov chain e_n is Harris. By [16, Theorem 10.0.1], the Markov chain e_n admits an invariant measure unique up to a constant multiple.

Step 2: The Markov chain e_n is aperiodic. By [16, Theorem 5.4.4], there exists an integer d , the period of the chain, such that there exist disjoint measurable sets D_0, \dots, D_{d-1} such that

- for $i = 0 \dots d-1$, if $e_i \in D_i$, then $\mathbb{P}_{e_i}(e_{i+1} \in D_{i+1}) = 1 \pmod{d}$;
- $l((\cup_{i=1}^d D_i)^c) = 0$.

By Lemma 12, for $n_1 \geq n_0$ large enough and $n \geq n_1$, the Lebesgue measure on $J \triangleq [(2\Theta^1 + \Theta^2)/3, (\Theta^1 + 2\Theta^2)/3]$ is absolutely continuous with respect to the law of e_n under \mathbb{P}_{e_0} . Therefore, for any $n \geq n_1$, if $e_n \in D_i$, then $J \subset D_i$, and then, if $d > 1$, $e_{n+1} \in D_{i+1}$ and thus also $J \subset D_{i+1}$, a contradiction. Hence, $d = 1$.

Step 3: The set $[0, 1]$ is regular for the Markov chain e_n . Take $B \in \mathcal{B}^+$. By Lemma 13, the time it will take for the

chain e_n to enter B is a.s. bounded above by n_0 times a geometric random variable of parameter p_{n_0} , hence its expectation is bounded above by n_0/p_{n_0} , hence $[0, 1]$ is regular.

Now we apply [16, Theorem 13.0.1] and get that the Markov chain e_n is positive Harris, hence has a unique invariant probability that we denote μ_δ . \square

We continue with the Proof of Lemma 6a). Recall that $0 \leq e_n \leq 1$, hence μ_δ is stochastically dominated by an atom at 1. μ_δ is the invariant measure, since the function $\phi_{\alpha, \beta}(\cdot)$ is increasing in e , μ_δ is stochastically dominated by the law of the chain started at 1 after one step

$$\mu_\delta \preceq \mathcal{L} \left(\frac{\delta + |a_0|^2}{\delta + |b_0|^2 + |a_0|^2} \right) \preceq \mathcal{L} \left(\frac{|a_0|^2}{|b_0|^2 + |a_0|^2} \right).$$

Thus, denoting by π_0 the law of $\frac{|a_0|^2}{|b_0|^2 + |a_0|^2}$, and using (H1)

$$\int -\log(1-x) d\mu_\delta(x) \leq \int -\log(1-x) d\pi_0(x) < \infty.$$

That is

$$-\log(1-\cdot) \in \mathbb{L}^1(\mu_\delta). \quad (56)$$

With Proposition 14, we get

$$\begin{aligned} & \frac{1}{n} \sum_{k=2}^{n+1} -\log(1-e_k) \\ & \xrightarrow{n \rightarrow \infty} \int_0^1 -\log(1-x) d\mu_\delta(x), \quad \mathbb{P}_{e_2}\text{-a.s.} \end{aligned} \quad (57)$$

With (53), it gives a Proof of Lemma 6a).

2) *Proof of Point b)*: The proof is based on the study of the logarithm of e_n by a coupling with an auxiliary Markov chain: w_n . We first study the jumps of the logarithm of e_n (Lemma 15), and define the Markov chain w_n . Then we state Lemma 16 which we will need in order to prove that w_n is Harris with the right properties and Lemma 17, which we use in the last part of the proof. From the Harris character of w_n , we deduce ergodic properties (Proposition 18), which we use along with martingale arguments to conclude the Proof of Lemma 6b).

Take $\eta > 0$ and $\varepsilon > 0$ small.

$$\begin{aligned} \int_0^1 -\log(1-x) d\mu_\delta(x) &= \int_0^\varepsilon -\log(1-x) d\mu_\delta(x) \\ &+ \int_\varepsilon^{1-\eta} -\log(1-x) d\mu_\delta(x) \\ &+ \int_{1-\eta}^1 -\log(1-x) d\mu_\delta(x) \\ &\leq -\varepsilon \log(1-\varepsilon) - \log \eta \mu_\delta([\varepsilon, 1]) \\ &+ \int_{1-\eta}^1 -\log(1-x) d\mu_\delta(x). \end{aligned} \quad (58)$$

By (56), the last term converges to 0 as η goes to 0. By (53), (57), and (58), to prove Lemma 6b), we only have to prove that for any given $\varepsilon > 0$ small enough

$$\mu_\delta([\varepsilon, 1]) \xrightarrow{\delta \rightarrow 0} 0.$$

For that, by Proposition 14, we need to show that the proportion of the time that the chain e_n spends above ε converges to 0 as δ goes to 0. Since we are interested in ε small, we can take $0 < \varepsilon < \varepsilon_0 < 1$, where ε_0 will be chosen in the Proof of Lemma 17. We consider the Markov chain $z_n \triangleq \log e_n$ and the random function g_n such that $z_n = g_n(z_{n-1})$. It is enough to show that the proportion of the time that z_n spends above $\log \varepsilon$ goes to 0 as δ goes to 0. Let us couple z_n with another Markov chain w_n , such that $w_n \geq z_n$ a.s. and that the proportion of the time that w_n spends above $\log \varepsilon$ goes to 0 as δ goes to 0.

For that, we need good information on the jumps of z_n .

Lemma 15: Assume (H1) and (H5). Set

$$\begin{aligned} j_n(z_{n-1}) &\triangleq z_n - z_{n-1} \\ &= \log \left(\frac{\delta}{e^{z_{n-1}}} + |a_{n-1}|^2 \right) \\ &\quad - \log(\delta + |b_{n-1}|^2 + |a_{n-1}|^2 e^{z_{n-1}}). \end{aligned}$$

($\forall \delta > 0$) ($\exists \varepsilon' > 0$) ($\forall x \geq \log \varepsilon'$)

a) $\mathbb{E} j_n(x) \leq 0$,

b) $\text{Var } j_n(x) \leq V$ where

$$V \triangleq \mathbb{E}((\log(|a_{n-1}|^2 + |b_{n-1}|^2))^2 + (\log(|a_{n-1}|^2))^2) + C.$$

C is a constant independent of everything. ε' is a function of δ but we will not write it to keep the notation clear. Moreover

$$\lim_{\delta \rightarrow 0} \varepsilon' = 0.$$

Proof: Note that by (H1), $V < \infty$. $j_n(x)$ is a nonincreasing continuous function of x and so is $\mathbb{E} j_n(x)$. $\mathbb{E} j_n(0) < 0$, hence given δ , there exist $0 < \varepsilon' < 1$ such that $\mathbb{E} j_n(\log \varepsilon') \leq 0$, and for $x \geq \log \varepsilon'$, $\mathbb{E} j_n(x) \leq 0$. That gives point 1. For point 2, take C such that for all $x \geq 0$

$$(\log(x+1))^2 \leq (\log(x))^2 + C.$$

To prove that $\lim_{\delta \rightarrow 0} \varepsilon' = 0$, it is enough to prove that for a given $L < 0$, we can find δ small enough such that $\mathbb{E} j_n(L) \leq 0$. That is true because for a given L , $\mathbb{E} j_n(x)$ is a continuous function of δ which, by (H4) is negative for $\delta = 0$. \square

