

The Shannon's mutual information of a multiple antenna time and frequency dependent channel: an ergodic operator approach

W. Hachem,^{*} A. Moustakas,[†] and L. Pastur[‡]

October 3, 2015

Abstract

Consider a random non-centered multiple antenna radio transmission channel. Assume that the deterministic part of the channel is itself frequency selective, and that the random multipath part is represented by an ergodic stationary vector process. In the Hilbert space $l^2(\mathbb{Z})$, one can associate to this channel a random ergodic self-adjoint operator having a so-called Integrated Density of States (IDS). Shannon's mutual information per receive antenna of this channel coincides then with the integral of a log function with respect to the IDS. In this paper, it is shown that when the numbers of antennas at the transmitter and at the receiver tend to infinity at the same rate, the mutual information per receive antenna tends to a quantity that can be identified and, in fact, is closely related to that obtained within the random matrix approach [37]. This result can be obtained by analyzing the behavior of the Stieltjes transform of the IDS in the regime of the large numbers of antennas.

1 Introduction and problem statement

In the landmark papers by Foschini and Gans [14] and by Telatar [37] the great promise of the use of multiple transmit and receive antennas was presented and established. The importance of such Multiple Input Multiple Output (MIMO) links is based on the fact that parallel data streams emanating from different transmit antennas can be decoded simultaneously from the receive array, thus making the throughput scale linearly with the number of transmit antennas. This scaling property is important in the quest to meet the expected thousand-fold increase of wireless network capacity in the coming years[33]. At the same time, the necessary transmit power per data stream is reduced by the same factor, thus also addressing energy-related pollution issues, which are becoming major societal and economical concerns [3, 9]. In fact, so-called *massive* MIMO systems of a few hundred antenna arrays [25, 35] have been identified as a key enabling technology for the next generation 5G wireless networks [2], promising unprecedented data transfer increases.

In this paper, we address the calculation of Shannon's mutual information of MIMO channels under some general assumptions on the channel statistics who have been considered only partially in the literature. To start with, we shall assume that the channel is subjected to time correlations that are more realistic than the commonly assumed multi-block-fading

^{*}CNRS LTCI; Telecom ParisTech, France (walid.hachem@telecom-paristech.fr),

[†]National & Capodistrian University of Athens, Department of Physics, Greece (arislm@phys.uoa.gr),

[‡]Institute of Low Temperature Physics, Kharkiv, Ukraine (pastur@ilt.kharkov.ua).

model [7, 4]. A second assumption on the model is related to the so-called frequency selectivity of the channel due to the delay spread induced by the reverberations of the signal from buildings, walls, etc. Typically, this effect is modelled using a tap-delay line system with different tap variances [4, 5]. A third component of the channel model is related with the so-called Ricean (or “deterministic”) part, due to *e.g.* a line of sight or a specular component. In this work, we consider in most generality that the deterministic part of the channel is also frequency selective [13].

To be more specific, let us denote by T the number of antennas at the transmitter and by N the number of antennas at the receiver. Then the \mathbb{C}^N -valued signal received at time k according to our model is

$$Y(k) = \sum_{\ell=-L}^L H(k, \ell)S(\ell) + V(k)$$

where the processes $\{S(k)\}_{k \in \mathbb{Z}}$ and $\{V(k)\}_{k \in \mathbb{Z}}$ represent respectively the \mathbb{C}^T -valued input signal fed to the channel and the \mathbb{C}^N -valued additive noise. The channel with $2L + 1$ matrix coefficients is represented $\mathbb{C}^{N \times (2L+1)T}$ -valued random process $\{\mathbf{H}(k) = [H(k, k-L), \dots, H(k, k+L)]\}_{k \in \mathbb{Z}}$, assumed to be Gaussian, stationary, ergodic, and generally non-centered. It is assumed that the processes $\{S(k)\}$, $\{V(k)\}$ and $\{\mathbf{H}(k)\}$ are mutually independent, and that $\{S(k)\}$ (respectively $\{V(k)\}$) is an independent process such that $S(k) \sim \mathcal{CN}(0, I_T)$ (resp. $V(k) \sim \mathcal{CN}(0, I_N)$) for any $k \in \mathbb{Z}$.

As is well-known, the Doppler effect induced by the mobility of the communicating terminals determines the form of the covariance function of the process $\{\mathbf{H}(k)\}$. The multipath effect that induce the frequency selectivity is captured by the $2L + 1$ matrix channel coefficients which are subjected in the practical situations to a certain power profile. In addition, the fact that $\mathbb{E}\mathbf{H}(0)$ can be non zero and can be also frequency selective is due *e.g.* a line of sight or a specular component. We shall formulate our assumptions on the channel more precisely in Section 2 below.

The signal $Y^n = [Y(-n)^T, \dots, Y(n)^T]^T$ observed during the time window $(-n, \dots, n)$ satisfies the equation $Y^n = H^n S^n + V^n$ where $V^n = [V(-n)^T, \dots, V(n)^T]^T$, $S^n = [S(-n-L)^T, \dots, S(n+L)^T]^T$, and

$$H^n = \begin{bmatrix} H(-n, -n-L) & \cdots & H(-n, -n+L) & & 0 \\ & & \ddots & & \\ & 0 & & H(n, n-L) & \cdots & H(n, n+L) \end{bmatrix} \quad (1)$$

is a $(2n+1)N \times (2n+2L+1)T$ matrix. Our goal is to study the mutual information of this channel assuming that it is perfectly known at the receiver. More precisely, we study $I(S; (Y, H))$, the mutual information between $\{S(k)\}$ and the couple $(\{Y(k)\}, \{\mathbf{H}(k)\})$. It is given by the equation

$$I(S; (Y, H)) = \limsup_n \frac{1}{2n+1} I(S^n; (Y^n, H^n))$$

where

$$I(S^n; (Y^n, H^n)) = \mathbb{E} \log \det(H^n H^{n*} + I)$$

is the mutual information between S^n and (Y^n, H^n) (see [15, Chap. 8], [37] and Section 4 below).

We shall see that the natural way to tackle this problem is to consider the MIMO channel as a *random operator* represented by the doubly infinite matrix $H = [H(k, \ell)]_{k, \ell \in \mathbb{Z}}$ and acting on the Hilbert space $l^2(\mathbb{Z})$ of the doubly infinite square summable sequences. Due to the ergodicity of $(\mathbf{H}(k))$, this operator that we also denote as H is *ergodic* in the sense of [29]. By the ergodicity, it turns out that the self-adjoint operator HH^* where H^* is the adjoint of H has a so-called *Integrated Density of States* (IDS). Specifically, there exists a positive deterministic measure μ of total mass one such that on a probability one set,

$$\frac{1}{(2n+1)N} \text{Tr} g(H^n H^{n*}) \xrightarrow{n \rightarrow \infty} \int g(\lambda) \mu(d\lambda)$$

for any continuous and bounded real function g . The IDS is the distribution function of μ . The crucial observation is that this convergence leads to the convergence of $I(S^n; (Y^n, H^n))/(2n+1)$, and the limit is

$$I(S; (Y, H)) = N \int \log(1 + \lambda) \mu(d\lambda)$$

as it will be shown below. Our goal is then to study the behavior of the integral at the right hand side. Unfortunately, a few can be said about this behavior in the general situation. To circumvent this problem, one needs to resort to a certain asymptotic regime.

In this paper, consistently with an established practice in the evaluation of the mutual information of MIMO channels, we consider an asymptotic regime where the numbers of antennas N and T tend to infinity at the same rate. A central object of study will be the Stieltjes transform of the measure μ , that is the complex analytical function

$$\mathbf{m}(z) = \int \frac{1}{\lambda - z} \mu(d\lambda)$$

on the upper half-plane $\mathbb{C}_+ = \{z : \Im z > 0\}$. This function completely characterizes μ and is intimately connected with the resolvent

$$Q(z) = (HH^* - zI)^{-1}$$

of the operator HH^* . Specifically, considering the matrix block representation $Q(z) = [Q(k, \ell)(z)]_{k, \ell \in \mathbb{Z}}$ of the resolvent where the blocks $Q(k, \ell)(z)$ are $N \times N$ complex matrices, it holds that

$$\mathbf{m}(z) = \frac{\text{Tr} \mathbb{E} Q(0, 0)(z)}{N}.$$

The core of our analysis consists in studying the behavior of the right hand side of this relation in the large $N(T), T$ regime. Re-denoting μ and $\mathbf{m}(z)$ as μ_T and $\mathbf{m}_T(z)$ to stress the dependency on T , it turns out that there exists a sequence of probability measures π_T which approximates μ_T in the sense that

$$\mathbf{m}_T(z) - \int \frac{1}{\lambda - z} \pi_T(d\lambda) \xrightarrow{T \rightarrow \infty} 0.$$

In addition, we can associate to each π_T a positive number \mathcal{I}_T of order one for large T , and such that

$$\frac{1}{N} I(S; (Y, H)) - \mathcal{I}_T \xrightarrow{T \rightarrow \infty} 0.$$

Thus, the mutual information increases at a linear scale with the number of antennas at one side of the transmission. The Stieltjes transform of π_T as well as \mathcal{I}_T satisfy a system of equations that can be easily implemented on a computer. Solving for \mathcal{I}_T , one can evaluate the impact of the various statistical parameters of the channel model on the mutual information.

It is also interesting to observe that asymptotic regimes other than the one where N and T converge to infinity at the same rate can be considered. For instance, the techniques of this paper allow to study the asymptotic behavior of the Stieltjes transform of the IDS of HH^* in the case where N and T are fixed and where L tends to infinity. This study would lead to describe the behavior of the mutual information of channels with a long memory. We restrict ourselves here to the regime of the large number of antennas to limit the size of this paper.

In the literature dealing with the ergodic mutual information of MIMO channels, it has been frequently assumed that the channel was a *frequency non selective* channel which satisfies the simplifying *block-fading* assumption (see *e.g.* [37]). In the language of the ergodic operator theory, the operator HH^* corresponding to such channels is a block diagonal operator with independent and identically distributed diagonal blocks, and the existence of the IDS is immediate. In order to approximate the Stieltjes transform of μ_T in the regime of the large number of antennas, one only needs to study the behavior of

$$\frac{1}{N} \text{Tr} \mathbb{E} Q(0, 0)(z) = \frac{1}{N} \text{Tr} \mathbb{E} (H(0, 0)H(0, 0)^* - zI_N)^{-1}.$$

The large N, T behavior of this object can in fact be studied with the help of random matrix theory techniques, without making any explicit reference to ergodic operators. In this framework, a number of works studied $N^{-1} \text{Tr} \mathbb{E} Q(0, 0)(z)$ with more and more sophisticated statistical model for $H(0, 0)$. Among these, centered models with the so-called single-sided or double-sided correlations were studied in [10, 26, 28, 16]. A more general centered model was considered in [38]. Non-centered models were considered in [27, 36, 12] among others. One contribution of our paper is that we consider frequency selective channels and a model for the Doppler effect more realistic than the block-fading model.

Ergodic operator theory (see *e.g.* [29]) has aroused a considerable interest in the fields of operator theory, quantum physics, and statistical mechanics where these operators are frequently used to model the Hamiltonians of randomly disordered systems. In this framework, the contribution [23] studied the asymptotic behavior of the IDS of certain random Hermitian operators when some design parameter is made converge to infinity. For the models considered in [23], it is shown that the limiting IDS coincides with that of a deformed large Wigner matrix, well known in random matrix theory. The connections between ergodic operator theory and random matrix theory are further explored in [30]. In this paper, we adopt the general approach of [23, 30], dealing with random operator models à la Marchenko-Pastur. In order to perform our asymptotic analysis, we rely on two tools that belong to the arsenal of random matrix theory, as shown in [31] and the references therein: an integration by parts formula for the expectation of functions of Gaussian vectors (already used in [23]), and the Poincaré-Nash inequality to control our variances.

In the communication theory literature, the ergodic operator formalism for studying the mutual information of multi-antenna channels was used in [20, 21]. These contributions

brought to the fore the relation between the mutual information and the IDS without elaborating on the properties of the latter.

In the next section of this paper, we formulate precisely our assumptions and we state our results. These results are illustrated in Section 3 by some computer simulations. An overview of the ergodic operator theory is provided in Section 4. The proofs of the main results are provided in Sections 5 and 6. The paper ends with a conclusion.

2 Assumptions and the results

We start by introducing the exact statistical model of our channel. Since the number of antennas will tend to infinity, the channel will be described by a sequence of random operators on $l^2(\mathbb{Z})$ indexed by the parameter T .

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and two sequences of positive integers $(N(T))_{T \in \mathbb{N}}$ and $(L(T))_{T \in \mathbb{N}}$, consider the sequence of random operators

$$H_T = A_T + X_T$$

with domain $\mathcal{D}(H_T) = \mathcal{K}$, the manifold of vectors of $l^2(\mathbb{Z})$ with finite support. The operator A_T is a deterministic operator whose matrix represented in the canonical basis of $l^2(\mathbb{Z})$ by the matrix $A_T = [A_T(k - \ell)]_{k, \ell \in \mathbb{Z}}$ where the block $A_T(k)$ is a $N(T) \times T$ complex matrix, and where the sequence $(\dots, A_T(-1), A_T(0), A_T(1), \dots)$ is supported by $\{-L(T), \dots, L(T)\}$. Since the blocks of A_T are identical along the diagonals, A_T is a convolution operator [6].

The operator X_T is a \mathcal{K} -defined random operator on $l^2(\mathbb{Z})$ (in the sense of [29, Sec. I.1.B], see Section 4 below) whose matrix representation in the canonical basis of $l^2(\mathbb{Z})$ is the band matrix

$$X_T = \begin{bmatrix} \ddots & & \ddots & & \ddots & \ddots & \ddots & & 0 \\ & \ddots & X_T(0, -1) & X_T(0, 0) & X_T(0, 1) & \ddots & & & \\ 0 & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & & & & \end{bmatrix} \\ = [X_T(k, \ell), k, \ell \in \mathbb{Z}, |k - \ell| \leq L(T)].$$

The $N(T) \times T$ random complex matrix $X_T(k, \ell)$ is written as

$$X_T(k, \ell) = [X_{T,n,t}(k, \ell), n = 0 : (N(T) - 1), t = 0 : (T - 1)]$$

and satisfies

$$X_T(k, \ell) = \frac{1}{\sqrt{T}} \phi_T(k - \ell) W_T(k, \ell).$$

Here $\phi_T : \mathbb{Z} \rightarrow [0, \infty)$ is a real valued function supported by the set $\{-L(T), \dots, L(T)\}$, and $(W_T(k, \ell) = [W_{T,n,t}(k, \ell), n = 0 : (N(T) - 1), t = 0 : (T - 1)])_{k, \ell \in \mathbb{Z}}$ is a complex Gaussian circularly symmetric centered random field such that

$$\mathbb{E}[W_{T,n_1,t_1}(k_1, \ell_1) \bar{W}_{T,n_2,t_2}(k_2, \ell_2)] = \delta_{n_1,n_2} \delta_{t_1,t_2} \delta_{k_1-\ell_1, k_2-\ell_2} \gamma_T(k_1 - k_2)$$

where the covariance function $\gamma_T(k)$ satisfies $\gamma_T(0) = 1$ without generality loss.

The operator A_T models the deterministic part of the channel. The function ϕ_T models the multipath amplitude profile, while the covariance function γ_T is related to the Doppler effect. As is well known, the function $\gamma_T(k)$ converges quickly to zero when the mobile terminal velocity is high.

The following set of assumptions will be needed:

1. $0 < \liminf_{T \rightarrow \infty} N(T)/T \leq \sup_T N(T)/T < \infty$. We write $\mathbf{c} = \sup_T N(T)/T$.
2. $\sigma_T^2 = \sum_{\ell} \phi_T(\ell)^2$ satisfies $0 < \liminf_T \sigma_T^2 \leq \sup_T \sigma_T^2 < \infty$. We let $\boldsymbol{\sigma}^2 = \sup_T \sigma_T^2$.
3. $\mathbf{g} = \sup_T \sum_{\ell} |\gamma_T(\ell)|$ is finite.
4. $\mathbf{a} = \sup_T \sum_{\ell=-L(T)}^{L(T)} \|A_T(\ell)\|$ is finite, where $\|\cdot\|$ is the spectral norm.

The asymptotic regime described by the above assumptions will be concisely denoted “ $T \rightarrow \infty$ ”. Notice that there is no bound on the number of channel coefficients $2L(T) + 1$.

Let us comment these assumptions. The parameter σ_T^2 in Assumption 2 is the part of the received power due to the centered part of the channel. The practical interpretation of Assumption 3 is that the so-called coherence time of the channel [4] does not grow with T . At this point, we stress the importance of the fact that our results on the mutual information assume the channel to be perfectly known at the receiver. In practice, this channel needs to be estimated, and this task will be getting harder as T grows if the coherence time is kept fixed (see however the discussion around Fig. 3 below). We believe that relaxing Assumption 3 requires other mathematical tools than those used in this paper. Finally, Assumption 4 could be certainly weakened and replaced with *e.g.* a bound on the Euclidean norms of the columns and the rows of the matrices $A(\ell)$ at the expense of a more involved proof for Theorem 2.3 below. We also point out that the convergence rates obtained in the proof of this theorem can be improved. We chose to keep Assumption 4 and the present proof for simplicity. Note also that we assume the Gaussian distribution of various random variables processes above largely for the simplicity of the corresponding proofs. By using a version of the so-called interpolation trick (see [31, Sections 18.3–18.4]), our results can be extended to the case where the corresponding ergodic random processes are not necessarily Gaussian just having the same covariance and with a certain number finite moments and sufficiently fast decaying correlations, although the respective proofs of the results become rather involved.