We continue with the Proof of Lemma 6b). We take $\delta > 0$ such that $0 < \varepsilon' < \varepsilon < \varepsilon_0 < 1$. We define w_n in a way that it stays between $\log \varepsilon'$ and 0. Set $w_0 = z_0$, for δ small enough, $w_0 > \log \varepsilon'$. For $x \in [\log \varepsilon'; 0]$, denote

$$h_n(x) = g_n(x) - \mathbb{E} j_n(x) \geq g_n(x).$$

That is

$$\begin{aligned} h_n(x) &= x + \log \left(\frac{\frac{\delta}{e^x} + |a_{n-1}|^2}{\delta + |b_{n-1}|^2 + |a_{n-1}|^2 e^x} \right) \\ &\quad - \mathbb{E} \log \left(\frac{\frac{\delta}{e^x} + |a_{n-1}|^2}{\delta + |b_{n-1}|^2 + |a_{n-1}|^2 e^x} \right). \end{aligned} \quad (59)$$

Note that

$$\mathbb{E}(h_n(z_{n-1}) - z_{n-1} | z_{n-1}) = 0. \quad (60)$$

- If $h_n(w_{n-1}) > 0$, set $w_n = 0$.
- If $h_n(w_{n-1}) < \log \varepsilon'$, set $w_n = \log \varepsilon'$.
- Otherwise, set $w_n = h_n(w_{n-1})$.

In the first two cases, we say that the chain is *truncated*. Note that for all n , $w_n \geq z_n$. Indeed, either $w_n = 0 \geq z_n$ or $w_n \geq h_n(w_{n-1}) \geq g_n(w_{n-1}) \geq g_n(z_{n-1}) = z_n$ by induction, and using the fact that g_n is a.s. nondecreasing. Therefore, the proportion of the time that the chain w_n spends above $\log \varepsilon$ is larger than the proportion of the time that chain z_n spends above $\log \varepsilon$.

Let us give the intuition for the end of the proof. w_n lives on $[\log \varepsilon', 0]$ and has drift 0 (except at the extremities). Moreover, as δ goes to 0, $\log \varepsilon'$ goes to $-\infty$. Consider now the discrete symmetric random walk on $\{\dots, -2, -1, 0\}$; the proportion of the time it spends on a finite set of the form $\{-N, \dots, -2, -1, 0\}$ is 0. Similarly, as δ goes to 0, w_n spends less and less time above $\log \varepsilon$.

Let us prove two lemmas, which will be used in the proof.

Lemma 16: Assume (H2).

- a) There exist $c > 0$ and $\theta > 0$ such that for all $x \in [\log \varepsilon'; 0]$

$$\mathbb{P}(h_n(x) \geq x + c) > \theta.$$

- b) Set $N = \lceil \frac{-\log \varepsilon'}{c} \rceil$. 0 is a recurrent point for the chain w_n and the time between two visits at 0 is a.s. bounded above by N times a geometric random variable of parameter θ^N .

Proof: a) We consider here $\delta \geq 0$. We denote by $\text{Supp}(X)$ the support of the law of a random variable X . We take δ_0 small enough. We consider for $x \in [\log \varepsilon'; 0]$ the function

$$\phi(x) = \max\{y; y \in \text{Supp}(h_n(x) - x)\}$$

which by (H2) and (59) is a continuous function of x . Moreover, since $\mathbb{E}(h_n(x) - x) = 0$, ϕ is strictly positive. By compactness, there exists $c > 0$ such that for $x \in [\log \varepsilon'; 0]$

$$\phi(x) > 2c$$

$$\mathbb{P}(h_n(x) \geq x + c) > 0.$$

By (H2) and (59), $\mathbb{P}(h_n(x) \geq x + c)$ is continuous and once again, by compactness, there exists $\theta > 0$ such that for $x \in [\log \varepsilon'; 0]$

$$\mathbb{P}(h_n(x) \geq x + c) > \theta.$$

- b) If there are at least N steps in a row such that $h_n(w_{n-1}) \geq x + c$, then the chain reaches 0. By the point a), that happens with probability at least $\theta^N > 0$, hence 0 is a recurrent point for the chain w_n and the time between two visits at 0 is a.s. bounded above by N times a geometric random variable of parameter θ^N . \square

Lemma 17: Assume (H2).

- a) There exist $u > 0$ and $\alpha > 0$ dependent on ε and independent of δ such that for all $x \in [2 \log \varepsilon; 0]$

$$\mathbb{P}(h_n(x) \geq x + u) > \alpha.$$

- b) There exists $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$, there exist $v > 0$ and $\beta > 0$ dependent on ε and independent of δ such that

$$\mathbb{P}(\log \varepsilon < h_1(0) < -v) > \beta.$$

Proof: We consider here $\delta \geq 0$. We denote by $\text{Supp}(X)$ the support of the law of a random variable X . We take δ_0 small enough.

- a) We consider for $x \in [2 \log \varepsilon; 0]$ and $0 \leq \delta \leq \delta_0$ the function

$$\phi(x, \delta) = \max\{y; y \in \text{Supp}(h_n(x) - x)\}$$

which by (H2) is a continuous function of (x, δ) because $(h_n(x) - x)$ is continuous in (x, δ) . Moreover, since $\mathbb{E}(h_n(x) - x) = 0$, ϕ is strictly positive. By compactness, there exists $u > 0$ such that for $x \in [2 \log \varepsilon; 0]$ and $0 \leq \delta \leq \delta_0$

$$\phi(x, \delta) > 2u$$

$$\mathbb{P}(h_n(x) \geq x + u) > 0.$$

By (H2), $\mathbb{P}(h_n(x) \geq x + u)$ is continuous and once again, by compactness, there exists $\alpha > 0$ such that for $x \in [2 \log \varepsilon; 0]$ and $0 \leq \delta \leq \delta_0$

$$\mathbb{P}(h_n(x) \geq x + u) > \alpha.$$

- b) For all $0 \leq \delta \leq \delta_0$, there exist $\varepsilon_0 > 0$ and $v > 0$ such that $\mathbb{P}(\log \varepsilon_0 < h_1(0) < -v) > 0$. Like in the proof of point a), by (H2), we can chose $\varepsilon_0 > 0$ and $v > 0$ continuous in δ . By compactness, we can chose $\varepsilon_0 > 0$ and $v > 0$ independent of δ such that for all $0 \leq \delta \leq \delta_0$, $\mathbb{P}(\log \varepsilon_0 < h_1(0) < -v) > 0$ and like in the proof of point a), by (H2), that probability can be chosen continuous in δ . Therefore, by compactness again, there exists $\beta > 0$ dependent on ε and independent of δ such that $\mathbb{P}(\log \varepsilon_0 < h_1(0) < -v) > \beta$. Take $\varepsilon < \varepsilon_0$, we have $\mathbb{P}(\log \varepsilon < h_1(0) < -v) > \beta$. \square

The following proposition states the properties of w_n which will be used in the Proof of Lemma 6b).

Proposition 18: Assume (H2).

- a) The Markov chain w_n has a unique stationary probability, say, ν_δ and for $s \in L^1(\nu_\delta)$, for every starting point $w_0 \in [\log \varepsilon', 0]$, \mathbb{P}_{w_0} -a.s.

$$\frac{1}{n} \sum_{i=0}^n s(w_i) \xrightarrow{n \rightarrow \infty} \int s d\nu_\delta.$$

- b) We denote T the return time to 0, starting from 0. Then $\nu_\delta(0) = 1/\mathbb{E}_0 T$.

Proof: See [16] and Definitions 7–10 for the theory of Harris Markov chains that we will use extensively in the proof. Define the following probability measure on $[\log \varepsilon', 0]$. For B a Borel set

$$\bar{l}(B) \triangleq \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \mathbb{P}_0(w_n \in B).$$

Let us prove that the Markov chain w_n is \bar{l} -irreducible, positive Harris with invariant probability ν_δ . By [16, Theorem 17.0.1], that will prove a).

The sets which have positive \bar{l} -measure are exactly the sets that have a positive probability to be visited starting from 0. Moreover, 0 is a recurrent point. Therefore, the Markov chain w_n is \bar{l} -irreducible and \bar{l} is a maximal irreducibility measure.

Moreover, take B with positive \bar{l} -measure, B is uniformly accessible from $\{0\}$. Therefore, we can apply [16, Theorem 9.1.3 (i)] and since 0 is Harris-recurrent, B is also Harris-recurrent, therefore, the chain w_n is Harris-recurrent. By Lemma 16 b), the time between two visits at 0 has finite expectation (bounded above by N/θ^n). Therefore, by [16, Theorem 10.2.2], the chain w_n is positive-Harris and admits a unique invariant probability measure. That finishes the proof of point a). The point b) is a consequence of

$$1 = \nu_\delta([\log \varepsilon', 0]) = \nu_\delta(0)\mathbb{E}_0[T]$$

which comes from [16, Theorem 10.0.1], which we apply to $A = \{0\}$, which has positive \bar{l} -measure. \square

We continue with the Proof of Lemma 6b). By Proposition 18a), to prove that the proportion of the time that w_n spends above $\log \varepsilon$ goes to 0 as δ goes to 0, we only need to prove that

$$\nu_\delta([\log \varepsilon, 0]) \xrightarrow{\delta \rightarrow 0} 0.$$

Let us first prove that $\mathbb{E}_0 T \xrightarrow{\delta \rightarrow 0} \infty$, which by Proposition 18 b) will prove that

$$\nu_\delta(0) \xrightarrow{\delta \rightarrow 0} 0.$$

We denote \mathcal{A} as the event $\log \varepsilon < h_1(0) < -v$. On \mathcal{A} , we define the stopping time

$$\tilde{T} = 1 + \inf\{n \geq 1; h_{n+1}(w_n) > 0 \text{ or } h_{n+1}(w_n) < \log \varepsilon'\}.$$

We now condition on the event \mathcal{A} and on $x_0 = h_1(0)$, denote by $\tilde{\mathbb{P}}$ and $\tilde{\mathbb{E}}$ the associated probability and expectation. $\tilde{T} \leq T$ is the first time the chain is truncated. Moreover, for $n < \tilde{T}$, $w_n = h_n(w_{n-1})$, so with (60), by classical martingale arguments,

$$\tilde{\mathbb{E}}(h_{\tilde{T}}(w_{\tilde{T}-1})) = x_0.$$

We denote by \mathcal{A}_0 the event that w_n reaches $\log \varepsilon'$ before 0, we set $p = \tilde{\mathbb{P}}(\mathcal{A}_0)$, $X_0 = \tilde{\mathbb{E}}(h_{\tilde{T}}(w_{\tilde{T}-1}) | \mathcal{A}_0^c)$ and $X_1 = \tilde{\mathbb{E}}(h_{\tilde{T}}(w_{\tilde{T}-1}) | \mathcal{A}_0)$.

$$x_0 = pX_1 + (1-p)X_0$$

$$p = \frac{X_0 - x_0}{X_0 - X_1}.$$

$X_0 \geq 0$ and $X_1 \leq \log \varepsilon' \leq \log \varepsilon < x_0 < -v$ hence

$$p \geq \frac{-x_0}{-X_1} \geq \frac{v}{-X_1}. \quad (61)$$

Using $X_1 \leq \log \varepsilon'$, (61), and $w_n^2 - Vn$, which is a super-martingale by Lemma 15b)

$$\begin{aligned} \tilde{\mathbb{E}}(T) &\geq \tilde{\mathbb{E}}(\tilde{T}) \geq \frac{\tilde{\mathbb{E}}(h_{\tilde{T}}(w_{\tilde{T}-1}))^2 - x_0^2}{V} \\ &\geq \frac{pX_1^2 + (1-p)X_0^2 - x_0^2}{V} \\ &\geq \frac{pX_1^2 - x_0^2}{V} \\ &\geq \frac{v(-X_1) - x_0^2}{V} \\ &\geq \frac{v(-\log \varepsilon') - x_0^2}{V}. \end{aligned}$$

We integrate over x_0 and use $\mathbb{P}(\mathcal{A}) > \beta$ (Lemma 17b) and $\mathbb{E}(h_1(0)^2 | \mathcal{A}) < (\log \varepsilon)^2$.

$$\begin{aligned} \mathbb{E}(T | \mathcal{A}) &\geq \frac{v(-\log \varepsilon') - \mathbb{E}(h_1(0)^2 | \mathcal{A})}{V} \\ \mathbb{E}(T) &\geq \beta \frac{v(-\log \varepsilon') - (\log \varepsilon)^2}{V}. \end{aligned}$$

We have proved that $\mathbb{E}T \xrightarrow{\delta \rightarrow 0} \infty$, which proves that $\nu_\delta(0) \xrightarrow{\delta \rightarrow 0} 0$.