As mentioned in the introduction, the Stieltjes transform of a probability measure ν on \mathbb{R} plays a fundamental role in this paper. This is the function

$$\mathbf{s}(z) = \int \frac{1}{\lambda - z} \nu(d\lambda)$$

defined on \mathbb{C}_+ . This function is *i)* holomorphic on \mathbb{C}_+ , *ii)* it satisfies $\mathbf{s}(z) \in \mathbb{C}_+$ for any $z \in \mathbb{C}_+$, and *iii)* $\lim_{y \rightarrow \infty} (-iy)\mathbf{s}(iy) = 1$. In addition, if ν is supported by $\mathbb{R}_+ = [0, \infty)$, then *iv)* $\Im(z\mathbf{s}(z)) \geq 0$ for any $z \in \mathbb{C}_+$. Conversely, any function $\mathbf{s}(z)$ satisfying *i)–iv)* is the Stieltjes Transform of a probability measure supported by \mathbb{R}_+ [24]. Observe that the Stieltjes

Transform of ν can be trivially extended from \mathbb{C}_+ to $\mathbb{C} - \text{supp}(\nu)$ where $\text{supp}(\nu)$ is the support of ν . Finally, “probability measure” can be replaced with “positive measure ν such that $0 < \nu(\mathbb{R}) < \infty$ ” in the preceding statements if we replace *iii*) with $\lim_{y \rightarrow \infty} (-iy)\mathbf{s}(iy) = \nu(\mathbb{R})$.

We shall now state our results. Due to the fact that γ_T is summable, the Gaussian stationary random process

$$\{\mathbf{H}_T(k) = [H_T(k, k - L(T)), \dots, H_T(k, k + L(T))]\}_{k \in \mathbb{Z}}$$

is ergodic. This makes the operator H_T ergodic in a sense made precise below. The characterization of the mutual information in the framework of ergodic operators is given by the two following results.

Proposition 2.1. *On a probability one set, the operator H_T is closable. Still denoting by H_T the closure of this operator and by H_T^* its adjoint, the positive self-adjoint operator $H_T H_T^*$ is ergodic and has an IDS defining a probability measure μ_T . Denoting by $Q_T(z) = (H_T H_T^* - zI)^{-1} = [Q_T(k, \ell)(z)]_{k, \ell \in \mathbb{Z}}$ the resolvent of $H_T H_T^*$ for $z \in \mathbb{C}_+$, where the blocks $Q_T(k, \ell)(z)$ are $N \times N$ matrices, the Stieltjes transform of μ_T is $\mathbf{m}_T(z) = N^{-1} \mathbb{E} \text{Tr} Q_T(0, 0)(z)$.*

Theorem 2.1. *The sequence $\{I(S^n; (Y^n, H^n)) / (2n + 1)\}$ converges as $n \rightarrow \infty$, and its limit is the mutual information $I_T(S; (Y, H)) = N \int \log(1 + \lambda) \mu_T(d\lambda) < \infty$.*

Let

$$\gamma_T(f) = \sum_k \exp(2i\pi k f) \gamma_T(k)$$

be the Fourier transform of the sequence $\{\gamma_T(k)\}$, and

$$\mathbf{A}_T(f) = \sum_k \exp(2i\pi k f) \mathbf{A}_T(k)$$

be the $N \times T$ Fourier transform of the sequence $\{\mathbf{A}_T(k)\}$. Write $(\mathbf{A}_T \mathbf{A}_T^*)(f) = \mathbf{A}_T(f) \mathbf{A}_T^*(f)$ and $(\mathbf{A}_T^* \mathbf{A}_T)(f) = \mathbf{A}_T^*(f) \mathbf{A}_T(f)$ for compactness.

Theorem 2.2. *Let $\mathbf{S}_T(f, z)$ and $\tilde{\mathbf{S}}_T(f, z)$ be respectively the $N \times N$ and $T \times T$ matrices*

$$\mathbf{S}_T(f, z) = \left[-z(1 + \sigma_T^2 \gamma_T(f) \star \tilde{\varphi}_T(f, z)) I_N + (1 + \sigma_T^2 \gamma_T(-f) \star \varphi_T(f, z))^{-1} (\mathbf{A}_T \mathbf{A}_T^*)(f) \right]^{-1}, \quad (2)$$

$$\tilde{\mathbf{S}}_T(f, z) = \left[-z(1 + \sigma_T^2 \gamma_T(-f) \star \varphi_T(f, z)) I_T + (1 + \sigma_T^2 \gamma_T(f) \star \tilde{\varphi}_T(f, z))^{-1} (\mathbf{A}_T^* \mathbf{A}_T)(f) \right]^{-1}, \quad (3)$$

and

$$\begin{aligned} \gamma_T(f) \star \tilde{\varphi}_T(f, z) &= \int_0^1 \gamma_T(f - u) \tilde{\varphi}_T(u, z) du, \text{ and} \\ \gamma_T(-f) \star \varphi_T(f, z) &= \int_0^1 \gamma_T(u - f) \varphi_T(u, z) du. \end{aligned}$$

Then for any $z \in \mathbb{C}_+$, the system of equations

$$\varphi_T(f, z) = \frac{\text{Tr} \mathbf{S}_T(f, z)}{T} \quad \text{and} \quad \tilde{\varphi}_T(f, z) = \frac{\text{Tr} \tilde{\mathbf{S}}_T(f, z)}{T}$$

admits a unique solution $(\varphi_T(\cdot, z), \tilde{\varphi}_T(\cdot, z))$ such that $\varphi_T(\cdot, z), \tilde{\varphi}_T(\cdot, z) : [0, 1] \rightarrow \mathbb{C}$ are both measurable and Lebesgue-integrable on $[0, 1]$ and such that $\Im\varphi(f, z), \Im\tilde{\varphi}(f, z), \Im(z\varphi(f, z))$ and $\Im(z\tilde{\varphi}(f, z))$ are nonnegative for any $f \in [0, 1]$.

The solutions $\varphi_T(\cdot, z)$ and $\tilde{\varphi}_T(\cdot, z)$ are continuous on $[0, 1]$, and $\Im\varphi(f, z), \Im\tilde{\varphi}(f, z), \Im(z\varphi(f, z))$ and $\Im(z\tilde{\varphi}(f, z))$ are positive for any $f \in [0, 1]$. Furthermore, for any $f \in [0, 1]$, the functions $(T/N)\varphi(f, z)$ and $\tilde{\varphi}(f, z)$ are defined on $\mathbb{C} - \mathbb{R}_+$ and are the Stieltjes transforms of probability measures supported by $[0, \infty)$.

This system of equations turns out to be formally close to that described in [17, Th. 2.4]. The uniqueness established in the present paper is a point-wise uniqueness (*i.e.*, for any $z \in \mathbb{C}_+$) which is stronger than the type of uniqueness shown in [17]. The proof of Theorem 2.2 is given in Section 5.

The discussion preceding Proposition 2.1 shows that the function

$$\mathbf{p}_T(z) = \frac{1}{N} \int_0^1 \text{Tr } \mathbf{S}_T(f, z) df$$

is the Stieltjes transform of a probability measure carried by $[0, \infty)$. We denote it π_T .

Theorem 2.3. *For any $z \in \mathbb{C}_+$,*

$$\mathbf{m}_T(z) - \mathbf{p}_T(z) \xrightarrow{T \rightarrow \infty} 0. \quad (4)$$

Moreover, the sequences $\{\mu_T\}$ and $\{\pi_T\}$ are tight, and

$$\int g(\lambda) \mu_T(d\lambda) - \int g(\lambda) \pi_T(d\lambda) \xrightarrow{T \rightarrow \infty} 0$$

for any continuous and bounded real function g .

The approximation of the mutual information for large T is finally provided by the following theorem.

Theorem 2.4. *We have*

$$N^{-1}I_T(S; (Y, H)) - \mathcal{I}_T \xrightarrow{T \rightarrow \infty} 0,$$

where

$$\mathcal{I}_T = \int \log(1 + \lambda) \pi_T(d\lambda).$$

and the integral is given by

$$\begin{aligned} \mathcal{I}_T &= \frac{1}{N} \int_0^1 \log \det \left((1 + \sigma_T^2 \gamma_T(f) \star \tilde{\varphi}_T(f, -1)) I_N + \frac{(\mathbf{A}_T \mathbf{A}_T^*)(f)}{1 + \sigma_T^2 \gamma_T(-f) \star \varphi_T(f, -1)} \right) df \\ &\quad + \frac{T}{N} \int_0^1 \log(1 + \sigma_T^2 \gamma_T(-f) \star \varphi_T(f, -1)) df \\ &\quad - \frac{T}{N} \int_0^1 \int_0^1 \sigma_T^2 \gamma_T(f - v) \tilde{\varphi}_T(v, -1) \varphi_T(f, -1) dv df. \end{aligned}$$

Before turning to the numerical illustrations of these results, two remarks are in order. We first observe that the variance profile represented by the function $\phi_T^2 : \mathbb{Z} \rightarrow [0, \infty)$ has no influence on \mathcal{I}_T except through the total received power σ_T^2 due to the random part of the channel.

We also observe that if the channel is centered, *i.e.*, if $A_T = 0$, then $\boldsymbol{\pi}_T$ is the Marchenko-Pastur distribution

$$\boldsymbol{\pi}_T(d\lambda) = \frac{1}{2\pi c_T \sigma_T^2 \lambda} \sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)} \mathbb{1}_{[\lambda_-, \lambda_+]}(\lambda) d\lambda + ((1 - c_N^{-1}) \vee 0) \delta_0$$

where $c_T = N/T$, $\lambda_+ = \sigma_T^2(1 + \sqrt{c_T})^2$, and $\lambda_- = \sigma_T^2(1 - \sqrt{c_T})^2$. It is indeed well known that the Stieltjes Transform of $\boldsymbol{\pi}_T$ is $c_N^{-1}\alpha(z)$, where $\alpha(z)$ is the unique solution in \mathbb{C}_+ of the equations

$$\alpha(z) = c_N(-z - z\sigma_T^2\tilde{\alpha}(z))^{-1}, \quad \tilde{\alpha}(z) = (-z - z\sigma_T^2\alpha(z))^{-1}.$$

Recalling that $\int \gamma_T(f) df = 1$, one can immediately check that when $A_T = 0$, the couple $(\varphi_T(f, z), \tilde{\varphi}_T(f, z)) = (\alpha(z), \tilde{\alpha}(z))$ is a solution of the system described in the statement of Theorem 2.2. The result follows by the uniqueness of this solution.

3 Numerics

In this section we will present the behavior of the mutual information for some representative cases and will compare with numerically generated instantiations. To begin with we present the usually accepted model for the temporal correlation of fast fading, *i.e.*, the so-called Jakes model [19], which, in the time domain has the following correlation form:

$$\gamma_T(t) = J_0\left(\frac{2\pi vt}{\lambda}\right)$$

which in the frequency domain becomes

$$\gamma_T(f) = \frac{1}{\pi} \frac{1}{\sqrt{f_d^2 - f^2}}$$

within the region $|f| < f_d$ and zero elsewhere, where $f_d = v/\lambda\tau$, the ratio of the velocity of the mobile to the wavelength times the time duration of the channel usage. It should be noted that the discontinuity at f_d results to $\gamma_T(t)$ not being absolutely summable, and hence strictly speaking it cannot be used in this paper. However, demanding the frequency response to be continuous at $f = f_d$, (by adding a small rounding factor in the frequency domain), we can make this to be an acceptable model for the system. For simplicity we will not include this in the simulations.

To move on, we need to also present a model for the deterministic matrix function $\mathbf{A}_T(f)$. The typical situation for wireless communications is that the constant matrices are due to line-of-sight rank-one components. Along these lines we assume that each of the time resolvable paths have the following matrix elements

$$A_T(k)_{m,n} = \exp[-|k|\xi/L_{tot}] \exp[2\pi j(m-n)\sin\theta_k] / \sqrt{N}$$

where $\theta_k = k\pi/L_{tot}$, for $k = -L, \dots, L$ and $L_{tot} = 2L + 1$. The exponential dependence on the delay spread ξ has been seen experimentally [32, 8]. Then $\mathbf{A}_T(f)$ follows from

$$\mathbf{A}_T(f) = \sum_{\ell=-L}^L A_T(k) \exp[-2\pi j\ell f]$$

Note that the rank of the above matrix is L_{max} .

In the next figures, we observe the behavior of the mutual information as a function of the Signal to Noise Ratio (SNR) parameter

$$\rho = \sigma_T^2 + \frac{1}{N} \int_0^1 \text{Tr}(\mathbf{A}_T \mathbf{A}_T^*)(f) df$$

in various cases. In Fig. 1 we plot the mutual information for different values of the Ricean coefficient K defined as the ratio between the power of the deterministic part channel to that of the random part of the channel for the case. We see that for increasing values of K the mutual information initially increases because the deterministic channel provides higher throughput, but then decreases because for higher K the rank-deficiency of the deterministic channel becomes apparent.

In Fig. 2 we see that the mutual information decreases with increasing inverse delay spread ξ , which is essentially the frequency correlation bandwidth. This behavior is expected, because the smaller ξ is the smaller the attenuation and thus the higher the strength of the channel at the receiver.

We have also found that the dependence of the mutual information on the value of f_d is rather benign and not discernable for the above type of deterministic channel. This can be attributed to the fact that the dependence of the eigenvalue distribution of $(\mathbf{A}_T \mathbf{A}_T^*)(f)$ on the frequency value f is quite small. For other less realistic examples of deterministic channels, *e.g.*, the case where $A(k) = \exp(-|k|\xi)I_N$, where I_N is the identity matrix, does depend strongly on f_d .

Finally, in Fig. 3 we compare the results obtained from the methodology presented in this paper with numerically generated values. In this figure we plot the average mutual information generated numerically for various block sizes (here depicted by M). In this case for simplicity, we have used the exponentially decaying temporal correlation model, *i.e.*, with $\gamma_T(k) = \exp(-|k|f_d)$. This model is easier to implement numerically and absolutely summable, but is not very realistic because it is very wide-band. The first important observation here is that this approach gives near exact results for antenna arrays as small as $N = T = 2$ presented in this figure. Second, we see that depending on the value of f_d the convergence to the asymptotic result depends on the actual block size. For small f_d (temporally correlated channels) the $M = 5$ set of points (diamonds) is significantly deviating from the analytic curve, while for large f_d the $M = 5$ simulations are much closer to the asymptotic result. Increasing the block size to $M = 40$ makes the numerical values right on the analytical ones.

4 An overview of ergodic operators. Proofs of Proposition 2.1 and Theorem 2.1

To make this paper reasonably self-contained, we provide a brief review of the ergodic operator theory that will be useful to our purpose. In passing, we shall prove Proposition 2.1. For a

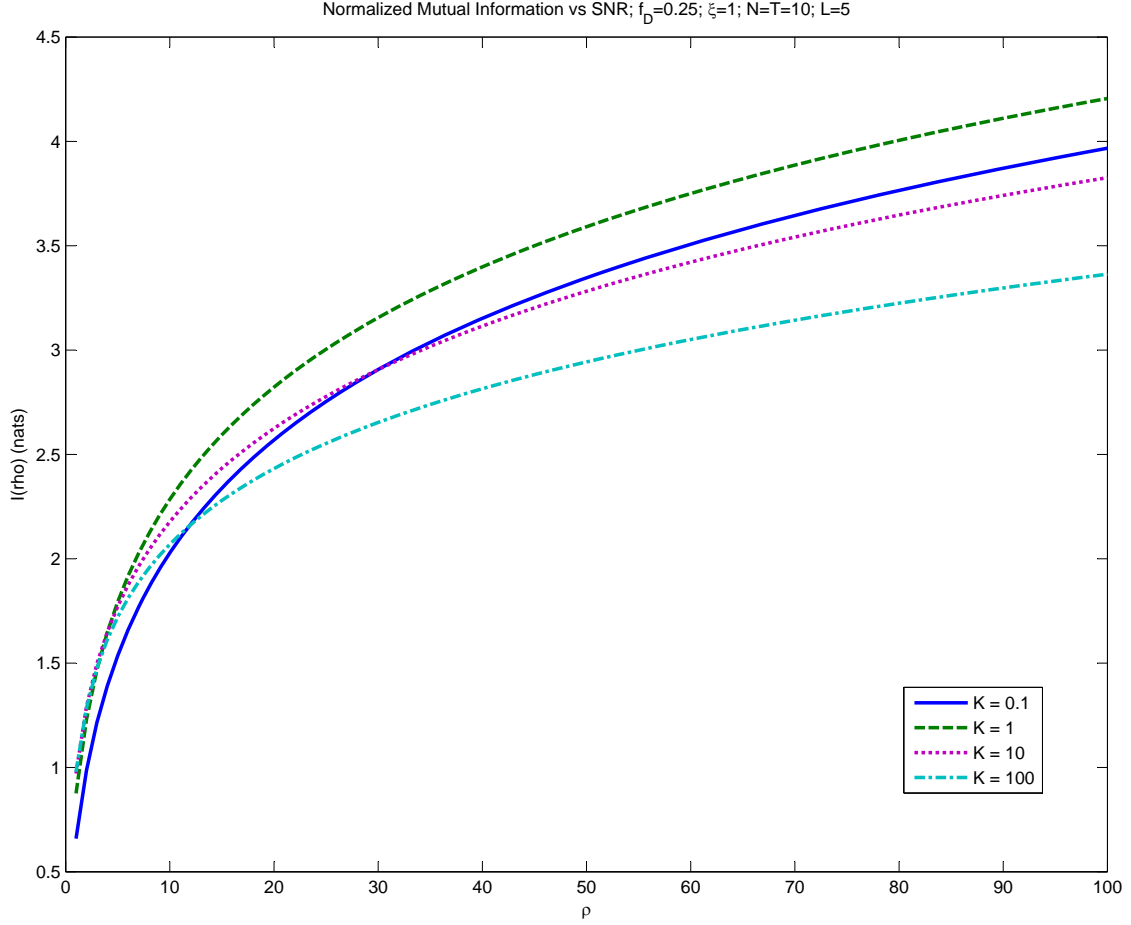


Figure 1: Mutual Information versus ρ for various K

comprehensive exposition of this theory, the reader may consult the book [29].

Since the operator H_T is defined for every $\omega \in \Omega$ on the dense linear subspace \mathcal{K} of $l^2(\mathbb{Z})$, and since the vector $H_T a$ is a random vector in $l^2(\mathbb{Z})$ for any $a \in \mathcal{K}$, the operator H_T is random, following the definition of [29, Sec. I.1.B].

For the presentation simplicity, let us temporarily assume that $N = T = 1$ and suppress the index T . Since \mathcal{K} is dense, H has an adjoint H^* with domain

$$\begin{aligned} \mathcal{D}(H^*) &= \left\{ a \in l^2(\mathbb{Z}) : \sum_k |\langle H e_k, a \rangle|^2 < \infty \right\} \\ &= \left\{ a = \sum_{k \in \mathbb{Z}} \alpha_k e_k : \sum_{k \in \mathbb{Z}} \left| \sum_{\ell=-L}^L H(k+\ell, k) \bar{\alpha}_{k+\ell} \right|^2 < \infty \right\} \end{aligned}$$

where e_k is the k^{th} canonical basis vector of $l^2(\mathbb{Z})$. Clearly $\mathcal{K} \subset \mathcal{D}(H^*)$, so this domain is dense in $l^2(\mathbb{Z})$. Therefore, H is closable, and its closure coincides with H^{**} . In the remainder, we reuse the notation H to designate this closure. Note that with this extension, the set \mathcal{K} becomes a core for H [34, Chap. VIII]. Since H is closed and densely defined, the operator HH^* is a positive self-adjoint operator with \mathcal{K} as a core [1, Th. 46.2].

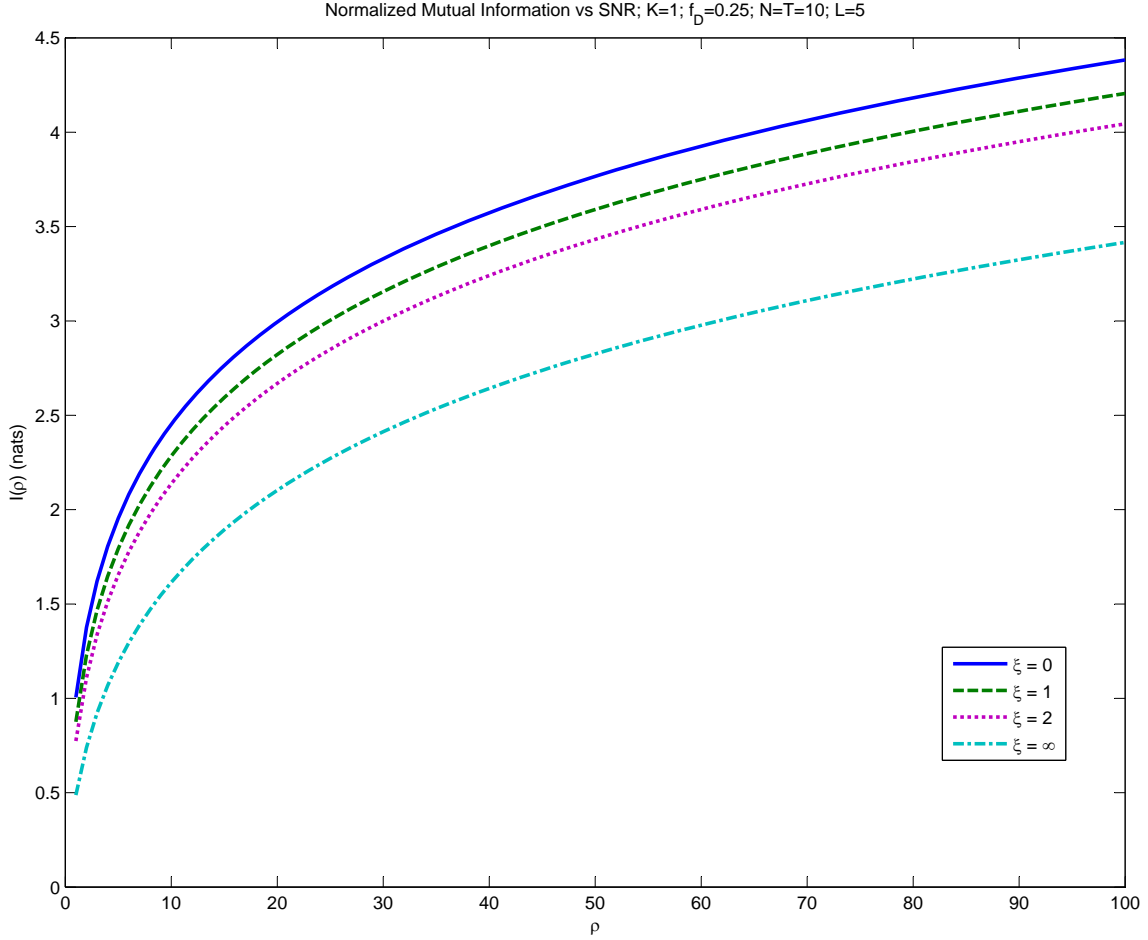


Figure 2: Mutual Information versus ρ for various ξ

We shall now consider the ergodic properties of HH^* . Making explicit the dependence on the elementary event ω in the definition of the process

$$\mathbf{H}(\omega) = \{\mathbf{H}(\omega, k)\}_{k \in \mathbb{Z}}, \quad \mathbf{H}(\omega, k) = [H(k, k-L), \dots, H(k, k+L)],$$

it is easy to see that the summability of γ implies that the shift $B : \Omega \rightarrow \Omega$ defined by the equation $\mathbf{H}(B\omega, k) = \mathbf{H}(\omega, k+1)$ is ergodic. Moreover, the random operator $H(\omega)$ clearly satisfies the equation

$$H(B\omega) = UH(\omega)U^{-1}$$

where U is the (unitary) shift operator $Ua = \sum_k \alpha_{k+1}e_k$ for $a = \sum_k \alpha_k e_k \in l^2(\mathbb{Z})$. An operator on $l^2(\mathbb{Z})$ that satisfies such an equation where B is ergodic is said ergodic in the sense of [29, page 33]. Now, writing $(HH^*)(\omega) = H(\omega)H^*(\omega)$, we also have $(HH^*)(B\omega) = U(HH^*)(\omega)U^{-1}$. In addition, it is clear that $UK = \mathcal{K}$. Therefore, the operator HH^* is also ergodic in the sense of [29, page 33].

Consider now the sequence of finite dimensional matrices $(H^n H^{n*})_{n \in \mathbb{N}}$ obtained by truncating the matrix HH^* and keeping the elements (i, j) such that $|i|, |j| \leq n$, see Eq. (1). Let

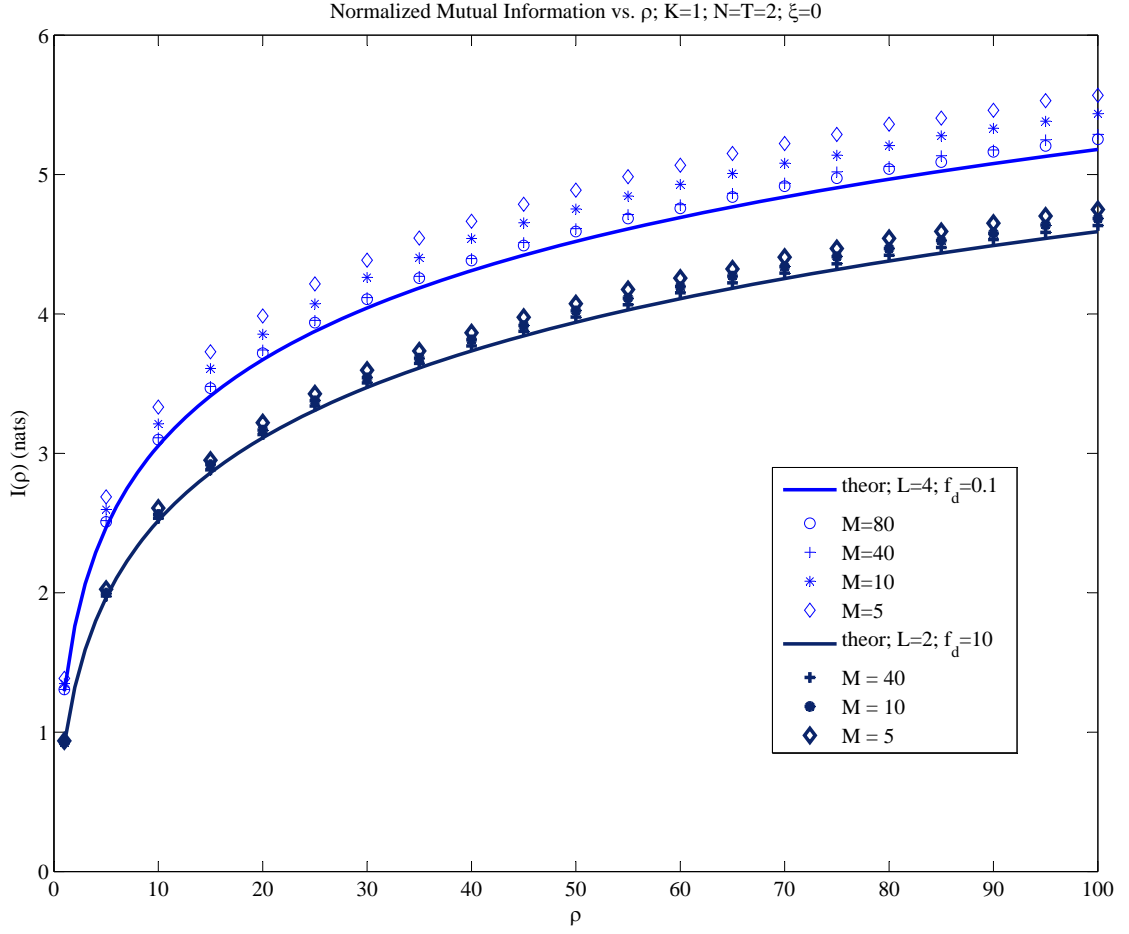


Figure 3: Mutual Information versus ρ for various L and block sizes

$\lambda_{-n,n}, \dots, \lambda_{n,n}$ be the eigenvalues of $H^n H^{n*}$ and let

$$\nu^n(\lambda) = \frac{1}{2n+1} \sum_{i=-n}^n \delta_{\lambda_{i,n}}$$

be the Normalized Counting Measure (NCM) of this matrix. Since the operator HH^* is represented by a band matrix, we have the following result, stemming directly from [29, Th. II-4.8]:

Proposition 4.1. *The operator HH^* has an IDS. In other words, there exists a deterministic probability measure μ on $[0, \infty)$ such that*

$$\int g(\lambda) \nu^n(d\lambda) \xrightarrow[n \rightarrow \infty]{a.s.} \int g(\lambda) \mu(d\lambda) \quad (5)$$

for all continuous and bounded functions g . Moreover, the Stieltjes transform of μ coincides with $\mathbb{E}Q(0,0)(z)$.

Proposition 2.1 for $N = T = 1$ follows. It will be useful to outline the proof of Proposition 4.1. The idea is to establish the convergence in the statement of this proposition for a large enough family of continuous functions. Given a continuous function g , start writing the left hand side of Eq. (5) as

$$\int g(\lambda) \nu^n(d\lambda) = \frac{1}{2n+1} \sum_{i=-n}^n g(H^n H^{n*})(i, i)$$

where $g(H^n H^{n*})(i, i)$ is the element (i, i) of the matrix $g(H^n H^{n*})$. Suppose that for large n , the operator at the right hand side of this equation can be replaced with HH^* . Denote by $g(HH^*)(i, i)$ the (i, i) element of the matrix representation of $g(HH^*)$. Observe now that since HH^* is ergodic, then by [29, Theorem 2.7], its *resolution of the identity* E_λ is an ergodic projection for every $\lambda \in \mathbb{R}$. As a result, the operator $g(HH^*) = \int g(\lambda) dE_\lambda$ is also ergodic. Thus,

$$\begin{aligned} \frac{1}{2n+1} \sum_{i=-n}^n g(HH^*)(\omega)(i, i) &= \frac{1}{2n+1} \sum_{i=-n}^n \langle e_i, g(HH^*)(\omega) e_i \rangle \\ &= \frac{1}{2n+1} \sum_{i=-n}^n \langle e_0, U^i g(HH^*)(\omega) U^{-i} e_0 \rangle \\ &= \frac{1}{2n+1} \sum_{i=-n}^n \langle e_0, g(HH^*)(B^i \omega) e_0 \rangle \\ &\xrightarrow[n \rightarrow \infty]{\text{a.s.}} \mathbb{E} g(HH^*)(0, 0) = \mathbb{E} \int g(\lambda) \langle e_0, E(d\lambda) e_0 \rangle \end{aligned}$$

by the ergodic theorem, provided the $(0, 0)$ element $g(HH^*)(0, 0)$ of $g(HH^*)$ is integrable. It follows that $\mu(d\lambda) = \mathbb{E} \langle e_0, E(d\lambda) e_0 \rangle$.

By taking $g(\lambda) = (\lambda - z)^{-1}$ where $z \in \mathbb{C}_+$, we get that the Stieltjes transform of μ coincides with $\mathbb{E} g(HH^*)(0, 0) = \mathbb{E} Q_{HH^*}(0, 0)(z)$.

It remains to generalize these results to the case where the $H_T(k, \ell)$ are $N \times T$ matrices. The closedness of H_T and the existence of the self-adjoint operator $H_T H_T^*$ are direct generalizations of the scalar case. Writing $\mathbf{H}_T(\omega, k) = [H_T(k, k + \ell)]_{\ell=-L(T)}^{L(T)}$, the shift B defined by the equation $\mathbf{H}_T(B\omega, k) = \mathbf{H}_T(\omega, k + 1)$ is ergodic. Since the operator $H_T H_T^*$ satisfies the identity $(H_T H_T^*)(B\omega) = U^N (H_T H_T^*)(\omega) U^{-N}$, it is also ergodic. The proof of Proposition 4.1 goes along nearly without modification. Let us just check that the Stieltjes transform of μ_T is $N^{-1} \text{Tr} \mathbb{E} Q(0, 0)(z)$. Writing $g(H_T H_T^*)_{p,q}(k, \ell) = \langle e_{kN+p}, g(H_T H_T^*) e_{\ell N+q} \rangle$, we get by mimicking the derivation above that

$$\frac{1}{(2n+1)N} \sum_{r=0}^{N-1} \sum_{i=-n}^n g(H_T H_T^*)_{r,r}(i, i) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \frac{1}{N} \text{Tr} \mathbb{E} g(H_T H_T^*)(0, 0).$$

Taking $g(\lambda) = (\lambda - z)^{-1}$, we get the result.

Proof of Theorem 2.1

As is well known (see *e.g.* [37]),

$$\begin{aligned} I(S; (Y, H)) &= \limsup_n \frac{1}{2n+1} \left(I(S^n; H^n) + I(S^n; Y^n | H^n) \right) \\ &= \limsup_n \frac{1}{2n+1} I(S^n; Y^n | H^n) \\ &= \limsup_n \frac{1}{2n+1} \mathbb{E} \log \det(H^n H^{n*} + I_{(2n+1)N}) \end{aligned}$$

by the independence of S^n and H^n . Let ν_T^n be the NCM of the random matrix $H^n H^{n*}$. Then

$$\frac{1}{(2n+1)N} \log \det(H^n H^{n*} + I) = \int \log(1 + \lambda) \nu_T^n(d\lambda).$$

The measure ν_T^n satisfies

$$\begin{aligned} \int \lambda \nu_T^n(d\lambda) &= \frac{\text{Tr}(H^n H^{n*})}{(2n+1)N} \\ &= \frac{1}{(2n+1)N} \sum_{m=-n}^n \sum_{\ell=-L}^L \text{Tr} H(m, m-\ell) H(m, m-\ell)^* \\ &\leq \frac{2}{(2n+1)N} \sum_{m=-n}^n \sum_{\ell=-L}^L (\text{Tr} A(\ell) A(\ell)^* + \text{Tr} X(m, m-\ell) X(m, m-\ell)^*) \\ &= \frac{2}{N} \sum_{\ell=-L}^L \text{Tr} A(\ell) A(\ell)^* \\ &\quad + 2 \sum_{\ell=-L}^L \frac{\phi_T(\ell)^2}{T} \frac{1}{(2n+1)N} \sum_{m=-n}^n \text{Tr} W(m, m-\ell) W(m, m-\ell)^*. \end{aligned} \quad (6)$$

Since the process $\{W(m, m-\ell) W(m, m-\ell)^*\}_{m \in \mathbb{Z}}$ is ergodic, it holds that

$$\frac{1}{2n+1} \sum_{m=-n}^n W(m, m-\ell) W(m, m-\ell)^* \xrightarrow[n \rightarrow \infty]{\text{a.s.}} T I_N.$$

We therefore get that

$$\limsup_n \int \lambda \nu_T^n(d\lambda) \leq 2 \left(\sum_{\ell=-L}^L \|A_T(\ell)\|^2 + \sigma_T^2 \right) \quad \text{w.p. 1.}$$

Let us consider an elementary event in the probability one set where ν_T^n converges weakly to μ_T by Prop. 4.1 and where $\sup_n \int \lambda d\nu_T^n < \infty$. By uniform integrability, we get that $\int \log(1 + \lambda) \nu_T^n(d\lambda) \rightarrow \int \log(1 + \lambda) \mu_T(d\lambda) < \infty$ on the elementary event we just considered. Consequently, $\int \log(1 + \lambda) \nu_T^n(d\lambda) \rightarrow \int \log(1 + \lambda) \mu_T(d\lambda) < \infty$ almost surely.