Using Lemma 17a) and the invariance of ν_δ , let us prove by induction that for $\mathcal{N} \leq \lceil \frac{-\log \varepsilon}{u} \rceil$

$$\begin{aligned} \nu_\delta([-\mathcal{N}u; 0]) &\leq \alpha^{-\mathcal{N}} \nu_\delta(0) \\ \nu_\delta([-(\mathcal{N}-1)u; 0]) &\geq \int \nu_\delta(dw_0) \mathbb{P}_{w_0}(w_1 \in [-(\mathcal{N}-1)u; 0]) \\ &\geq \int_{[-\mathcal{N}u; 0]} \nu_\delta(dw_0) \mathbb{P}_{w_0}(w_1 \in [-(\mathcal{N}-1)u; 0]) \\ &\geq \int_{[-\mathcal{N}u; 0]} \nu_\delta(dw_0) \mathbb{P}_{w_0}(h_1(w_0) \geq u + w_0) \\ &\geq \alpha \nu_\delta([-\mathcal{N}u; 0]). \end{aligned}$$

Therefore

$$\nu_\delta([\log \varepsilon; 0]) \leq \alpha^{\lceil \frac{-\log \varepsilon}{u} \rceil} \nu_\delta(0).$$

Thus

$$\nu_\delta([\log \varepsilon; 0]) \xrightarrow{\delta \rightarrow 0} 0.$$

That concludes the Proof of Lemma 6b).

C. Proof of Theorem 2

We reformulate the problem in the spirit of Appendix A. Let $K > 1$. The \mathbf{a}_i (resp., \mathbf{b}_i) are now independent complex vectors of size K whose coefficients are independent and distributed according to π_a (resp., π_b). We denote by \mathbb{P} the probability associated with those random sequences and by \mathbb{E} the associated expectation. We consider the following $M \times K(M+1)$ channel transfer matrix:

$$\mathbf{H}_M = \begin{pmatrix} \mathbf{a}_1 & \mathbf{b}_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathbf{a}_M & \mathbf{b}_M \end{pmatrix}.$$

We consider the following random variable

$$\mathcal{C}_M(P) = \frac{1}{M} \text{tr} \left\{ \log \left(I + \frac{P}{K} \mathbf{H}_M \mathbf{H}_M^\dagger \right) \right\}$$

where $P = K\rho$, and $\mathbf{H}_M \mathbf{H}_M^\dagger$ is explicitly given by (62) at the top of the following page, with

$$\begin{aligned} |\mathbf{a}_i|^2 &= \sum_{k=1}^K |a_{i,k}|^2 \\ \langle \mathbf{a}_i, \mathbf{b}_j \rangle &= \sum_{k=1}^K (a_{i,k})^\dagger b_{j,k}. \end{aligned}$$

Proposition 19: Assume (H1), (H2), and (H4)

$$H_M H_M^\dagger = \begin{pmatrix} |\mathbf{a}_1|^2 + |\mathbf{b}_1|^2 & \langle \mathbf{a}_2; \mathbf{b}_1 \rangle & 0 & \cdots & 0 \\ \langle \mathbf{a}_2; \mathbf{b}_1 \rangle^\dagger & |\mathbf{a}_2|^2 + |\mathbf{b}_2|^2 & \langle \mathbf{a}_3; \mathbf{b}_2 \rangle & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \langle \mathbf{a}_M; \mathbf{b}_{M-1} \rangle \\ 0 & \cdots & 0 & \langle \mathbf{a}_M; \mathbf{b}_{M-1} \rangle^\dagger & |\mathbf{a}_M|^2 + |\mathbf{b}_M|^2 \end{pmatrix} \quad (62)$$

- a) For every $\rho > 0$, $\mathcal{C}_M(P)$ converges \mathbb{P} -a.s. as M goes to infinity. We call the limit $\mathcal{C}(P)$.
b) As P goes to infinity

$$\mathcal{C}(P) = \log P + \mathbb{E} \log \left(\frac{e + |\mathbf{b}|^2}{K} \right) + o(1)$$

where the expectation is taken in the following way. e and \mathbf{b} are independent. \mathbf{b} is a complex K -vector whose coefficients are independent and distributed according to π_b . The law of e is m_0 , which is the unique invariant probability of the Markov chain defined by

$$e_{n+1} = |\mathbf{a}_n|^2 \left(\frac{e_n + |\mathbf{b}_{n-1}|^2 \sin^2(\mathbf{a}_n, \mathbf{b}_{n-1})}{e_n + |\mathbf{b}_{n-1}|^2} \right).$$

The rest of this Appendix is devoted to the Proof of Proposition 19.

As in Appendix A, we define the sequence x_n as follows. $x_0 = 0, x_1 = 1$, and for $n \geq 1$

$$x_{n+1} = \frac{\lambda - |\mathbf{a}_n|^2 - |\mathbf{b}_n|^2}{\langle \mathbf{a}_{n+1}; \mathbf{b}_n \rangle} x_n - \frac{\langle \mathbf{a}_n; \mathbf{b}_{n-1} \rangle^\dagger}{\langle \mathbf{a}_{n+1}; \mathbf{b}_n \rangle} x_{n-1}. \quad (63)$$

We get, like in (50), for $\lambda = -1/\rho$

$$\begin{aligned} \mathcal{C}_M(P) &= \log(P/K) + \frac{1}{M} \log |x_{M+1}(\lambda)| \\ &\quad + \frac{1}{M} \sum_{i=1}^M \log |\langle \mathbf{a}_{i+1}; \mathbf{b}_i \rangle|, \quad \mathbb{P}\text{-a.s.} \end{aligned} \quad (64)$$

Set $c_n \triangleq x_n/x_{n-1}$, for $n \geq 2$. By (63), we get

$$c_{n+1} = \frac{\lambda - |\mathbf{a}_n|^2 - |\mathbf{b}_n|^2}{\langle \mathbf{a}_{n+1}; \mathbf{b}_n \rangle} - \frac{\langle \mathbf{a}_n; \mathbf{b}_{n-1} \rangle^\dagger}{c_n \langle \mathbf{a}_{n+1}; \mathbf{b}_n \rangle}.$$

Let $d_n = c_n \langle \mathbf{a}_n; \mathbf{b}_{n-1} \rangle$. Then

$$\begin{aligned} d_{n+1} &= \lambda - |\mathbf{a}_n|^2 - |\mathbf{b}_n|^2 - \frac{|\langle \mathbf{a}_n; \mathbf{b}_{n-1} \rangle|^2}{d_n} \\ &= \lambda - |\mathbf{b}_n|^2 - |\mathbf{a}_n|^2 \left(1 + \frac{|\mathbf{b}_{n-1}|^2 \cos^2(\mathbf{a}_n, \mathbf{b}_{n-1})}{d_n} \right) \end{aligned}$$

where

$$\cos^2(\mathbf{a}_n, \mathbf{b}_{n-1}) \triangleq |\langle \mathbf{a}_n; \mathbf{b}_{n-1} \rangle|^2 / |\mathbf{a}_n|^2 |\mathbf{b}_{n-1}|^2.$$