Getting back to the inequality (6), it is clear that $\sup_n \mathbb{E} \int \lambda \nu_T^n(d\lambda) \leq 2(\sum_{\ell} \|A_T(\ell)\|^2 + \sigma_T^2)$. Therefore, $\sup_n \mathbb{E} (\int \log(1 + \lambda) \nu_T^n(d\lambda))^2 < \infty$ which shows that

$$\mathbb{E} \int \log(1 + \lambda) \nu_T^n(d\lambda) \xrightarrow[n \rightarrow \infty]{} \int \log(1 + \lambda) \mu_T(d\lambda) < \infty$$

by uniform integrability.

5 Proof of Theorem 2.2

We start by showing the uniqueness of the solution $(\varphi_T(\cdot, z), \tilde{\varphi}_T(\cdot, z))$ satisfying the integrability and the non-negativity conditions. We then provide a constructive existence proof, and prove in passing that the constructed solution satisfies the Stieltjes transform properties provided in the statement.

5.1 The uniqueness

The general idea of the uniqueness proof can be found in *e.g.* [11] or [18]. Applying this idea to the case of this article requires some specific work. We fix $z \in \mathbb{C}_+$ and we re-denote herein $\varphi(f, z)$, $\mathbf{S}(f, z)$, etc. as $\varphi(f)$, $\mathbf{S}(f)$, etc. for simplicity.

We start with some lemmas. Assume that $(\varphi_T(f), \tilde{\varphi}_T(f))$ is a solution. Then

Lemma 5.1. *The solution $(\varphi_T(f), \tilde{\varphi}_T(f))$ satisfies the equation*

$$\begin{aligned} \begin{bmatrix} \Im \varphi(f) \\ \Im(z \tilde{\varphi}(f)) \end{bmatrix} &= \int_0^1 \begin{bmatrix} K_{11}^{\varphi, \tilde{\varphi}}(f, u) & K_{12}^{\varphi, \tilde{\varphi}}(f, u) \\ K_{21}^{\varphi, \tilde{\varphi}}(f, u) & K_{22}^{\varphi, \tilde{\varphi}}(f, u) \end{bmatrix} \begin{bmatrix} \Im \varphi(u) \\ \Im(z \tilde{\varphi}(u)) \end{bmatrix} du \\ &+ \Im z \begin{bmatrix} F_1^{\varphi, \tilde{\varphi}}(f) \\ F_2^{\varphi, \tilde{\varphi}}(f) \end{bmatrix}, \quad f \in [0, 1], \end{aligned}$$

where

$$\begin{aligned} K_{11}^{\varphi, \tilde{\varphi}}(f, u) &= \frac{\text{Tr} \mathbf{S}(f)(\mathbf{A}\mathbf{A}^*)(f)\mathbf{S}(f)^*}{T|1 + \zeta(f)|^2} \sigma^2 \gamma(u - f), \\ K_{12}^{\varphi, \tilde{\varphi}}(f, u) &= \frac{\text{Tr} \mathbf{S}(f)\mathbf{S}(f)^*}{T} \sigma^2 \gamma(f - u), \\ K_{21}^{\varphi, \tilde{\varphi}}(f, u) &= \frac{\text{Tr} |z|^2 \tilde{\mathbf{S}}(f)\tilde{\mathbf{S}}(f)^*}{T} \sigma^2 \gamma(u - f), \\ K_{22}^{\varphi, \tilde{\varphi}}(f, u) &= \frac{\text{Tr} \tilde{\mathbf{S}}(f)(\mathbf{A}^*\mathbf{A})(f)\tilde{\mathbf{S}}(f)^*}{T|1 + \tilde{\zeta}(f)|^2} \sigma^2 \gamma(f - u), \end{aligned}$$

$$\begin{aligned} F_1^{\varphi, \tilde{\varphi}}(f) &= \frac{\text{Tr} \mathbf{S}(f)\mathbf{S}(f)^*}{T}, \\ F_2^{\varphi, \tilde{\varphi}}(f) &= \frac{\text{Tr} \tilde{\mathbf{S}}(f)(\mathbf{A}^*\mathbf{A})(f)\tilde{\mathbf{S}}(f)^*}{T|1 + \tilde{\zeta}(f)|^2}, \end{aligned}$$

$$\zeta(f) = \sigma^2 \gamma(-f) \star \varphi(f), \quad \text{and} \quad \tilde{\zeta}(f) = \sigma^2 \gamma(f) \star \tilde{\varphi}(f).$$

Proof. We have

$$\begin{aligned}
\Im\varphi(f) &= \frac{\varphi - \bar{\varphi}}{2i} \\
&= \frac{1}{2iT} \operatorname{Tr} \mathbf{S}(\mathbf{S}^{-*} - \mathbf{S}^{-1})\mathbf{S}^* \\
&= \frac{1}{2iT} \operatorname{Tr} \mathbf{S} \left((z - \bar{z})I + \sigma^2 \boldsymbol{\gamma}(f) \star (z\tilde{\varphi}(f) - \bar{z}\tilde{\varphi}(f))I \right. \\
&\quad \left. + \frac{(\mathbf{A}\mathbf{A}^*)(f)}{1 + \tilde{\zeta}(f)} - \frac{(\mathbf{A}\mathbf{A}^*)(f)}{1 + \zeta(f)} \right) \mathbf{S}^* \\
&= \Im z \frac{\operatorname{Tr} \mathbf{S}(f)\mathbf{S}^*(f)}{T} + (\Im(z\tilde{\zeta}(f))) \frac{\operatorname{Tr} \mathbf{S}(f)\mathbf{S}^*(f)}{T} \\
&\quad + (\Im\zeta(f)) \frac{\operatorname{Tr} \mathbf{S}(f)(\mathbf{A}\mathbf{A}^*)(f)\mathbf{S}^*(f)}{T|1 + \zeta(f)|^2}.
\end{aligned}$$

By a similar derivation, we also have

$$\begin{aligned}
\Im(z\tilde{\varphi}(f)) &= \frac{1}{2iT} \operatorname{Tr} |z|^2 \tilde{\mathbf{S}}(\bar{z}^{-1}\tilde{\mathbf{S}}^{-*} - z^{-1}\tilde{\mathbf{S}}^{-1})\tilde{\mathbf{S}}^* \\
&= (\sigma^2 \boldsymbol{\gamma}(-f) \star \Im\varphi(f)) \frac{|z|^2 \operatorname{Tr} \tilde{\mathbf{S}}(f)\tilde{\mathbf{S}}^*(f)}{T} + \Im z \frac{\operatorname{Tr} \tilde{\mathbf{S}}(f)(\mathbf{A}^*\mathbf{A})(f)\tilde{\mathbf{S}}^*(f)}{T|1 + \tilde{\zeta}(f)|^2} \\
&\quad + (\sigma^2 \boldsymbol{\gamma}(f) \star \Im(z\tilde{\varphi}(f))) \frac{\operatorname{Tr} \tilde{\mathbf{S}}(f)(\mathbf{A}^*\mathbf{A})(f)\tilde{\mathbf{S}}^*(f)}{T|1 + \tilde{\zeta}(f)|^2}
\end{aligned}$$

hence the result. \square

Lemma 5.2. For any $f \in [0, 1]$, $\Im\varphi(f) > 0$, $F_1^{\varphi, \tilde{\varphi}}(f) > 0$, $0 < \Im(z\tilde{\varphi}(f)) < C$, and $F_2^{\varphi, \tilde{\varphi}}(f) > C' \operatorname{Tr}(\mathbf{A}^*\mathbf{A})(f)$ where C and C' are positive numbers that do not depend on f .

Proof. Let $\lambda_0, \dots, \lambda_{N-1}$ be the eigenvalues of $(\mathbf{A}\mathbf{A}^*)(f)$. We start by establishing an upper bound on $|-z(1 + \zeta) + \lambda_i/(1 + \zeta)|$. Observing that $\max_f \boldsymbol{\gamma}(f) \leq \sum |\boldsymbol{\gamma}_T(\ell)| \leq \mathbf{g}$ and writing $\tilde{\mathbf{b}}(z) = \int_0^1 |\tilde{\varphi}(f, z)| df < \infty$, we have $|z(1 + \zeta)| \leq |z|(1 + \sigma^2 \mathbf{g}\tilde{\mathbf{b}}(z))$. Moreover, $|1 + \zeta| = |z + z\zeta|/|z| \geq \Im z/|z|$ since $\Im(z\varphi) \geq 0$ by assumption, and $\boldsymbol{\gamma}(f) \geq 0$. Therefore, $|\lambda_i/(1 + \zeta)| \leq |z|\mathbf{a}^2/\Im z$ and $|-z(1 + \zeta) + \lambda_i/(1 + \zeta)| \leq |z|(1 + \sigma^2 \mathbf{g}\tilde{\mathbf{b}}(z) + \mathbf{a}^2/\Im z)$. The inequality $F_1^{\varphi, \tilde{\varphi}}(f) > 0$ follows readily. Moreover,

$$\Im\varphi(f) = \frac{1}{T} \sum_{i=0}^{N-1} \frac{\Im(-\bar{z}(1 + \tilde{\zeta}) + \frac{\lambda_i}{1 + \tilde{\zeta}})}{|-z(1 + \tilde{\zeta}) + \frac{\lambda_i}{1 + \tilde{\zeta}}|^2} \geq \frac{1}{T} \sum_{i=0}^{N-1} \frac{\Im z}{|-z(1 + \tilde{\zeta}) + \frac{\lambda_i}{1 + \tilde{\zeta}}|^2} > 0. \quad (7)$$

The inequalities on $\Im(z\tilde{\varphi}(f))$ and $F_2^{\varphi, \tilde{\varphi}}(f)$ are proven by similar means. \square

Lemma 5.3. The functions φ and $\tilde{\varphi}$ are continuous on $[0, 1]$.

Proof. Since φ is integrable and $\boldsymbol{\gamma}$ is continuous, ζ is continuous by the dominated convergence theorem, and similarly for $\tilde{\zeta}$. Moreover, $(\mathbf{A}\mathbf{A}^*)$ is continuous as a trigonometric polynomial. It follows that the matrix functions \mathbf{S} and $\tilde{\mathbf{S}}$ are continuous, hence the result. \square

To establish the uniqueness of the solution, we observe that by setting

$$J_1^{\varphi, \tilde{\varphi}}(f, u) = \frac{K_{11}^{\varphi, \tilde{\varphi}}(f, u)}{\Im \varphi(f)} \Im \varphi(u) + \frac{K_{12}^{\varphi, \tilde{\varphi}}(f, u)}{\Im \varphi(f)} \Im(z \tilde{\varphi}(u))$$

and

$$J_2^{\varphi, \tilde{\varphi}}(f, u) = \frac{K_{21}^{\varphi, \tilde{\varphi}}(f, u)}{\Im(z \tilde{\varphi}(f))} \Im \varphi(u) + \frac{K_{22}^{\varphi, \tilde{\varphi}}(f, u)}{\Im(z \tilde{\varphi}(f))} \Im(z \tilde{\varphi}(u))$$

on $[0, 1]^2$, Lemmas 5.1 and 5.2 show that

$$0 < \int_0^1 J_1^{\varphi, \tilde{\varphi}}(f, u, z) du \leq 1 - \Im z \frac{F_1^{\varphi, \tilde{\varphi}}(f, z)}{\Im \varphi(f, z)} < 1 \quad (8)$$

and

$$0 < \int_0^1 J_2^{\varphi, \tilde{\varphi}}(f, u, z) du \leq 1 - \Im z \frac{F_2^{\varphi, \tilde{\varphi}}(f, z)}{\Im(z \tilde{\varphi}(f, z))} \leq 1 - C \operatorname{Tr}(\mathbf{A}^* \mathbf{A})(f) \quad (9)$$

where $C > 0$. Assume now that $(\varphi, \tilde{\varphi})$ and $(\varphi', \tilde{\varphi}')$ are two solutions. Denote respectively by $(\mathbf{S}, \tilde{\mathbf{S}})$ and $(\mathbf{S}', \tilde{\mathbf{S}}')$ the matrix functions associated to these solutions as in the statement of the proposition. Write $\Delta \varphi(f) = \varphi(f) - \varphi'(f)$ and $z \Delta \tilde{\varphi}(f) = z \tilde{\varphi}(f) - z \tilde{\varphi}'(f)$. Mimicking the proof of Lemma 5.1, we get

$$\begin{aligned} \Delta \varphi(f) &= \frac{1}{T} \operatorname{Tr} \mathbf{S}(\mathbf{S}'^{-1} - \mathbf{S}^{-1}) \mathbf{S}' \\ &= \int_0^1 (\mathcal{K}_{11}(f, u) \Delta \varphi(u) + \mathcal{K}_{12}(f, u) z \Delta \tilde{\varphi}(u)) du \end{aligned}$$

and

$$\begin{aligned} z \Delta \tilde{\varphi}(f) &= \frac{z}{T} \operatorname{Tr} \tilde{\mathbf{S}}(\tilde{\mathbf{S}}'^{-1} - \tilde{\mathbf{S}}^{-1}) \tilde{\mathbf{S}}' \\ &= \int_0^1 (\mathcal{K}_{21}(f, u) \Delta \varphi(u) + \mathcal{K}_{22}(f, u) z \Delta \tilde{\varphi}(u)) du \end{aligned}$$

where

$$\begin{aligned} \mathcal{K}_{11}(f, u) &= \frac{\operatorname{Tr} \mathbf{S}(f)(\mathbf{A} \mathbf{A}^*)(f) \mathbf{S}'(f)}{T(1 + \sigma^2 \gamma(-f) \star \varphi(f))(1 + \sigma^2 \gamma(-f) \star \varphi'(f))} \sigma^2 \gamma(u - f), \\ \mathcal{K}_{12}(f, u) &= \frac{\operatorname{Tr} \mathbf{S}(f) \mathbf{S}'(f)}{T} \sigma^2 \gamma(f - u), \\ \mathcal{K}_{21}(f, u) &= \frac{\operatorname{Tr} z^2 \tilde{\mathbf{S}}(f) \tilde{\mathbf{S}}'(f)}{T} \sigma^2 \gamma(u - f), \\ \mathcal{K}_{22}(f, u) &= \frac{\operatorname{Tr} \tilde{\mathbf{S}}(f)(\mathbf{A}^* \mathbf{A})(f) \tilde{\mathbf{S}}'(f)}{T(1 + \sigma^2 \gamma(f) \star \tilde{\varphi}(f))(1 + \sigma^2 \gamma(f) \star \tilde{\varphi}'(f))} \sigma^2 \gamma(f - u). \end{aligned} \quad (10)$$

Let

$$\varepsilon(f) = \frac{\Delta \varphi(f)}{\sqrt{\Im \varphi(f) \Im \varphi'(f)}} \quad \text{and} \quad \tilde{\varepsilon}(f) = \frac{z \Delta \tilde{\varphi}(f)}{\sqrt{\Im(z \tilde{\varphi}(f)) \Im(z \tilde{\varphi}'(f))}}.$$

Using the inequality $|\text{Tr}(M_1 M_2)| \leq \sqrt{\text{Tr}(M_1 M_1^*)} \sqrt{\text{Tr}(M_2 M_2^*)}$ for any two matrices M_1 and M_2 with compatible dimensions, we get

$$|\varepsilon(f)| \leq \int_0^1 \left(\frac{\sqrt{K_{11}^{\varphi, \tilde{\varphi}}(f, u) \Im \varphi(u)}}{\sqrt{\Im \varphi(f)}} \frac{\sqrt{K_{11}^{\varphi', \tilde{\varphi}'}(f, u) \Im \varphi'(u)}}{\sqrt{\Im \varphi'(f)}} |\varepsilon(u)| \right. \\ \left. + \frac{\sqrt{K_{12}^{\varphi, \tilde{\varphi}}(f, u) \Im(z \tilde{\varphi}(u))}}{\sqrt{\Im \varphi(f)}} \frac{\sqrt{K_{12}^{\varphi', \tilde{\varphi}'}(f, u) \Im(z \tilde{\varphi}'(u))}}{\sqrt{\Im \varphi'(f)}} |\tilde{\varepsilon}(u)| \right) du$$

Let $\varepsilon = \sup_{f \in [0, 1]} (|\varepsilon(f)| \vee |\tilde{\varepsilon}(f)|)$. Using the inequality $\sqrt{ac} + \sqrt{bd} \leq \sqrt{a+b} \sqrt{c+d}$ where $a, b, c, d \geq 0$ along with the Cauchy-Schwarz inequality, we get

$$|\varepsilon(f)| \leq \left(\int_0^1 J_1^{\varphi, \tilde{\varphi}}(f, u) du \right)^{1/2} \left(\int_0^1 J_1^{\varphi', \tilde{\varphi}'}(f, u) du \right)^{1/2} \varepsilon$$

We shall assume that $\varepsilon > 0$ and obtain a contradiction. By the last inequality and (8), we have $|\varepsilon(f)| < \varepsilon$. Lemmas 5.2 and 5.3 show that $\varepsilon(f)$ is continuous on $[0, 1]$, hence $\sup_{f \in [0, 1]} |\varepsilon(f)| < \varepsilon$. Turning to $\tilde{\varepsilon}(f)$, we get by a similar argument that

$$|\tilde{\varepsilon}(f)| \leq \int_0^1 \left(\frac{\sqrt{K_{21}^{\varphi, \tilde{\varphi}}(f, u) \Im \varphi(u)}}{\sqrt{\Im(z \tilde{\varphi}(f))}} \frac{\sqrt{K_{21}^{\varphi', \tilde{\varphi}'}(f, u) \Im \varphi'(u)}}{\sqrt{\Im(z \tilde{\varphi}'(f))}} |\varepsilon(u)| \right. \\ \left. + \frac{\sqrt{K_{22}^{\varphi, \tilde{\varphi}}(f, u) \Im(z \tilde{\varphi}(u))}}{\sqrt{\Im(z \tilde{\varphi}(f))}} \frac{\sqrt{K_{22}^{\varphi', \tilde{\varphi}'}(f, u) \Im(z \tilde{\varphi}'(u))}}{\sqrt{\Im(z \tilde{\varphi}'(f))}} |\tilde{\varepsilon}(u)| \right) du.$$

If $(\mathbf{A}^* \mathbf{A})(f) = 0$, then $K_{22}^{\varphi, \tilde{\varphi}}(f, u) = K_{22}^{\varphi', \tilde{\varphi}'}(f, u) = 0$ for any $u \in [0, 1]$. But then, Inequalities (9) show that $|\tilde{\varepsilon}(f)| \leq \sup_{u \in [0, 1]} |\varepsilon(u)| < \varepsilon$. On the other hand, if $\text{Tr}(\mathbf{A}^* \mathbf{A})(f) > 0$, then

$$|\tilde{\varepsilon}(f)| \leq \left(\int_0^1 J_2^{\varphi, \tilde{\varphi}}(f, u) du \right)^{1/2} \left(\int_0^1 J_2^{\varphi', \tilde{\varphi}'}(f, u) du \right)^{1/2} \varepsilon < \varepsilon$$

by 9 again. Therefore, we also have $\sup_{f \in [0, 1]} |\tilde{\varepsilon}(f)| < \varepsilon$ by the continuity of $\tilde{\varepsilon}(f)$, which leads to a contradiction. Uniqueness is established.