Note that $0 \leq \cos^2 \leq 1$. Let $e_n = -d_n - |\mathbf{b}_{n-1}|^2$.

$$e_{n+1} = -\lambda + |\mathbf{a}_n|^2 \left(\frac{e_n + |\mathbf{b}_{n-1}|^2 \sin^2(\mathbf{a}_n, \mathbf{b}_{n-1})}{e_n + |\mathbf{b}_{n-1}|^2} \right) \quad (65)$$

where $\sin^2 \triangleq 1 - \cos^2$. With the initial conditions, $d_2 < -|\mathbf{b}_1|^2$, hence $e_2 > 0$ and for all $n, e_n > 0$. Note that (e_n) is a Markov chain and that for all n , e_n is independent of \mathbf{a}_n and \mathbf{b}_{n-1} . By (64), we get

$$\begin{aligned} \mathcal{C}_M(P) &= \log(P/K) + \frac{1}{M} \sum_{i=2}^{M+1} \log |c_i(\lambda)| \\ &\quad + \frac{1}{M} \sum_{i=1}^M \log |\langle \mathbf{a}_{i+1}; \mathbf{b}_i \rangle| \\ &= \log(P/K) + \frac{1}{M} \sum_{i=2}^{M+1} \log(|d_i|) + o(1) \\ &= \log(P) + \frac{1}{M} \sum_{i=2}^{M+1} \log \left(\frac{e_i(\lambda) + |\mathbf{b}_{i-1}|^2}{K} \right) + o(1). \end{aligned} \quad (66)$$

We only need to study the Markov chain (e_n, \mathbf{b}_{n-1}) . For convenience, we set $\delta = -\lambda$ and we allow $\delta = 0$. We also assume without loss of generality that the chain starts at (e_1, \mathbf{b}_0) .

Proposition 20: Assume (H2) and (H4). Take $\delta \geq 0$. The Markov chain $(e_n(\delta), \mathbf{b}_{n-1})$ has a unique stationary probability, say, μ_δ , and for $s \in L^1(\mu_\delta)$, for every starting point $(e_1, \mathbf{b}_0) \in \mathbb{R}_+ \times \mathbb{C}^K$, $\mathbb{P}_{(e_1, \mathbf{b}_0)}$ -a.s.

$$\frac{1}{n} \sum_{i=0}^n s(e_i, \mathbf{b}_{i-1}) \xrightarrow[n \rightarrow \infty]{} \int s d\mu_\delta.$$

Moreover, μ_δ is weakly continuous in $\delta = 0$.

Proof: We consider the Markov chain (e_n) on the compact $[0, \infty]$. By (65), for $n \geq 1$ and $e \in [0, \infty]$, $\mathbb{P}_e(e_n = \infty) = 0$. Consider (65), by (H2), for $e_1 \in [0, \infty]$, the law of e_2 under P_{e_1} is absolutely continuous with respect to the Lebesgue measure on $[\delta, \infty]$. Moreover, by (H4), the law of e_2 under P_{e_1} and the Lebesgue measure on $[\delta, \infty]$ are mutually absolutely continuous. Therefore, for $e_1 \in [0, \infty)$ and $n \geq 3$, the law of e_n under P_{e_1} and the Lebesgue measure on $[\delta, \infty]$ are mutually absolutely continuous. That fact allows us to prove, as in Appendix B, that the Markov chain (e_n) is l -irreducible, positive Harris with invariant probability m_δ , where l is the Lebesgue measure on $[\delta, \infty]$. Since $\mathbb{P}_e(e_n = \infty) = 0$, m_δ does not charge $\{\infty\}$. We identify m_δ and the measure it induces on \mathbb{R}_+ . We denote by $\Pi_{\mathbf{b}}$ the law of \mathbf{b} . Since for $n \geq 1$, e_n and \mathbf{b}_{n-1} are independent, the Markov chain (e_n, \mathbf{b}_{n-1}) is $l \times \Pi_{\mathbf{b}}$ -irreducible, positive Harris with invariant probability $\mu_\delta = m_\delta \times \Pi_{\mathbf{b}}$. By [16, Theorem 17.0.1], the Markov chain $(e_n(\lambda), \mathbf{b}_{n-1})$ has a unique

stationary probability μ_δ and for $s \in L^1(\mu_\delta)$, for every starting point $(e_1, \mathbf{b}_0) \in \mathbb{R}_+ \times \mathbb{C}^K$, $\mathbb{P}_{(e_1, \mathbf{b}_0)}$ -a.s.

$$\frac{1}{n} \sum_{i=0}^n s(e_i, \mathbf{b}_{i-1}) \xrightarrow{n \rightarrow \infty} \int s d\mu_\delta.$$

Let us prove that m_δ converges weakly to m_0 when δ converges to 0, which will finish the proof. $\{m_\delta, \delta \geq 0\}$ are measures on the compact $[0, \infty]$ hence it is enough to show that m_0 is the only limit point when δ goes to 0. By (H2), for a point x and an interval A in $[0, \infty]$, $\mathbb{P}_{e_1}(e_2(\delta) \in A)$ converges to $\mathbb{P}_{e_1}(e_2(0) \in A)$. It implies that a limit point must be an invariant measure for the chain with $\delta = 0$. The only possibility is m_0 . \square

By (65), m_δ is stochastically dominated by the law of $|\mathbf{a}_n|^2 + \delta$. Therefore, by (H1), $(x, y) \rightarrow \log(x + y) \in L^1(\mu_\delta)$. Equation (66) and Proposition 20 conclude the Proof of Proposition 19.

D. Product of Random Matrices

We prove Lemma 6a) assuming only (H1) and (H2). We use the theory of product of random matrices theory. For a general introduction to the aspects of the theory we use here, the reader may consult [26], [27], [36]–[38].

Let us take $|\cdot|$ any norm on \mathbb{C}^2 and $\|\cdot\|$ the associated operator norm on matrices. For a given λ

$$\begin{pmatrix} x_{n+1} \\ x_n \end{pmatrix} = \begin{pmatrix} \frac{\lambda - |a_n|^2 - |b_n|^2}{a_{n+1}^\dagger b_n} & -\frac{a_n b_{n-1}^\dagger}{a_{n+1}^\dagger b_n} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_n \\ x_{n-1} \end{pmatrix}.$$

For $a, a', b, b' \in \mathbb{C} - 0$, we define the following invertible matrices:

$$\mathbf{g}(\lambda, a, a', b, b') \triangleq \begin{pmatrix} \frac{\lambda - |a|^2 - |b|^2}{a'^\dagger b'} & -\frac{ab^\dagger}{a'^\dagger b'} \\ 1 & 0 \end{pmatrix}.$$

Finally, we define

$$\begin{aligned} \mathbf{g}_n(\lambda) &\triangleq \mathbf{g}(\lambda, a_n, a_{n-1}, b_{n-1}, b_n) \\ &= \begin{pmatrix} \frac{\lambda - |a_n|^2 - |b_n|^2}{a_{n+1}^\dagger b_n} & -\frac{a_n b_{n-1}^\dagger}{a_{n+1}^\dagger b_n} \\ 1 & 0 \end{pmatrix} \\ \mathbf{M}_n &\triangleq \mathbf{g}_n \cdots \mathbf{g}_1. \end{aligned} \quad (67)$$

So that

$$\begin{pmatrix} x_{n+1} \\ x_n \end{pmatrix} = \mathbf{M}_n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Set $\mathcal{E} = (\mathbb{C} - 0)^4$ which is a Borel set of a separable and complete metric space. $\mathbf{X}_n \triangleq (a_{n+1}, a_n, b_n, b_{n-1})$ is a Markov chain on \mathcal{E} , with invariant measure $\Pi \triangleq \pi_a \times \pi_a \times \pi_b \times \pi_b$. With (H1)

$$\mathbb{E}_\Pi(\log^+ \|\mathbf{g}(\lambda, a, a', b, b')\| + \log^+ \|\mathbf{g}(\lambda, a, a', b, b')^{-1}\|) < \infty. \quad (68)$$

Notice that $\mathbf{g}_n(\lambda)$ is a continuous function of \mathbf{X}_n , therefore $((\mathbf{X}_n, \mathbf{M}_n), \Pi)$ is a multiplicative Markovian process. By [39,

Example 1 and Proposition 2.5], $1/n \log \|\mathbf{M}_n(\lambda)\|$ converges \mathbb{P} -a.s. and in $\mathbb{L}_1(\Omega)$, we set

$$\gamma(\lambda) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|\mathbf{M}_n(\lambda)\|. \quad (69)$$

$\gamma(\lambda)$ is the first Lyapunov exponent.

The $\mathbb{L}_1(\Omega)$ convergence already gives an easy upper bound for $\gamma(\lambda)$. By the property of operator norm

$$\gamma(\lambda) \leq \mathbb{E}_\Pi \log \|\mathbf{g}_1(\lambda)\|.$$

Moreover, we can refine that bound into a whole family of upper bounds, for $k \in \mathbb{N}$

$$\gamma(\lambda) \leq \frac{1}{k} \mathbb{E}_\Pi \log \|\mathbf{g}_1(\lambda) \cdots \mathbf{g}_k(\lambda)\|. \quad (70)$$

Note that this upper bound is getting better as k increases and tight as $k \rightarrow \infty$.

Let us now prove that

$$\frac{1}{n} \log |x_{n+1}(\lambda)| \xrightarrow{n \rightarrow \infty} \gamma(\lambda).$$

Definition 21: The multiplicative system $((\mathbf{X}_n, \mathbf{M}_n), \Pi)$ is irreducible if there is no measurable nonrandom family $\{V(\mathbf{X}), \mathbf{X} \in E\}$ of proper subspaces of \mathbb{C}^2 such that

$$\mathbf{M}_n V(\mathbf{X}_0) = V(\mathbf{X}_n), \quad \mathbb{P}\text{-a.s.}, \quad \forall n \in \mathbb{N}.$$

Lemma 22: Assume (H2). The multiplicative system $((\mathbf{X}_n, \mathbf{M}_n), \Pi)$ is irreducible

The proof is an adaptation of the proof of [40, Proposition 6.1.1].

Proof: The proof is by contradiction. Assume that there is a measurable family $\{V(\mathbf{X}), \mathbf{X} \in E\}$ of proper subspaces of \mathbb{C}^2 such that

$$\mathbf{g}_3 V(\mathbf{X}_2) = V(\mathbf{X}_3), \quad \mathbb{P}\text{-a.s.}, \quad \forall n \in \mathbb{N}.$$

We parameterize the proper subspaces of \mathbb{C}^2 by $\binom{c}{1}$ for c in $(-\infty, \infty]$. There is a measurable family $\{c(\mathbf{X}), \mathbf{X} \in E\}$ such that $\mathbf{g}_3 \binom{c(\mathbf{X}_2)}{1}$ and $\binom{c(\mathbf{X}_3)}{1}$ are \mathbb{P} -a.s. collinear. A direct computation gives

$$c(a_4, a_3, b_3, b_2) = \frac{\lambda - |a_3|^2 - |b_3|^2}{a_4^\dagger b_3} - \frac{a_3 b_2^\dagger}{c(a_3, a_2, b_2, b_1) a_4^\dagger b_3}, \quad \mathbb{P}\text{-a.s.}$$

that is

$$c(a_3, a_2, b_2, b_1) = \frac{a_3 b_2^\dagger}{a_4^\dagger b_3 \left(\frac{\lambda - |a_3|^2 - |b_3|^2}{a_4^\dagger b_3} - c(a_4, a_3, b_3, b_2) \right)}, \quad \mathbb{P}\text{-a.s.}$$

Note that the right-hand side (RHS) does not depend on a_2 and b_1 , hence, $c(a, a', b, b')$ does not depend on a' and b' . Setting $d(a, b) = a^\dagger b c(a, b)$, we get

$$d(a_4, b_3) = \lambda - |a_3|^2 - |b_3|^2 - \frac{|a_3|^2 |b_2|^2}{d(a_3, b_2)}, \quad \mathbb{P}\text{-a.s.} \quad (71)$$

The RHS does not depend on a_4 , hence, $d(a, b)$ does not depend on a . From (71), we get

$$\frac{d(b_2)}{|b_2|^2} = -\frac{|a_3|^2}{d(b_3) - \lambda + |a_3|^2 + |b_3|^2}, \quad \mathbb{P}_{\pi}\text{-a.s.}$$

The RHS does not depend on b_2 , hence, $d(b)/|b|^2$ does not depend on b , set $d(b) = L|b|^2$, where L is a fixed constant. Then

$$(L + 1)|b_3|^2 = \lambda - |a_3|^2 \left(1 + \frac{1}{L}\right), \quad \mathbb{P}_{\pi}\text{-a.s.}$$

If $L \neq -1$, $|b_3|^2$ is a measurable function of a_3 and since it is also independent of a_3 , it is a constant, which is in contradiction with (H2). Hence, $L = -1$, which gives a contradiction with $\lambda < 0$. \square

By [39, Lemma 2.6], irreducibility implies that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left| \binom{x_{n+2}}{x_{n+1}} \right| = \gamma; \quad \mathbb{P}\text{-a.s.}$$

The following lemma completes the proof.

Lemma 23: Assume (H1).