5.2 The existence

Starting with the functions $\varphi^{(0)}(f, z) = -(N/T)z^{-1}$ and $\tilde{\varphi}^{(0)}(f, z) = -z^{-1}$ on $[0, 1] \times \mathbb{C}_+$, define the recursion

$$(\varphi^{(k+1)}(f, z), \tilde{\varphi}^{(k+1)}(f, z)) = h_z(\varphi^{(k)}(f, z), \tilde{\varphi}^{(k)}(f, z))$$

with

$$h_z(\varphi^{(k)}(f, z), \tilde{\varphi}^{(k)}(f, z)) = \left(T^{-1} \text{Tr} \mathbf{S}^{(k)}(f, z), T^{-1} \text{Tr} \tilde{\mathbf{S}}^{(k)}(f, z) \right),$$

$$\mathbf{S}^{(k)}(f, z) = \left[-z(1 + \tilde{\zeta}^{(k)}(f, z))I + \frac{(\mathbf{A} \mathbf{A}^*)(f)}{1 + \zeta^{(k)}(f, z)} \right]^{-1},$$

$$\tilde{\mathbf{S}}^{(k)}(f, z) = \left[-z(1 + \zeta^{(k)}(f, z))I + \frac{(\mathbf{A}^* \mathbf{A})(f)}{1 + \tilde{\zeta}^{(k)}(f, z)} \right]^{-1},$$

$$\zeta^{(k)}(f, z) = \sigma^2 \gamma(-f) \star \varphi^{(k)}(f, z), \quad \text{and} \quad \tilde{\zeta}^{(k)}(f, z) = \sigma^2 \gamma(f) \star \tilde{\varphi}^{(k)}(f, z).$$

We shall show that the sequence $(\varphi^{(k)}(f, z), \tilde{\varphi}^{(k)}(f, z))$ converges on $[0, 1] \times \mathbb{C}_+$, and that the limit $(\varphi(f, z), \tilde{\varphi}(f, z))$ is the solution of the system described in the statement of Theorem 2.2.

Lemma 5.4. *The following facts hold true:*

i) For any $k \in \mathbb{N}$ and any $f \in [0, 1]$, the functions $\varphi^{(k)}(f, \cdot)$ and $\tilde{\varphi}^{(k)}(f, \cdot)$ are holomorphic on \mathbb{C}_+ and satisfy $|\varphi^{(k)}(f, z)| \leq \mathbf{c}/\Im z$ and $|\tilde{\varphi}^{(k)}(f, z)| \leq 1/\Im z$. Furthermore, $\Im \varphi^{(k)}(f, z)$, $\Im(z\varphi^{(k)}(f, z))$, $\Im \tilde{\varphi}^{(k)}(f, z)$, and $\Im(z\tilde{\varphi}^{(k)}(f, z))$ are all positive for $z \in \mathbb{C}_+$.

ii) In the region

$$\mathcal{R} = \left\{ z \in \mathbb{C}_+ : (\mathbf{c} \vee 1)\sigma^2 \left(\frac{\mathbf{a}^2 |z|^2}{(\Im z)^4} + \frac{|z|}{(\Im z)^2} \right) < \frac{1}{2} \right\},$$

it holds that

$$\begin{aligned} |\varphi^{(k+1)}(f, z) - \varphi^{(k)}(f, z)| \vee |\tilde{\varphi}^{(k+1)}(f, z) - \tilde{\varphi}^{(k)}(f, z)| \\ < \frac{1}{2} (|\varphi^{(k)}(f, z) - \varphi^{(k-1)}(f, z)| \vee |\tilde{\varphi}^{(k)}(f, z) - \tilde{\varphi}^{(k-1)}(f, z)|). \end{aligned}$$

iii) For any $f, f' \in [0, 1]$ and any $z \in \mathbb{C}_+$,

$$\begin{aligned} |\varphi^{(k)}(f, z) - \varphi^{(k)}(f', z)| \vee |\tilde{\varphi}^{(k)}(f, z) - \tilde{\varphi}^{(k)}(f', z)| \\ \leq \frac{(\mathbf{c} \vee 1)|z|}{(\Im z)^3} \left(\left(\sigma^2 + \frac{|z|\sigma^2\mathbf{a}^2}{(\Im z)^2} \right) \int_0^1 |\gamma(f - f' + u) - \gamma(u)| du + 4\pi L\mathbf{a}^2 |f - f'| \right). \end{aligned}$$

Proof. The analyticity as well as the inequalities stated in *i*) are trivially true for $\varphi^{(0)}$ and for $\tilde{\varphi}^{(0)}$. Assume they are for $\varphi^{(k)}$ and $\tilde{\varphi}^{(k)}$. Denoting again by $\lambda_0, \dots, \lambda_{N-1}$ the eigenvalues of $(\mathbf{A}\mathbf{A}^*)(f)$, we have

$$|-z(1 + \tilde{\zeta}^{(k)}) + \lambda_i/(1 + \zeta^{(k)})| \geq \Im z + \Im(z\tilde{\zeta}^{(k)}) + \lambda_i \Im \zeta^{(k)} / |1 + \zeta^{(k)}| \geq \Im z. \quad (11)$$

Consequently, $\varphi^{(k+1)}$ and $\tilde{\varphi}^{(k+1)}$ are holomorphic on $z \in \mathbb{C}_+$ for any $f \in [0, 1]$ and they satisfy $|\varphi^{(k+1)}(f, z)| \leq \mathbf{c}/\Im z$ and $|\tilde{\varphi}^{(k+1)}(f, z)| \leq 1/\Im z$. By reproducing the inequalities (7), we also show that $\Im \varphi^{(k+1)} > 0$. The other inequalities are proven similarly, which establishes Item *i*).

We now show *ii*). Writing $\Delta^{(k)}(f, z) = \varphi^{(k)}(f, z) - \varphi^{(k-1)}(f, z)$ and $\tilde{\Delta}^{(k)}(f, z) = \tilde{\varphi}^{(k)}(f, z) - \tilde{\varphi}^{(k-1)}(f, z)$, we get by a derivation similar to the one made in Section 5.1 that

$$\begin{aligned} \Delta^{k+1}(f, z) &= \int_0^1 (\mathcal{K}_{11}^{(k)}(f, u, z)\Delta^{(k)}(u, z) + z\mathcal{K}_{12}^{(k)}(f, u, z)\tilde{\Delta}^{(k)}(u, z)) du \\ \tilde{\Delta}^{k+1}(f, z) &= \int_0^1 (z^{-1}\mathcal{K}_{21}^{(k)}(f, u, z)\Delta^{(k)}(u, z) + \mathcal{K}_{22}^{(k)}(f, u, z)\tilde{\Delta}^{(k)}(u, z)) du \end{aligned}$$

where the $\mathcal{K}_{ij}^{(k)}$ have the same expressions as the \mathcal{K}_{ij} in (10) except that the \mathbf{S}, \dots there are replaced with $\mathbf{S}^{(k)}, \dots$ and the \mathbf{S}', \dots are replaced with $\mathbf{S}^{(k-1)}, \dots$. Using the bounds established in *i*), we readily obtain

$$\begin{aligned} |\mathcal{K}_{11}^{(k)}(f, u, z)| &\leq \mathbf{c}\sigma^2\mathbf{a}^2|z|^2\gamma(u-f)/(\Im z)^4, \\ |z\mathcal{K}_{12}^{(k)}(f, u, z)| &\leq \mathbf{c}\sigma^2|z|\gamma(f-u)/(\Im z)^2, \\ |z^{-1}\mathcal{K}_{21}^{(k)}(f, u, z)| &\leq \sigma^2|z|\gamma(u-f)/(\Im z)^2, \\ |\mathcal{K}_{22}^{(k)}(f, u, z)| &\leq \sigma^2\mathbf{a}^2|z|^2\gamma(f-u)/(\Im z)^4. \end{aligned}$$

Item *ii*) follows from these inequalities after integrating over u .

To show *iii*), we start by writing $\varphi^{(k)}(f) - \varphi^{(k)}(f') = T^{-1} \operatorname{Tr} \mathbf{S}^{(k)}(f) (\mathbf{S}^{(k)}(f')^{-1} - \mathbf{S}^{(k)}(f)^{-1}) \mathbf{S}^{(k)}(f')$ where we omit the parameter z . Developing the right hand side, we get

$$\begin{aligned} |\tilde{\zeta}^{(k)}(f) - \tilde{\zeta}^{(k)}(f')| &= \sigma^2 \left| \int_0^1 (\gamma(f-u) - \gamma(f'-u)) \tilde{\varphi}^{(k)}(u) du \right| \\ &\leq \frac{\sigma^2}{\Im z} \int_0^1 |\gamma(f-f'+u) - \gamma(u)| du \end{aligned}$$

and similarly for $|\zeta^{(k)}(f) - \zeta^{(k)}(f')|$. Furthermore,

$$\begin{aligned} \frac{(\mathbf{A}\mathbf{A}^*)(f)}{1 + \zeta^{(k)}(f)} - \frac{(\mathbf{A}\mathbf{A}^*)(f')}{1 + \zeta^{(k)}(f')} &= \frac{(\mathbf{A}\mathbf{A}^*)(f) - (\mathbf{A}\mathbf{A}^*)(f')}{1 + \zeta^{(k)}(f)} - \frac{(\zeta^{(k)}(f) - \zeta^{(k)}(f'))(\mathbf{A}\mathbf{A}^*)(f')}{(1 + \zeta^{(k)}(f))(1 + \zeta^{(k)}(f'))}. \end{aligned}$$

Recalling that $\mathbf{A}(f)$ is a trigonometric matrix polynomial, it can be checked that $\|(\mathbf{A}\mathbf{A}^*)(f) - (\mathbf{A}\mathbf{A}^*)(f')\| \leq 4\pi L\alpha^2 |f - f'|$. Item *iii*) is then obtained by a small calculation. \square

For any $z \in \mathbb{C}_+$, the map h_z defined before the statement of Lemma 5.4 is a map on the Banach space of the continuous $[0, 1] \rightarrow \mathbb{C}_+^2$ functions endowed with the maximum of the supremum norms of the two components. Showing the existence of the solution amounts to showing the existence of a fixed point of h_z that satisfies the inequalities provided in the statement of Theorem 2.2. By Lemma 5.4-*ii*), the map h_z is a contraction for any $z \in \mathcal{R}$. Therefore, $(\varphi^{(k)}(\cdot, z), \tilde{\varphi}^{(k)}(\cdot, z))$ converges for any $z \in \mathcal{R}$, and the limit that we denote $(\varphi(\cdot, z), \tilde{\varphi}(\cdot, z))$ is a fixed point of h_z .

We now show that from every sequence of integers, one can extract a subsequence $v(k)$ such that $\varphi^{(v(k))}(f, z)$ and $\tilde{\varphi}^{(v(k))}(f, z)$ converge uniformly on the compact subsets of $[0, 1] \times \mathbb{C}_+$. To that end, we start by showing that on every compact set $K \subset \mathbb{C}_+$, the families $\varphi^{(k)}(f, z)$ and $\tilde{\varphi}^{(k)}(f, z)$ are equicontinuous on $[0, 1] \times K$. Let $2d > 0$ be the distance from K to \mathbb{R} , and let $z, z' \in K$ be such that $|z - z'| \leq d/2$. Denote by \mathcal{C} the positively oriented circle with center z and radius d . Since $\varphi^{(k)}(f, z)$ is holomorphic, we have by Cauchy's formula

$$\varphi^{(k)}(f, z) - \varphi^{(k)}(f, z') = \frac{z - z'}{2i\pi} \oint_{\mathcal{C}} \frac{\varphi^{(k)}(f, w)}{(w - z)(w - z')} dw$$

which shows that

$$|\varphi^{(k)}(f, z) - \varphi^{(k)}(f, z')| \leq \frac{2c}{d^2} |z - z'|. \quad (12)$$

Given two couples $(f, z), (f', z') \in [0, 1] \times K$, we have $|\varphi^{(k)}(f, z) - \varphi^{(k)}(f', z')| \leq |\varphi^{(k)}(f, z) - \varphi^{(k)}(f', z)| + |\varphi^{(k)}(f', z) - \varphi^{(k)}(f', z')|$. In conjunction with Inequality (12), Lemma 5.4-*iii*) shows then that the families $\{\varphi^{(k)}(f, z)\}_{k \in \mathbb{N}}$ and $\{\tilde{\varphi}^{(k)}(f, z)\}_{k \in \mathbb{N}}$ are equicontinuous on the compact set $[0, 1] \times K$ (note that the integral at the right hand side of the inequality in Lemma 5.4-*iii*) converges to zero as $f - f' \rightarrow 0$ by the dominated convergence theorem). These families are moreover bounded on $[0, 1] \times K$. Therefore, by the Arzelà-Ascoli theorem, one can extract from every sequence of integers a subsequence $p(k)$ such that $\varphi^{(p(k))}(f, z)$ and $\tilde{\varphi}^{(p(k))}(f, z)$ converge uniformly on $[0, 1] \times K$. Considering a sequence of compact subsets of

\mathbb{C}_+ who is increasing with respect to the inclusion and whose union is \mathbb{C}_+ , one can establish the existence of the sequence $v(k)$ by the diagonal process.

Note that the respective limit functions $\varphi_v(f, z)$ and $\tilde{\varphi}_v(f, z)$ of $\varphi^{(v(k))}(f, z)$ and $\tilde{\varphi}^{(v(k))}(f, z)$ are continuous in the variable f and holomorphic in the variable z . They also satisfy $|\varphi_v(f, z)| \leq c/\Im z$, $|\tilde{\varphi}_v(f, z)| \leq 1/\Im z$, and $\Im \varphi_v$, $\Im \tilde{\varphi}_v$, $\Im(z\varphi_v)$, and $\Im(z\tilde{\varphi}_v)$ are nonnegative by Lemma 5.4-*i*) and a passage to the limit. Furthermore, $h_z(\varphi_v(f, z), \tilde{\varphi}_v(f, z))$ exists and is holomorphic on \mathbb{C}_+ .

For $z \in \mathcal{R}$, $(\varphi_v(f, z), \tilde{\varphi}_v(f, z))$ clearly coincides with $(\varphi(f, z), \tilde{\varphi}(f, z))$. Thus $(\varphi_v(f, z), \tilde{\varphi}_v(f, z)) - h_z(\varphi_v(f, z), \tilde{\varphi}_v(f, z)) = 0$ on \mathcal{R} , hence on all \mathbb{C}_+ by analyticity.

We just showed that $(\varphi^{(k)}, \tilde{\varphi}^{(k)})$ converges to the unique solution $(\varphi, \tilde{\varphi})$ of the system in the statement of Theorem 2.2. The functions φ and $\tilde{\varphi}$ are analytical in the variable z and they satisfy $|\varphi(f, z)| \leq c/\Im z$ and $|\tilde{\varphi}(f, z)| \leq 1/\Im z$ along with the non-negativity conditions. Hence, for any $f \in [0, 1]$, $(T/N)\varphi(f, z)$ and $\tilde{\varphi}(f, z)$ are the Stieltjes transforms of finite positive measures carried by $[0, \infty)$. By Lemma 5.2, $\Im \varphi$, $\Im \tilde{\varphi}$, $\Im(z\varphi)$, and $\Im(z\tilde{\varphi})$ are positive. It remains to show that these measures are probability measures. We have

$$-\frac{T}{N}iy\varphi(f, iy) = \frac{1}{N} \sum_{i=1}^N \frac{-iy}{-iy(1 + \sigma^2\gamma(f) \star \tilde{\varphi}(f, iy)) + \frac{\lambda_i}{1 + \sigma^2\gamma(-f) \star \varphi(f, iy)}}.$$

For any i , the denominator behaves as $-iy$ when $y \rightarrow \infty$. Hence $-(T/N)iy\varphi(f, iy) \rightarrow 1$ as $y \rightarrow \infty$, which establishes the result for $(T/N)\varphi$. The result for $\tilde{\varphi}$ is proven similarly.

6 Proofs of Theorems 2.3 and 2.4

We start with some notations. Recall that $Q_T(z) = (H_T H_T^* - zI)$. Denote by $\tilde{Q}(z) = (H_T^* H_T - zI)^{-1}$ the resolvent of the self-adjoint operator $H_T^* H_T$. The complex number $z \in \mathbb{C}_+$ will be always written as $z = x + iy$. When there is no ambiguity, the parameter z or the index T will be omitted for notational simplicity. We shall denote as $B_{m,n}(k, \ell)$ or $[B]_{m,n}(k, \ell)$ the (m, n) element of the (k, ℓ) block of the matrix representation of B , the size of the blocks being clear from the context. The block itself will be denoted as $B(k, \ell)$ or $[B](k, \ell)$.

6.1 Identities, inequalities and basic tools

We start with a some technical results.

Lemma 6.1. *The operator $Q(z)H$ can be continuously extended to a bounded operator on $l^2(\mathbb{Z})$ (since the domain of H is only dense in $l^2(\mathbb{Z})$) with adjoint $H^*Q(\bar{z})$. The norm of this operator satisfies $\|Q(z)H\|^2 \leq (y + |z|)/y^2$.*

Proof. Given any $a \in \mathcal{D}(H)$ and any $b \in l^2(\mathbb{Z})$, we have $\langle Q(z)Ha, b \rangle = \langle Ha, Q(\bar{z})b \rangle = \langle a, H^*Q(\bar{z})b \rangle$, the second inequality being due to the fact that $Q(\bar{z})b \in \mathcal{D}(HH^*) \subset \mathcal{D}(H^*)$. Therefore, $H^*Q(\bar{z})$ is the adjoint of $Q(z)H$, and since it is defined on all $l^2(\mathbb{Z})$, the operator $Q(z)H$ can be extended to a bounded operator. We furthermore have

$$\|Q(z)HH^*Q(\bar{z})\| = \|Q(z)(I + \bar{z}Q(\bar{z}))\| \leq \frac{y + |z|}{y^2}$$

hence the last result. \square

The following bounds will also be often used in the proof without mention:

$$\begin{aligned} |Q_{n,n'}(k, \ell)(z)| &\leq \|Q(k, \ell)(z)\| \leq \|Q(z)\| \leq \frac{1}{y}, \\ \sum_{k \in \mathbb{Z}} \sum_{n=0}^{N-1} |Q_{n,p}(k, \ell)|^2 &= [Q^*Q]_{p,p}(\ell, \ell) \leq \frac{1}{y^2} \\ \sum_{\ell \in \mathbb{Z}} \sum_{p=0}^{N-1} |[QH]_{n,p}(m, \ell)|^2 &= [QHH^*Q^*]_{p,p}(m, m) \leq \frac{y + |z|}{y^2} \end{aligned}$$

Lemma 6.2. $H^*Q(z)H = I + z\tilde{Q}(z)$ and $H\tilde{Q}H^* = I + zQ(z)$ by continuous extensions of the left hand members.

Proof. The proof is based on the polar decomposition $H = U|H|$ of H [22, §VI.2.7]. We recall that $|H| = (H^*H)^{1/2}$, that U is a partial isometry from $\overline{\mathcal{R}(|H|)}$ to $\overline{\mathcal{R}(H)}$ (here $\mathcal{R}(\cdot)$ is the range of an operator and $\overline{\mathcal{R}(\cdot)}$ the closure of this range), and that $H^* = |H|U^*$. Letting $a \in \mathcal{D}(H)$ and $b = Q(z)Ha$, we have $Ha = HH^*b - zb$, or equivalently $b = z^{-1}H(H^*b - a)$ that we rewrite as $b = Uh$ with $h = z^{-1}|H|(|H|U^*b - a)$. Since $h \in \mathcal{R}(|H|)$, $U^*Uh = h$, therefore $h = z^{-1}|H|(|H|h - a)$ or equivalently $h = (|H|^2 - zI)^{-1}|H|a = |H|(|H|^2 - zI)^{-1}a$. Finally, the vector $w = H^*Q(z)Ha = H^*b$ satisfies $w = |H|U^*Uh = |H|^2(|H|^2 - zI)^{-1}a = H^*H\tilde{Q}(z)a = (H^*H - z)\tilde{Q}(z)a + z\tilde{Q}(z)a = a + z\tilde{Q}(z)a$. \square

The following lemma can be shown by direct calculation.

Lemma 6.3 (Differentiation formulas). *Given an integer $M \in \mathbb{N}$, let H^M be the random operator represented by the matrix $H^M = [\mathbb{1}_{|k| \leq M} \mathbb{1}_{|\ell| \leq M} H(k, \ell)]$ and let $Q^M(z) = (H^M H^{M*} - zI)^{-1} = [[Q_{p,q}^M(k, \ell)(z)]_{p,q=0}^{N-1}]_{k,\ell \in \mathbb{Z}}$ be the resolvent of the self adjoint operator $H^M H^{M*}$. For any integers i, j, k, l, n, t, p, q satisfying $-M \leq i, j, k, \ell \leq M$, $|i - j| \leq L$, $0 \leq n, t, p \leq N - 1$, and $0 \leq q \leq T - 1$, it holds that*

$$\begin{aligned} \frac{\partial Q_{n,t}^M(k, \ell)}{\partial X_{p,q}(i, j)} &= -Q_{n,p}^M(k, i)[H^{M*}Q^M]_{q,t}(j, \ell) \quad \text{and} \\ \frac{\partial Q_{n,t}^M(k, \ell)}{\partial \tilde{X}_{p,q}(i, j)} &= -[Q^M H^M]_{n,q}(k, j)Q_{p,t}^M(i, \ell). \end{aligned}$$

The following basic property of convolution operators is well known [6].

Proposition 6.1. *A convolution operator $B = [B_{k-\ell}]_{k,\ell \in \mathbb{Z}}$ where the blocks are $N \times N$ matrices is a bounded operator on $l^2(\mathbb{Z})$ if and only if there exists a bounded $N \times N$ matrix function $\mathbf{B}(f)$ on $[0, 1]$ such that*

$$B_k = \int_0^1 \exp(-2i\pi kf) \mathbf{B}(f) df \quad k \in \mathbb{Z},$$

i.e., the B_k are the Fourier coefficients of the 1-periodic function equal to $\mathbf{B}(f)$ on $[0, 1]$.

Let us identify the function $\mathbf{B}(f)$ on $[0, 1]$ with the multiplication operator $\mathbf{B} : \mathcal{L}^2([0, 1] \rightarrow \mathbb{C}^N) \rightarrow \mathcal{L}^2([0, 1] \rightarrow \mathbb{C}^N)$ who sends $\mathbf{g}(f)$ to $(\mathbf{B}\mathbf{g})(f) = \mathbf{B}(f)\mathbf{g}(f)$. Clearly, the spectral norm of this operator coincides with $\sup_{f \in [0,1]} \|\mathbf{B}(f)\|$. Let

$$\mathcal{F}_N : \mathcal{L}^2([0, 1] \rightarrow \mathbb{C}^N) \rightarrow l^2(\mathbb{Z}), \quad \mathcal{F}_N(\mathbf{g}) = \left(\int_0^1 \exp(-2i\pi kf) \mathbf{g}(f) df \right)_{k \in \mathbb{Z}} = (g_k)_{k \in \mathbb{Z}}$$

be the isometric operator that sends \mathbf{g} to the sequence of its Fourier coefficients. Then it holds that $B = \mathcal{F}_N \mathbf{B} \mathcal{F}_N^*$.

We now introduce the two basic tools used in the proof of Theorem 2.3.

Proposition 6.2. *Let $\xi \in \mathbb{C}^M$ be a complex Gaussian centered vector such that $\mathbb{E}\xi\xi^T = 0$ and $\mathbb{E}\xi\xi^* = \Xi$. Let $\Gamma : \mathbb{C}^M \rightarrow \mathbb{R}$ be a C^1 function polynomially bounded together with its derivatives (when seeing Γ as a function from \mathbb{R}^{2M} to \mathbb{R}). Then*

$$\mathbb{E}[\xi_i \Gamma(\xi)] = \sum_{m=1}^M [\Xi]_{im} \mathbb{E} \left[\frac{\partial \Gamma(\xi)}{\partial \bar{\xi}_m} \right]$$

This proposition can be proven by an integration by parts. We shall call it for this reason the IP formula.

Proposition 6.3. *Let $\xi \in \mathbb{C}^M$ and $\Gamma : \mathbb{C}^M \rightarrow \mathbb{R}$ be as in the statement of Proposition 6.2, and let $\nabla_z \Gamma = [\partial \Gamma / \partial z_1, \dots, \partial \Gamma / \partial z_M]^T$ and $\nabla_{\bar{z}} \Gamma = [\partial \Gamma / \partial \bar{z}_1, \dots, \partial \Gamma / \partial \bar{z}_M]^T$. Then the following inequality holds true:*

$$\text{Var} \Gamma(\xi) \leq \mathbb{E}[\nabla_z \Gamma(\xi)^T \Xi \overline{\nabla_z \Gamma(\xi)}] + \mathbb{E}[(\nabla_{\bar{z}} \Gamma(\xi))^* \Xi \nabla_{\bar{z}} \Gamma(\xi)] .$$

This inequality is known as the Poincaré-Nash (PN) inequality. For a proof, see [31, Prop. 2.1.6].

6.2 System of equations controlling $\mathbb{E}Q(k, k+m)$ and $\mathbb{E}\tilde{Q}(k, k+m)$.

We now enter the core of the proof, which consists in showing that the $\mathbb{E}Q(k, \ell)$ and the $\mathbb{E}\tilde{Q}(k, \ell)$ satisfy a perturbed infinite system of equations. Our derivations will be greatly simplified by the fact that due to the ergodicity of H_T , the operators $\mathbb{E}Q(z)$ and $\mathbb{E}\tilde{Q}(z)$ are bounded convolution operators for any $z \in \mathbb{C}_+$.

We start with some variance controls. The following lemma is proven in Appendix A. The proof is based on the use of Proposition 6.3 along with an operator truncation argument.

Lemma 6.4. *For any $k, \ell \in \mathbb{Z}$ and any $n \in \{0, \dots, N-1\}$ and $t \in \{0, \dots, T-1\}$, we have*

$$\text{Var} Q_{p,q}(k, \ell) \leq \frac{2}{T} \frac{\sigma^2 \mathbf{g}(|z| + y)}{y^4}, \quad \text{Var} \text{Tr} Q(k, \ell) \leq \frac{2c\sigma^2 \mathbf{g}(y + |z|)}{y^4},$$

$$\text{and } \text{Var}[H^* Q]_{p,q}(k, \ell) \leq \frac{2}{T} \frac{\sigma^2 \mathbf{g}(|z| + y)^2}{y^4}.$$

Our purpose is now to find an equation satisfied by $\mathbb{E}Q_{p,q}(k, \ell)$. To that end, we start by writing $\mathbb{E}[[HH^* Q]_{p,q}(k, \ell)] = \chi_1 + \chi_2 + \chi_3 + \chi_4$ where χ_1 , χ_2 and χ_3 are given by the finite sums

$$\chi_1 = \sum_{i,j} \sum_{n,t=0}^{N-1, T-1} \mathbb{E}[X_{p,t}(k, i) \bar{X}_{n,t}(j, i) Q_{nq}(j, \ell)],$$

$$\chi_2 = \sum_{i,j} \sum_{n,t=0}^{N-1, T-1} \mathbb{E}[X_{p,t}(k, i) \bar{A}_{n,t}(j, i) Q_{nq}(j, \ell)],$$

$$\chi_3 = \sum_{i,j} \sum_{n,t=0}^{N-1, T-1} \mathbb{E}[A_{p,t}(k, i) \bar{X}_{n,t}(j, i) Q_{nq}(j, \ell)],$$

and where

$$\chi_4 = \mathbb{E}[[AA^*Q]_{p,q}(k, \ell)].$$

Starting with χ_1 , the idea is to develop $\mathbb{E}[X_{p,t}(k, i)\bar{X}_{n,t}(j, i)Q_{nq}(j, \ell)]$ with the help of the IP formula. Thanks to this formula, we expect that

$$\begin{aligned} & \mathbb{E}[X_{p,t}(k, i)\bar{X}_{n,t}(j, i)Q_{nq}(j, \ell)] \\ &= \sum_{u,v} \mathbb{E}[X_{p,t}(k, i)\bar{X}_{p,t}(u, v)] \mathbb{E}\left[\frac{\partial(\bar{X}_{n,t}(j, i)Q_{nq}(j, \ell))}{\partial\bar{X}_{p,t}(u, v)}\right] \\ &= \sum_{u,v} \frac{\gamma_T(k-u)\phi_T(k-i)^2}{T} \delta_{k-i, u-v} \left\{ \delta_{p,n}\delta_{u,j}\delta_{v,i} \mathbb{E}[Q_{nq}(j, \ell)] \right. \\ & \quad \left. - \mathbb{E}[\bar{X}_{n,t}(j, i)[QH]_{n,t}(j, v)Q_{p,q}(u, \ell)] \right\} \\ &= \sum_r \frac{\gamma_T(r)\phi_T(k-i)^2}{T} \left\{ \delta_{p,n}\delta_{k,j}\delta_{r,0} \mathbb{E}[Q_{nq}(j, \ell)] \right. \\ & \quad \left. - \mathbb{E}[\bar{X}_{n,t}(j, i)[QH]_{n,t}(j, i-r)Q_{p,q}(k-r, \ell)] \right\}. \end{aligned} \quad (13)$$

Note that the vector ξ in the statement of Proposition 6.2 is finite dimensional. Therefore, Equation (13) needs to be justified. Given an integer $M \in \mathbb{N}$, let H^M and Q^M be as in the statement of Lemma 6.3. When $M \geq |k| \vee |i| \vee |j|$, we get by applying Proposition 6.2 and repeating the derivation above that

$$\begin{aligned} & \mathbb{E}[X_{p,t}(k, i)\bar{X}_{n,t}(j, i)Q_{nq}^M(j, \ell)] \\ &= \sum_{r=k \vee i - M}^{k \wedge i + M} \frac{\gamma_T(r)\phi_T(k-i)^2}{T} \left\{ \delta_{p,n}\delta_{k,j}\delta_{r,0} \mathbb{E}[Q_{nq}^M(j, \ell)] \right. \\ & \quad \left. - \mathbb{E}[\bar{X}_{n,t}(j, i)[Q^M H^M]_{n,t}(j, i-r)Q_{p,q}^M(k-r, \ell)] \right\}. \end{aligned} \quad (14)$$

The argument provided in the proof of Lemma 6.4 shows that $Q_{nq}^M(j, \ell) \xrightarrow{\text{a.s.}} Q_{nq}(j, \ell)$ as $M \rightarrow \infty$. Since $|X_{p,t}(k, i)\bar{X}_{n,t}(j, i)Q_{nq}^M(j, \ell)| \leq |X_{p,t}(k, i)\bar{X}_{n,t}(j, i)|/y$, it holds by the dominated convergence theorem that $\mathbb{E}[X_{p,t}(k, i)\bar{X}_{n,t}(j, i)Q_{nq}^M(j, \ell)] \rightarrow \mathbb{E}[X_{p,t}(k, i)\bar{X}_{n,t}(j, i)Q_{nq}(j, \ell)]$ as $M \rightarrow \infty$. Using Lemma 6.1 and the summability of γ , we can show by a similar argument that the right hand side of (14) converges to the right hand side of (13). Equation (13) is shown.