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(\log \left| \binom{x_{n+2}}{x_{n+1}} \right| - \log |x_{n+1}| \right) = 0, \quad \mathbb{P}\text{-a.s.}$$

Proof:

$$\log \left| \binom{x_{n+2}}{x_{n+1}} \right| - \log |x_{n+1}| = \log \left| \binom{c_{n+2}}{1} \right| \geq 0.$$

Let us prove that for $\varepsilon > 0$, $\mathbb{P}(\frac{1}{n} \log |c_n| \geq \varepsilon)$ is a summable series, which by the Borel–Cantelli lemma will prove the lemma. We have

$$\begin{aligned} & \mathbb{P} \left(\frac{1}{n} \log |c_n| \geq \varepsilon \right) \\ & \leq \mathbb{P} \left(\frac{1}{n} \log (|c_n| + 1) \geq \varepsilon \right) \\ & \leq \mathbb{P} (|c_n| \geq e^{n\varepsilon} - 1) \leq \mathbb{P} (|c_n| \geq e^{\frac{n\varepsilon}{2}}) \\ & \leq \mathbb{P} \left(\frac{|b_{n-1}|}{|a_n|} \frac{1}{1 - e_n} \geq e^{\frac{n\varepsilon}{2}} \right) \\ & \leq \mathbb{P} \left(\frac{|b_{n-1}|}{|a_n|} \geq e^{\frac{n\varepsilon}{4}} \right) + \mathbb{P} \left(\frac{1}{1 - e_n} \geq e^{\frac{n\varepsilon}{4}} \right). \quad (72) \end{aligned}$$

We analyze the right side of (72). We use the fact that $\log |a_n|$ and $\log |b_{n-1}|$ have a second moment by (H1) and that it does not depend on n . By the Bienaymé–Chebishev inequality, we get

$$\begin{aligned} \mathbb{P} \left(\frac{|b_{n-1}|}{|a_n|} \geq e^{\frac{n\varepsilon}{4}} \right) &= \mathbb{P} \left(\log |b_{n-1}| - \log |a_n| \geq \frac{n\varepsilon}{4} \right) \\ &\leq \frac{16\mathbb{E}((\log |b_{n-1}| - \log |a_n|)^2)}{n^2\varepsilon^2} \quad (73) \end{aligned}$$

implying that the first term in the right side of (72) forms a summable series. Moreover

$$\log \frac{1}{1 - e_n} \leq \log \frac{-\lambda + |b_{n-1}|^2 + |a_{n-1}|^2}{|b_{n-1}|^2}$$

which has a second moment by (H1), hence, by a computation like (73) and the Bienaymé–Chebishev inequality, $\mathbb{P}(\frac{1}{1 - e_n} \geq e^{\frac{n\varepsilon}{4}})$ is a summable series. The Borel–Cantelli lemma applied to the right side of (72) concludes the proof. \square

E. Determinants of Jacobi Matrices

An interesting and useful characterization of an $M \times M$ Jacobi matrix is that its determinant can be expressed by the following recursive formula [41]:

$$\det \mathbf{G}_m = [\mathbf{G}_m]_{m,m} \det \mathbf{G}_{m-1} - [\mathbf{G}_m]_{m,m-1} [\mathbf{G}_m]_{m-1,m} \det \mathbf{G}_{m-2} \quad (74)$$

for $m = 3, \dots, M$, with

$$\begin{aligned} \det \mathbf{G}_1 &= [\mathbf{G}_m]_{1,1} \\ \det \mathbf{G}_2 &= [\mathbf{G}_m]_{1,1} [\mathbf{G}_m]_{2,2} - [\mathbf{G}_m]_{1,2} [\mathbf{G}_m]_{2,1} \quad (75) \end{aligned}$$

where \mathbf{G}_m is the principle submatrix of \mathbf{G}_M , obtained by deleting its last $(M - m)$ columns. This characterization already used by Narula [14], can be easily proved by straight forward calculation of the determinant of \mathbf{G}_M , starting from its last row.

Examining (74), it is observed that the determinant of a square Jacobi matrix is dependent on a weighted sum of its two largest principle matrices' determinants only. Furthermore, $\det \mathbf{G}_{m-1}$ and $\det \mathbf{G}_{m-2}$ are independent of the entries $[\mathbf{G}_m]_{m,m}$, $[\mathbf{G}_m]_{m,m-1}$, and $[\mathbf{G}_m]_{m+1,m}$.

It is worth mentioning that this approach cannot be extended for matrices with more than three nonzero diagonals.

ACKNOWLEDGMENT

The authors would like to thank David Bitton for his help with the implementation of the Monte Carlo simulation used in Section II-D1 (application of Proposition 3). We would also like to thank the Associated Editor Dr. Gerhard Kramer and the anonymous referee for constructive comments on the first draft of this paper.

REFERENCES

- [1] S. Shamai (Shitz), O. Somekh, and B. M. Zaidel, "Multi-cell communications: An information theoretic perspective," in *Proc. Joint Workshop on Communications and Coding (JWCC'04)*, Florence, Italy, Oct. 2004.
- [2] 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. Boca Raton, FL: Auerbach Publications, CRC Press, 2007.
- [3] A. J. Viterbi, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, no. 3, pp. 303–311, May 1991.
- [4] S. V. Hanly and P. A. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommun. Syst.*, vol. 1, pp. 1–42, 1993.
- [5] M. Zorzi, "Power control and diversity in mobile radio cellular systems in the presence of ricean fading and log-normal shadowing," *IEEE Trans. Veh. Technol.*, vol. 45, no. 3, pp. 373–382, May 1996.
- [6] M. Zorzi, "On the analytical computation of the interference statistics in cellular systems," *IEEE Trans. Commun.*, vol. 45, no. 1, pp. 103–109, Jan. 1996.