Turning to χ_2 , we obtain by a similar derivation

$$\chi_2 = - \sum_{i,j,r} \sum_{n,t=0}^{N-1, T-1} \frac{\gamma_T(r)\phi_T(k-i)^2}{T} \mathbb{E}[\bar{A}_{n,t}(j, i)[QH]_{n,t}(j, i-r)Q_{p,q}(k-r, \ell)].$$

Considering $\chi_1 + \chi_2$, and taking the sum over j and n , we get

$$\begin{aligned}
\chi_1 + \chi_2 &= \sum_{i,r} \sum_{t=0}^{T-1} \frac{\gamma_T(r) \phi_T(k-i)^2}{T} \left\{ \delta_{r,0} \mathbb{E}[Q_{p,q}(k, \ell)] - \right. \\
&\quad \left. \mathbb{E}[[H^*QH]_{t,t}(i, i-r) Q_{p,q}(k-r, \ell)] \right\} \\
&= \sum_{r,s} \sum_{t=0}^{T-1} \frac{\gamma_T(r) \phi_T(s)^2}{T} \mathbb{E} \left[Q_{p,q}(k-r, \ell) \times \right. \\
&\quad \left. (I_{t,t}(k-s, k-s-r) - [H^*QH]_{t,t}(k-s, k-s-r)) \right] \\
&= -z \sum_{r,s} \gamma_T(r) \phi_T(s)^2 \mathbb{E} \left[Q_{pq}(k-r, \ell) \frac{\text{Tr} \tilde{Q}(k-s, k-s-r)}{T} \right]
\end{aligned}$$

where the third equality is due to Lemma 6.2. In the second equation, we considered that $\delta_{r,0} = I_{t,t}(k-s, k-s-r)$, the (t, t) element of the $(k-s, k-s-r)$ block of the “ $N \times N$ block-matrix representation” of the identity operator. Turning to χ_3 , we have

$$\begin{aligned}
\chi_3 &= \sum_{i,j,u,v} \sum_{n,t} A_{p,t}(k, i) \mathbb{E}[X_{n,t}(u, v) \bar{X}_{n,t}(j, i)] \mathbb{E} \left[\frac{\partial Q_{n,q}(j, \ell)}{\partial X_{n,t}(u, v)} \right] \\
&= - \sum_{i,j,u,v} \sum_{n,t} A_{p,t}(k, i) \frac{\gamma_T(u-j) \phi_T(j-i)^2}{T} \delta_{j-i, u-v} \mathbb{E}[Q_{n,n}(j, u) [H^*Q]_{t,q}(v, \ell)] \\
&= - \sum_{i,r,s} \sum_t \gamma_T(r) \phi_T(s)^2 A_{p,t}(k, i) \mathbb{E} \left[\frac{\text{Tr} Q(s+i, s+i+r)}{T} [H^*Q]_{t,q}(i+r, \ell) \right].
\end{aligned}$$

We now “decouple” the terms within the expectations in the expressions of $\chi_1 + \chi_2$ and χ_3 by using the inequality $|\mathbb{E}XY - \mathbb{E}X\mathbb{E}Y| = |\mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]| \leq (\text{Var } X)^{1/2} (\text{Var } Y)^{1/2}$ along with Lemma 6.4. By the ergodicity of H , terms such as $\mathbb{E}[Q_{p,q}(k, \ell)]$, $\mathbb{E}[[H^*Q]_{p,q}(k, \ell)]$, or $\mathbb{E}[[AH^*Q]_{p,q}(k, \ell)]$ depend on $k - \ell$ only. With a small notation abuse, we shall henceforth denote the first of these terms as $\mathbb{E}[Q_{p,q}(k, \ell)]$ or $\mathbb{E}[Q_{p,q}(k - \ell)]$ interchangeably, and similarly for the other terms. With these notations, we get

$$\chi_1 + \chi_2 = -z \sigma_T^2 \sum_r \gamma_T(r) \mathbb{E}[Q_{pq}(k-r, \ell)] \left[\frac{\text{Tr} \mathbb{E} \tilde{Q}(r)}{T} \right] + \varepsilon_{p,q}(k - \ell)$$

where

$$|\varepsilon_{p,q}(k - \ell)| \leq \frac{2\sqrt{c}\sigma^4 \mathbf{g}^2}{T^{3/2}} \frac{|z|(|z| + y)}{y^4},$$

and

$$\chi_3 = -\sigma_T^2 \sum_{i,r} \sum_t \gamma_T(r) A_{p,t}(k, i) \left[\frac{\text{Tr} \mathbb{E} Q(-r)}{T} \right] \mathbb{E}[[H^*Q]_{t,q}(i, \ell-r)] + \varepsilon'_{p,q}(k - \ell)$$

where ε' satisfies by Lemma 6.4

$$|\varepsilon'_{p,q}(k - \ell)| \leq \frac{2\sqrt{c}\sigma^4 \mathbf{g}^2 (y + |z|)^{3/2}}{y^4} \frac{1}{T^{3/2}} \sum_{i,t} |A_{p,t}(k, i)| \leq \frac{2\sqrt{c}\sigma^4 \mathbf{g}^2 \mathbf{a} (y + |z|)^{3/2}}{y^4} \frac{1}{T}$$

since $\sum_{i,t} |A_{p,t}(k,i)| \leq \sqrt{T} \sum_i \|A(k,i)\| \leq \sqrt{T} \mathbf{a}$. Recalling that A is a convolution operator, we have

$$\chi_3 = -\sigma_T^2 \sum_r \gamma_T(r) \left[\frac{\text{Tr} \mathbb{E} Q(-r)}{T} \right] \mathbb{E} [[AH^*Q]_{p,q}(k+r-\ell)] + \varepsilon'_{p,q}(k-\ell).$$

By the identity $HH^*Q(z) = (HH^* - zI)Q(z) + zQ = I + zQ(z)$, we get

$$\begin{aligned} \mathbb{E} Q_{pq}(k-\ell) &= -z^{-1} I_{p,q}^N(k-\ell) + z^{-1}(\chi_1 + \chi_2) + z^{-1} \mathbb{E} [AH^*Q]_{p,q}(k-\ell), \\ \mathbb{E} [AH^*Q]_{p,q}(k-\ell) &= \chi_3 + \chi_4. \end{aligned}$$

Specifically,

$$\begin{aligned} \mathbb{E} Q_{p,q}(k) &= -z^{-1} I_{p,q}^N(k) - \sigma_T^2 \sum_r \gamma_T(r) \left[\frac{\text{Tr} \mathbb{E} \tilde{Q}(r)}{T} \right] \mathbb{E} Q_{pq}(k-r) \\ &\quad + z^{-1} \mathbb{E} [AH^*Q]_{p,q}(k) + \varepsilon_{p,q}(k), \\ \mathbb{E} [AH^*Q]_{p,q}(k) &= -\sigma_T^2 \sum_r \gamma_T(-r) \left[\frac{\text{Tr} \mathbb{E} Q(r)}{T} \right] \mathbb{E} [AH^*Q]_{p,q}(k-r) \\ &\quad + \mathbb{E} [AA^*Q]_{p,q}(k) + \varepsilon'_{p,q}(k). \end{aligned}$$

Similar derivations lead to the identities

$$\begin{aligned} \mathbb{E} \tilde{Q}_{p,q}(k) &= -z^{-1} I_{p,q}^T(k) - \sigma_T^2 \sum_r \gamma_T(-r) \left[\frac{\text{Tr} \mathbb{E} Q(r)}{T} \right] \mathbb{E} \tilde{Q}_{p,q}(k-r) \\ &\quad + z^{-1} \mathbb{E} [A^*H\tilde{Q}]_{p,q}(k) + \tilde{\varepsilon}_{p,q}(k), \\ \mathbb{E} [A^*H\tilde{Q}]_{p,q}(k) &= -\sigma_T^2 \sum_r \gamma_T(r) \left[\frac{\text{Tr} \mathbb{E} \tilde{Q}(r)}{T} \right] \mathbb{E} [A^*H\tilde{Q}]_{p,q}(k-r) \\ &\quad + \mathbb{E} [A^*A\tilde{Q}]_{p,q}(k) + \tilde{\varepsilon}'_{p,q}(k) \end{aligned}$$

where

$$|\tilde{\varepsilon}_{p,q}(k)| \leq \frac{2\sigma^4 \mathbf{g}^2 |z|(|z|+y)}{y^4} \frac{1}{T^{3/2}} \quad \text{and} \quad |\tilde{\varepsilon}'_{p,q}(k)| \leq \frac{2\sigma^4 \mathbf{g}^2 \mathbf{a} (y+|z|)^{3/2}}{y^4} \frac{1}{T}.$$

6.3 Passage to the frequency domain; End of the proof

For any $z \in \mathbb{C}_+$, we can identify the matrix functions $\mathbf{S}_T(\cdot, z)$ and $\tilde{\mathbf{S}}_T(\cdot, z)$ defined in the statement of Theorem 2.2 with the multiplication operators $(\mathbf{S}_T(\cdot, z)\mathbf{g})(f) = \mathbf{S}_T(f, z)\mathbf{g}(f)$ and $(\tilde{\mathbf{S}}_T(\cdot, z)\tilde{\mathbf{g}})(f) = \tilde{\mathbf{S}}_T(f, z)\tilde{\mathbf{g}}(f)$ on $\mathcal{L}^2([0, 1] \rightarrow \mathbb{C}^N)$ and $\mathcal{L}^2([0, 1] \rightarrow \mathbb{C}^T)$ respectively. Given a function $\mathbf{g} \in \mathcal{L}^2([0, 1] \rightarrow \mathbb{C}^N)$ such that $\|\mathbf{g}\| > 0$, one can show by derivations similar to those of Section 5 that

$$\langle (\Im \mathbf{S}(\cdot, z))\mathbf{g}, \mathbf{g} \rangle = \int_0^1 \frac{\mathbf{g}(f)^* \mathbf{S}(f) (\mathbf{S}(f)^{-*} - \mathbf{S}(f)^{-1}) \mathbf{S}(f)^* \mathbf{g}(f)}{2i} df > 0$$

for any $z \in \mathbb{C}_+$, where $\Im \mathbf{S} = (\mathbf{S} - \mathbf{S}^*)/(2i)$. The self-adjoint operator $\Im \mathbf{S}(\cdot, z)$ is therefore positive for any $z \in \mathbb{C}_+$. Similarly, one can show that $\Im(z\mathbf{S}(\cdot, z))$ is positive on \mathbb{C}_+ , and so is the case for $\Im \tilde{\mathbf{S}}(\cdot, z)$ and $\Im(z\tilde{\mathbf{S}}(\cdot, z))$.

Let $S_T(z) = [S_T(k - \ell)(z)]_{k, \ell \in \mathbb{Z}}$ and $\tilde{S}_T(z) = [\tilde{S}_T(k - \ell)(z)]_{k, \ell \in \mathbb{Z}}$ the convolution operators obtained through the isometries

$$S_T(z) = \mathcal{F}_N \mathbf{S}_T(\cdot, z) \mathcal{F}_N^* \quad \text{and} \quad \tilde{S}_T(z) = \mathcal{F}_T \tilde{\mathbf{S}}_T(\cdot, z) \mathcal{F}_T^*.$$

Then we have the following result:

Lemma 6.5. *For any $z \in \mathbb{C}_+$, the matrix blocks of the operators S_T and \tilde{S}_T satisfy the equations (where the parameter z is omitted)*

$$\begin{aligned} S_T(k) &= -z^{-1} \delta_{k,0} I_N - \sigma_T^2 \sum_r \gamma_T(r) \tilde{\varphi}_T(r) S_T(k-r) + z^{-1} P_T(k), \\ P_T(k) &= -\sigma_T^2 \sum_r \gamma_T(-r) \varphi_T(r) P_T(k-r) + [A_T A_T^* S_T](k), \\ \tilde{S}_T(k) &= -z^{-1} \delta_{k,0} I_T - \sigma_T^2 \sum_r \gamma_T(-r) \varphi_T(r) \tilde{S}_T(k-r) + z^{-1} \tilde{P}_T(k), \\ \tilde{P}_T(k) &= -\sigma_T^2 \sum_r \gamma_T(r) \tilde{\varphi}_T(r) \tilde{P}_T(k-r) + [A_T^* A_T \tilde{S}_T](k) \end{aligned}$$

where

$$\varphi_T(r) = \frac{\text{Tr } S_T(r)}{T} \quad \text{and} \quad \tilde{\varphi}_T(r) = \frac{\text{Tr } \tilde{S}_T(r)}{T}.$$

The convolution operators $P_T = [P_T(k - \ell)]_{k, \ell \in \mathbb{Z}}$ and $\tilde{P}_T = [\tilde{P}_T(k - \ell)]_{k, \ell \in \mathbb{Z}}$ satisfy the inequality $\|P(z)\| \vee \|\tilde{P}(z)\| \leq \mathbf{a}^2 |z| / (\Im z)^2$. In addition, $\|S_T\| \vee \|\tilde{S}_T\| \leq (\Im z)^{-1}$ and furthermore, $\Im S_T$, $\Im \tilde{S}_T$, $\Im(z S_T)$ and $\Im(z \tilde{S}_T)$ are positive.

There exist four unique bounded convolution operators $S_T = [S_T(k - \ell)]_{k, \ell \in \mathbb{Z}}$, $\tilde{S}_T = [\tilde{S}_T(k - \ell)]_{k, \ell \in \mathbb{Z}}$, $P_T = [P_T(k - \ell)]_{k, \ell \in \mathbb{Z}}$ and $\tilde{P}_T = [\tilde{P}_T(k - \ell)]_{k, \ell \in \mathbb{Z}}$ satisfying the equations above in association with the positivity constraints on $\Im S_T$, $\Im \tilde{S}_T$, $\Im(z S_T)$ and $\Im(z \tilde{S}_T)$.

Proof. Equation (2) can be rewritten as

$$\begin{aligned} -z \mathbf{S}(f) - z \sigma^2 (\gamma(f) \star \tilde{\varphi}(f)) \mathbf{S}(f) + \mathbf{P}(f) &= I, \\ (1 + \sigma^2 \gamma(-f) \star \varphi(f)) \mathbf{P}(f) &= (\mathbf{A} \mathbf{A}^*)(f) \mathbf{S}(f). \end{aligned}$$

Taking the Fourier transforms of $\mathbf{S}(f)$ and $\mathbf{P}(f)$, we recover the first two equations in the statement. The other two equations are obtained similarly. We know that the multiplication operator \mathbf{S} satisfies $\|\mathbf{S}\| \leq 1/\Im z$, $\Im \mathbf{S} > 0$ and $\Im(z \mathbf{S}) > 0$. From the expression of $\mathbf{P}(f)$ we see that the associated multiplication operator has its norm bounded by $\mathbf{a}^2 |z| / (\Im z)^2$. Since \mathcal{F}_N and \mathcal{F}_T are isometries, the norms and positivity constraints are deduced at once.

Finally, since a bounded convolution operator is uniquely determined by its associated multiplication operator, and since the operators \mathbf{S} and $\tilde{\mathbf{S}}$ are uniquely determined by the conditions of Theorem 2.2, we obtain the uniqueness alluded to in the last part of the statement. \square

We are now in position to establish the first statement of Theorem 2.3. Since

$$\frac{\text{Tr } S_T(0)(z)}{N} = \frac{1}{N} \int_0^1 \text{Tr } \mathbf{S}_T(f) df = \frac{T}{N} \int_0^1 \varphi_T(f, z) df,$$

what we need to show is that $N^{-1} \text{Tr}(\mathbb{E}Q_T(0)(z) - S_T(0)(z)) \rightarrow 0$ for any $z \in \mathbb{C}_+$. To that end, we start by showing that

$$\Delta_T = \sup_{k \in \mathbb{Z}} \frac{\|\mathbb{E}Q(k) - S(k)\|_{\text{fro}}}{\sqrt{T}} \quad \text{and} \quad \tilde{\Delta}_T = \sup_{k \in \mathbb{Z}} \frac{\|\mathbb{E}\tilde{Q}(k) - \tilde{S}(k)\|_{\text{fro}}}{\sqrt{T}}$$

converge to zero as $T \rightarrow \infty$ in a certain region of \mathbb{C}_+ , where $\|\cdot\|_{\text{fro}}$ is the Frobenius norm. Indeed, we have

$$\begin{aligned} \mathbb{E}Q(k) - S(k) &= -\sigma_T^2 \sum_r \gamma_T(r) \frac{\text{Tr} \mathbb{E}\tilde{Q}(r)}{T} (\mathbb{E}Q(k-r) - S(k-r)) \\ &\quad - \sigma_T^2 \sum_r \gamma_T(r) \frac{\text{Tr} \mathbb{E}\tilde{Q}(r) - \text{Tr} \tilde{S}(r)}{T} S(k-r) \\ &\quad + z^{-1} (\mathbb{E}[AH^*Q](k) - P(k)) + E(k), \\ \mathbb{E}[AH^*Q](k) - P(k) &= -\sigma_T^2 \sum_r \gamma_T(-r) \frac{\text{Tr} \mathbb{E}Q(r)}{T} (\mathbb{E}[AH^*Q](k-r) - P(k-r)) \\ &\quad - \sigma_T^2 \sum_r \gamma_T(-r) \frac{\text{Tr} \mathbb{E}Q(r) - \text{Tr} S(r)}{T} P(k-r) \\ &\quad + [AA^*(\mathbb{E}Q - S)](k) + E'(k) \end{aligned}$$

where $\|E(k)\|_{\text{fro}} = \left(\sum_{p,q} |\varepsilon_{p,q}(k)|^2 \right)^{1/2} \leq C(z)/\sqrt{T}$ and $\|E'(k)\|_{\text{fro}} \leq C'(z)$, with $C(z)$ and $C'(z)$ being bounded on the compact subsets of \mathbb{C}_+ . Notice that

$$\begin{aligned} \left| \frac{\text{Tr}[\mathbb{E}Q - S](r)}{T} \right| &\leq \frac{\sum_p |\mathbb{E}Q - S]_{p,p}(r)|}{T} \\ &\leq \frac{(N \sum_p |\mathbb{E}Q - S]_{p,p}(r)|^2)^{1/2}}{T} \leq \sqrt{c} \frac{\|\mathbb{E}Q - S](r)\|_{\text{fro}}}{\sqrt{T}} \end{aligned} \quad (15)$$

and similarly, $|\text{Tr}[\mathbb{E}\tilde{Q} - \tilde{S}](r)/T| \leq \|\mathbb{E}\tilde{Q} - \tilde{S}](r)\|_{\text{fro}}/\sqrt{T}$. We also have $\|S(k)\|_{\text{fro}}^2/T \leq c\|S(k)\|^2 \leq c/y^2$ and $\|P(k)\|_{\text{fro}}^2/T \leq c\|P(k)\|^2 \leq ca^4|z|^2/y^4$.