- [7] B. Rimoldi and Q. Li, "Potential impact of rate-splitting multiple access on cellular communications," in *Proc. IEEE GLOBECOM*, London, U.K., Nov. 1996, pp. 178–182.
- [8] S. A. Jafar, G. Foschini, and A. J. Goldsmith, "Phantomnet: Exploring optimal multicellular multiple antenna systems," *EURASIP J. Applied Signal Process., Special Issue on MIMO Communications and Signal Processing*, vol. 2004 (2004), no. 5, pp. 591–605, May 2004.
- [9] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.
- [10] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), "Sum-rate characterization of multiple-cell processing," in *Proc. Canadian Workshop on Information Theory (CWIT'05)*, McGill University, Montreal, QC, Canada, Jun. 2005, pp. 24–27.
- [11] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), "Sum rate characterization of joint multiple cell-site processing," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4473–4497, Dec. 2007.
- [12] N. Nasser, A. Hasswa, and H. Hassanein, "Handoffs in fourth generation heterogeneous networks," *IEEE Commun. Mag.*, vol. 44, no. 10, pp. 96–103, Oct. 2006.
- [13] S. Jing, D. N. C. Tse, J. Hou, J. Soriaga, J. E. Smee, and R. Padovani, "Downlink macro-diversity in cellular networks," in *Proc. IEEE Int. Symp. Information Theory (ISIT'07)*, Nice, France, Jun. 2007, pp. 1–5.
- [14] A. Narula, "Information Theoretic Analysis of Multiple-Antenna Transmission Diversity," Ph.D. dissertation, Massachusetts Institute of Technology (MIT), Cambridge, MA, Jun. 1997.
- [15] Y. Liang and A. Goldsmith, "Symmetric rate capacity of cellular systems with cooperative base stations," in *Proc. GLOBECOM 2006*, Nov./Dec. 2006, p. CTH13-3.
- [16] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, ser. Communications and Control Engineering Series. London, U.K.: Springer-Verlag London Ltd., 1993.
- [17] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.
- [18] S. Hanly and P. A. Whiting, "Constraint on capacity in a multi-user channel," in *Proc. Int. Symp. Information Theory (ISIT 94)*, Trondheim, Norway, Jun./Jul. 1994, p. 54.
- [19] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," in *Foundations and Trends in Communications and Information Theory*. Hanover, MA: Now Publishers, 2004, vol. 1.
- [20] G. W. Anderson and O. Zeitouni, "A clt for a band matrix model," *Probl. Theory Rel. Fields*, vol. 134, pp. 283–338, 2005.
- [21] S. Shamai (Shitz) and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.
- [22] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1329–1343, Jun. 2002.
- [23] A. Lozano, A. Tulino, and S. Verdú, "High-SNR power offset in multiple-antenna communications," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
- [24] 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. Inf. Theory*, vol. 37, no. 6, pp. 1527–1539, Nov. 1991.
- [25] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [26] L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators*. Berlin, Germany: Springer-Verlag, 1992, ser. *Grundlehren der Mathematischen Wissenschaften* [Fundamental Principles of Mathematical Sciences], vol. 297.
- [27] R. Carmona and J. Lacroix, "Spectral Theory on Random Schrödinger Operators," in *Probability and its Applications*. Boston, MA: Birkhäuser, 1990.
- [28] O. Somekh and S. Shamai (Shitz), "Shannon-theoretic approach to a Gaussian cellular multiple-access channel with fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1401–1425, Jul. 2000.
- [29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th ed. San Diego, CA: Academic, 2000.
- [30] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, and S. Shamai (Shitz), "Distributed multiple-cell zero-forcing beamforming in cellular downlink channels," *IEEE Trans. Inf. Theory*, submitted for publication.
- [31] H. A. David, *Order Statistics*, 2nd ed. New York: Wiley, 1981.
- [32] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.
- [33] 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.
- [34] N. Levy, O. Zeitouni, and S. Shamai (Shitz), "On information rates of the fading Wyner cellular model via the Thouless formula for the strip," *IEEE Trans. Inf. Theory*, submitted for publication.
- [35] B. N. Parlett, *The Symmetric Eigenvalue Problem Vol. 20 of Classics in Applied Mathematics*. Philadelphia, PA: SIAM, 1998, Corrected reprint of the 1980 original.
- [36] 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, 1985.
- [37] 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 *Contemporary Mathematics*. Providence, RI: AMS, 1986, pp. 23–30.
- [38] I. Y. Goldsheid and B. A. Khorunzhenko, "The Thouless formula for random non-Hermitian Jacobi matrices," *Israel J. Math.*, vol. 148, pp. 331–346, 2005.
- [39] P. Bougerol, "Théorèmes limites pour les systèmes linéaires à coefficients markoviens," *Prob. Theory and Related Fields*, vol. 78, no. 2, pp. 193–221, 1988.
- [40] P. Bougerol, "Comparaison des exposants de Lyapounov des processus Markoviens multiplicatifs," *Ann. Inst. H. Poincaré Probab. Statist.*, vol. 24, no. 4, pp. 439–489, 1988.
- [41] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

Nathan Levy (S'08) received the B.A. degree in mathematics from the Ecole Normale Supérieure, Paris, France, in 2005 and the M.A. degree in mathematics from the Université Paris-Sud, Paris, France, in 2006.

He is now working towards the Ph.D. degree in electrical engineering at the Technion–Israel Institute of Technology, Haifa, Israel. His research interests are random matrices theory and information theory.

Oren Somekh (S'90–M'06) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion–Israel Institute of Technology, Haifa, Israel, in 1989, 1991, and 2005, respectively.

During 1991–1996, he served in the Israel Defense Forces (IDF) in the capacity of a Research Engineer. During 1998–2002, he was the Vice President of Research and Development (VP R&D) and later the Chief Technical Officer (CTO) of Surf Communication Solutions Ltd., Yokneam, Israel. From April 2005 to October 2006, he was a Visiting Research Fellow at the Center for Wireless Communications and Signal Processing Research (CWCSRP), New Jersey Institute of Technology, Newark. Since November 2006, he has been a Visiting Research Fellow at the Electrical Engineering Department at Princeton University, Princeton, NJ. His current main research interests are information-theoretical aspects of cooperative wireless networks.

Dr. Somekh is a recipient of the European Community (EC) Marie Curie Outgoing International Fellowship (OIF), he was the corecipient of the first IEEE Information Theory Society ISIT best student paper award (2007), and the coinventor of several patents in the field of communications.

Shlomo Shamai (Shitz) (S'80–M'82–SM'88–F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion–Israel Institute of Technology, Haifa, Israel, in 1975, 1981, and 1986 respectively.

During 1975–1985, he was with the Communications Research Labs in the capacity of a Senior Research Engineer. Since 1986, he has been with the Department of Electrical Engineering, Technion–Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompass a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is a member of the Union Radio Scientifique Internationale (URSI). He is the recipient of the 1999 van der Pol Gold Medal of URSI, and a corecipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IEEE IT/COM societies paper award, and the 2007 IEEE Information Theory Society Paper Award. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and also served on the Board of Governors of the IEEE Information Theory Society.

Ofer Zeitouni (F'00) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion–Israel Institute of Technology, Haifa, Israel, in 1980, 1983, and 1986, respectively.

After completing a postdoctoral stay at Brown University, Providence, RI, and the Massachusetts Institute of Technology (MIT), Cambridge, he joined the Technion–Israel Institute of Technology, where he stayed until 2007, as Professor of Electrical Engineering and of Mathematics. Since 2002, he has been a Professor of Mathematics at the University of Minnesota, Minneapolis, and in 2007 he joined the Weizmann Institute in Rehovot, Israel, as a Professor of Mathematics. His research interests are in probability theory, with special emphasis on motion in random media, random matrices, and applications in communication, information, and filtering theories.

Prof. Zeitouni is a fellow of the Institute of Mathematical Statistics, and a corecipient of the 2003 best paper award from the *Annales Institut Poincaré*. He served as Associate Editor for numerous journal (including the IEEE TRANSACTIONS ON INFORMATION THEORY), and as Editor in Chief of the journals *Stochastic Processes and their Applications* (1999–2002) and *Annals of Probability* (2009–2011).