Writing

$$\Delta'_T = \sup_k \frac{\|\mathbb{E}[AH^*Q](k) - P(k)\|_{\text{fro}}}{\sqrt{T}},$$

we easily get from the expression of $\mathbb{E}Q(k) - S(k)$ above

$$\Delta_T \leq \frac{\sigma^2 g}{y} \Delta_T + \frac{\sigma^2 g \sqrt{c}}{y} \tilde{\Delta}_T + \frac{\Delta'_T}{|z|} + \frac{C(z)}{T}.$$

Using the norm inequality $\|UV\|_{\text{fro}} \leq \|U\| \|V\|_{\text{fro}}$ where U and V are two matrices, we observe that

$$\begin{aligned} \|[AA^*(\mathbb{E}Q - S)](k)\|_{\text{fro}} &= \left\| \sum_{i,j} A(i)A(j)^* [\mathbb{E}Q - S](j-i+k) \right\|_{\text{fro}} \\ &\leq \sum_{i,j} \|A(i)\| \|A(j)\| \|\mathbb{E}Q - S](j-i+k)\|_{\text{fro}} \end{aligned}$$

and we get from the expression of $\mathbb{E}[AH^*Q](k) - P(k)$ that

$$\Delta'_T \leq \frac{\sigma^2 \mathbf{g} \mathbf{c}}{y} \Delta'_T + \frac{\sigma^2 \mathbf{g} \sqrt{\mathbf{c} \mathbf{a}^2} |z|}{y^2} \Delta_T + \mathbf{a}^2 \Delta_T + \frac{C'(z)}{\sqrt{T}},$$

and we have two similar inequalities for $\tilde{\Delta}_T$ and $\tilde{\Delta}'_T = \sup_k \|\mathbb{E}[A^* H \tilde{Q}](k) - \tilde{P}(k)\|_{\text{fro}} / \sqrt{T}$. These four inequalities can be written in a compact form as

$$\mathbf{\Delta}_T(z) \leq \mathbf{M}(z) \mathbf{\Delta}_T(z) + \frac{1}{\sqrt{T}} \mathbf{E}(z)$$

where $\mathbf{\Delta}_T(z) = [\Delta_T, \Delta'_T, \tilde{\Delta}_T, \tilde{\Delta}'_T]^T$, the inequality is element wise, $\mathbf{M}(z)$ is a 4×4 matrix independent on T whose norm converges to zero as $y \rightarrow \infty$, and $\mathbf{E}(z)$ is vector whose norm is bounded on any compact set of \mathbb{C}_+ .

This result coupled with the inequalities (15) shows that for y large enough, it holds that $N^{-1} \text{Tr}(\mathbb{E}Q_T(0)(z) - S_T(0)(z)) \rightarrow 0$. Recall that $N^{-1} \text{Tr} \mathbb{E}Q_T(0)(z)$ and $N^{-1} \text{Tr} S_T(0)(z)$ are holomorphic and bounded on the compact subsets of $\mathbb{C} - \mathbb{R}_+$. By the normal family theorem, from any sequence of integers depending on T , we can extract a subsequence $v(T)$ such that $N(v(T))^{-1} \text{Tr}(\mathbb{E}Q_{v(T)}(0)(z) - S_{v(T)}(0)(z)) \rightarrow 0$ on the compact subsets of $\mathbb{C} - \mathbb{R}_+$. But since the limit is zero for large enough y , it is identically zero on $\mathbb{C} - \mathbb{R}_+$. The convergence (4) in the statement of Theorem 2.3 is proven. To establish the second part of the statement, we need the tightness of π_T .

Lemma 6.6. *The sequence π_T is tight.*

Proof. For any small $\varepsilon > 0$, any $T \in \mathbb{N}$, any $f \in [0, 1]$ and any eigenvalue $\lambda_{i,T}(f)$ of $(\mathbf{A}_T \mathbf{A}_T^*)(f)$, writing $\zeta_T(f, y) = \sigma_T^2 \gamma_T(f) \star \tilde{\varphi}_T(f, y)$ and $\tilde{\zeta}_T(f, y) = \sigma_T^2 \gamma_T(-f) \star \varphi_T(f, y)$, we have

$$\left| \frac{-iy}{-iy - iy \tilde{\zeta}_T(f, y) + \frac{\lambda_{i,T}(f)}{1 + \zeta_T(f, y)}} - 1 \right| = \left| \frac{\tilde{\zeta}_T(f, y) - \frac{\lambda_{i,T}(f)}{iy + iy \zeta_T(f, y)}}{1 + \tilde{\zeta}_T(f, y) - \frac{\lambda_{i,T}(f)}{iy + iy \zeta_T(f, y)}} \right| < \varepsilon$$

when $y > C_\varepsilon = 2(\sigma^2 \mathbf{g} + \mathbf{a}^2) / \varepsilon$, thanks to the inequalities $|\zeta_T(f, y)| \leq \sigma^2 \mathbf{g} / y$ and $\Im(iy \tilde{\zeta}_T(f, y)) > 0$ as $y > 0$. This implies that the Stieltjes transform $\mathbf{p}_T(z) = N^{-1} \int_0^1 \text{Tr} \mathbf{S}_T(f, z) df$ of π_T satisfies $|iy \mathbf{p}_T(-iy) + 1| < \varepsilon$ when $y > C_\varepsilon$. Since the family $(m_T(z))_{T \in \mathbb{N}}$ is bounded on the compact subsets of \mathbb{C}_+ , from any sequence of integers, one can extract a subsequence $v(T)$ such that $\mathbf{p}_{v(T)}(z)$ converges to a function $\mathbf{p}^v(z)$ uniformly on the compacts of \mathbb{C}_+ . The function $\mathbf{p}^v(z)$ is moreover the Stieltjes transform of a positive measure μ^v . By a passage to the limit, it holds that $|iy \mathbf{p}^v(-iy) + 1| \leq \varepsilon$ when $y > C_\varepsilon$. This implies that $-iy \mathbf{p}^v(-iy) \rightarrow 1$ as $y \rightarrow \infty$, in other words, μ^v is a probability measure. We have shown that the sequence of probability measures μ_T is weakly compact. Therefore it is tight. \square

From the tightness of the sequence π_T and the convergence (4), we get that any converging subsequence of Stieltjes transforms of the μ_T has the Stieltjes transform of a probability measure as a limit. Hence the tightness of the μ_T and the second part of the statement of Theorem 2.3.

6.4 Sketch of the proof of Theorem 2.4

We remarked above that Equations (2)–(3) are similar to their counterparts in [17, Th. 2.4] (the latter are “discrete frequency” analogues of Equations (2)–(3)). For this reason, the proof of Theorem 2.4 follows closely the proof of [17, Th. 4.1]. We just provide here the main steps of this proof.

The first observation is that the mutual information per receive antenna $\int \log(1+\lambda)\mu_T(d\lambda)$ is related with \mathbf{m}_T by the equation

$$\int \log(1+\lambda)\mu_T(d\lambda) = \int_1^\infty \left(\frac{1}{t} - \mathbf{m}_T(-t)\right) dt$$

which is obtained from the relation $d[\log(1+\lambda/t)]/dt = -t^{-1} + (\lambda+t)^{-1}$. It also holds that $\int \lambda\mu_T(d\lambda) < 4(\mathbf{a}^2 + \sigma^2)$. Indeed, getting back to the proof of Theorem 2.1, consider an event where ν_T^n converges weakly to μ_T as $n \rightarrow \infty$, and where $\limsup_n \int \lambda\nu_T^n(d\lambda) \leq 2(\mathbf{a}^2 + \sigma^2)$. Let $C > 0$ be a continuity point of μ_T . Then for n large enough,

$$4(\mathbf{a}^2 + \sigma^2) > \int (\lambda \wedge C) \nu_T^n(d\lambda) \xrightarrow{n \rightarrow \infty} \int (\lambda \wedge C) \mu_T(d\lambda).$$

Taking a sequence of such continuity points that converges to infinity, we obtain the result by the monotone convergence theorem.

By mimicking the proof of [17, Lemma C.1], we can also show that $\int \lambda\pi_T(d\lambda) < \mathbf{a}^2 + \sigma^2$. This shows that

$$\begin{aligned} & |t^{-1} - \mathbf{p}_T(-t) - (t^{-1} - \mathbf{m}_T(-t))| \\ & \leq |t^{-1} - \mathbf{p}_T(-t)| + |(t^{-1} - \mathbf{m}_T(-t))| \\ & = \left| \int_0^\infty \left(\frac{1}{t} - \frac{1}{t+\lambda}\right) \pi_T(d\lambda) \right| + \left| \int_0^\infty \left(\frac{1}{t} - \frac{1}{t+\lambda}\right) \mu_T(d\lambda) \right| \\ & \leq t^{-2} \left| \int \lambda \pi_T(d\lambda) \right| + t^{-2} \left| \int \lambda \mu_T(d\lambda) \right| \leq 3(\mathbf{a}^2 + \sigma^2)/t^2. \end{aligned}$$

Therefore,

$$\mathcal{I}_T = \int \log(1+\lambda)\mu_T(d\lambda) = \int_1^\infty \left(\frac{1}{t} - \mathbf{p}_T(-t)\right) dt$$

is finite, and by the dominated convergence theorem, $N^{-1}I_T(S; (Y, H)) - \mathcal{I}_T \rightarrow 0$ as $T \rightarrow \infty$. In order to obtain the expression of \mathcal{I}_T provided in the statement, we need to find an antiderivative for $1/t - \mathbf{p}_T(-t)$. This antiderivative can be obtained by a lengthy but straightforward adaptation of the derivations of [17, § C.2].

7 Conclusion

In this paper, Shannon’s mutual information of a non centered time and frequency selective ergodic MIMO channel has been studied. In this context, the doubly infinite matrix bearing the channel coefficients represents an ergodic operator H on the Hilbert space $l^2(\mathbb{Z})$. The mutual information per receive antenna coincides with the integral of a log function with respect to the Integrated Density of States of the ergodic self adjoint operator HH^* . In the asymptotic regime where the numbers of antennas at both the transmitter and the receiver

converge to infinity at the same rate, the Stieltjes transform of the IDS can be approximated by another Stieltjes transform who can be derived by solving the system of functional equations given by Theorem 2.2. Based on these equations, an approximation of the mutual information is given by Theorem 2.4. With the help of these results, it is then possible to evaluate the impact of the channel statistical model on the mutual information. In the asymptotic regime, the approximation of the mutual information depends on the Doppler power spectrum and the spectral density matrix of the non centered part of the channel. The shape of the variance profile of the channel coefficients shows to have no effect on this approximation.

A Proof of Lemma 6.4

Given an integer $M \in \mathbb{N}$, let H^M and $Q^M(z)$ be as in the statement of Lemma 6.3. The variance of $Q_{p,q}^M(k, \ell)$ can be bounded by the PN inequality. Specifically, we have by this inequality $\text{Var} Q_{p,q}^M(k, \ell) \leq \chi_1 + \chi_2$ where

$$\begin{aligned}\chi_1 &= \sum_{i,j,u,v=-M}^M \sum_{n,t=0}^{N-1,T-1} \mathbb{E}[X_{n,t}(i,j)\bar{X}_{n,t}(u,v)] \mathbb{E}\left[\frac{\partial Q_{p,q}^M(k, \ell)}{\partial X_{n,t}(i,j)} \overline{\left(\frac{\partial Q_{p,q}^M(k, \ell)}{\partial X_{n,t}(u,v)}\right)}\right] \\ \chi_2 &= \sum_{i,j,u,v=-M}^M \sum_{n,t=0}^{N-1,T-1} \mathbb{E}[X_{n,t}(i,j)\bar{X}_{n,t}(u,v)] \mathbb{E}\left[\overline{\left(\frac{\partial Q_{p,q}^M(k, \ell)}{\partial \bar{X}_{n,t}(i,j)}\right)} \frac{\partial Q_{p,q}^M(k, \ell)}{\partial \bar{X}_{n,t}(u,v)}\right].\end{aligned}$$

Let us deal with χ_1 . Since

$$\frac{\partial Q_{p,q}^M(k, \ell)}{\partial X_{n,t}(i,j)} = -\mathbb{1}_{|i| \leq M} \mathbb{1}_{|j| \leq M} Q_{p,n}^M(k, i) [H^{M*} Q^M]_{t,q}(j, \ell),$$

we get

$$\begin{aligned}\chi_1 &= \frac{1}{T} \sum_{i,j,u,v} \sum_{n,t} \gamma_T(i-u) \phi_T(i-j)^2 \delta_{i-j, u-v} \\ &\quad \mathbb{E}\left[Q_{p,n}^M(k, i) [H^{M*} Q^M]_{t,q}(j, \ell) \overline{Q_{p,n}^M(k, u) [H^{M*} Q^M]_{t,q}(v, \ell)}\right] \\ &= \frac{1}{T} \sum_{r,s,u} \sum_{n,t} \gamma_T(r) \phi_T(s)^2 \\ &\quad \mathbb{E}\left[Q_{p,n}^M(k, r+u) [H^{M*} Q^M]_{t,q}(r+u-s, \ell) \overline{Q_{p,n}^M(k, u) [H^{M*} Q^M]_{t,q}(u-s, \ell)}\right].\end{aligned}$$

Using the inequality $|\mathbb{E} \sum_u X_u Y_u| \leq (\sum_u \mathbb{E}|X_u|^2)^{1/2} (\sum_u \mathbb{E}|Y_u|^2)^{1/2}$, we get

$$\begin{aligned}
\chi_1 &\leq \frac{1}{T} \sum_{r,s} |\gamma_T(r)| \phi_T(s)^2 \sum_u \sum_{n,t} \mathbb{E} |Q_{p,n}^M(k, u) [H^{M*} Q^M]_{t,q}(u-s, \ell)|^2 \\
&= \frac{1}{T} \sum_{r,s} |\gamma_T(r)| \phi_T(s)^2 \sum_u \mathbb{E} \left[[Q^M(k, u) Q^M(k, u)^*]_{p,p} [[H^{M*} Q^M](u-s, \ell)^* [H^{M*} Q^M](u-s, \ell)]_{q,q} \right] \\
&\leq \frac{1}{T} \frac{\sigma_T^2}{y^2} \sum_r |\gamma_T(r)| \sum_u \mathbb{E} \left[[[H^{M*} Q^M](u, \ell)^* [H^{M*} Q^M](u, \ell)]_{q,q} \right] \\
&\leq \frac{1}{T} \frac{\sigma^2 \mathbf{g}}{y^2} \mathbb{E} \left[[Q^{M*} H^M H^{M*} Q^M]_{q,q}(\ell, \ell) \right] \\
&\leq \frac{1}{T} \frac{\sigma^2 \mathbf{g}(|z| + y)}{y^4}.
\end{aligned}$$

The term χ_2 can be treated similarly, using the second identity in the statement of Lemma 6.3. This leads to the inequality

$$\mathbb{V}ar Q_{p,q}^M(k, \ell) \leq \frac{2}{T} \frac{\sigma^2 \mathbf{g}(|z| + y)}{y^4}.$$

Now, given any vector $a \in \mathcal{K}$, it is clear that $H^M H^{M*} a \rightarrow H H^* a$ strongly as $M \rightarrow \infty$. Since \mathcal{K} is a core for $H H^*$, the operator $H^M H^{M*}$ converges to $H H^*$ in the strong resolvent sense [34, Th. VIII.25]. It results that $Q_{p,q}^M(k, \ell) \rightarrow Q_{p,q}(k, \ell)$ almost surely as $M \rightarrow \infty$. Since both these random variables are bounded by $1/y$, we get that $\mathbb{V}ar Q_{p,q}^M(k, \ell) \rightarrow_{M \rightarrow \infty} \mathbb{V}ar Q_{p,q}(k, \ell)$, and the bound on $\mathbb{V}ar Q_{p,q}(k, \ell)$ is established.

We now establish the bound over $\mathbb{V}ar \text{Tr} Q(k, \ell)$. Using the PN inequality again, we get $\mathbb{V}ar \text{Tr} Q^M(k, \ell) \leq \chi_1 + \chi_2$ where

$$\chi_1 = \sum_{i,j,u,v=-M}^M \sum_{n,t=0}^{N-1, T-1} \mathbb{E} [X_{n,t}(i, j) \bar{X}_{n,t}(u, v)] \mathbb{E} \left[\frac{\partial \text{Tr} Q^M(k, \ell)}{\partial X_{n,t}(i, j)} \overline{\left(\frac{\partial \text{Tr} Q^M(k, \ell)}{\partial X_{n,t}(u, v)} \right)} \right]$$

and a similar formula for χ_2 . By the differentiation formula, we have

$$\frac{\partial \text{Tr} Q^M(k, \ell)}{\partial X_{n,t}(i, j)} = -\mathbb{1}_{|i| \leq M} \mathbb{1}_{|j| \leq M} \left[[H^{M*} Q^M](j, \ell) Q^M(k, i) \right]_{t,n}.$$

Therefore,

$$\begin{aligned}
\chi_1 &= \frac{1}{T} \sum_{i,j,u,v} \sum_{n,t} \gamma_T(i-u) \phi_T(i-j)^2 \delta_{i-j,u-v} \\
&\quad \mathbb{E} \left[\overline{[[H^{M^*}Q^M](j,\ell)Q^M(k,i)]_{t,n} [[H^{M^*}Q^M](v,\ell)Q^M(k,u)]_{t,n}} \right] \\
&= \frac{1}{T} \sum_{r,s,u} \sum_{n,t} \gamma_T(r) \phi_T(s)^2 \\
&\quad \mathbb{E} \left[\overline{[[H^{M^*}Q^M](u+r-s,\ell)Q^M(k,u+r)]_{t,n} [[H^{M^*}Q^M](u-s,\ell)Q^M(k,u)]_{t,n}} \right] \\
&= \frac{1}{T} \sum_{r,s,u} \sum_{n,t} \gamma_T(r) \phi_T(s)^2 \\
&\quad \mathbb{E} \left[\overline{[H^{M^*}Q^M D_\ell J_{k-\ell} Q^M]_{t,n}(u+r-s, u+r) \overline{[H^{M^*}Q^M D_\ell J_{k-\ell} Q^M]_{t,n}(u-s, u)}} \right]
\end{aligned}$$

where D_ℓ and J_ℓ are the operators on $l^2(\mathbb{Z})$

$$[J_\ell a](k) = a(k+\ell) \quad \text{and} \quad [D_\ell a](k) = \delta_{k,\ell} a(k)$$

where $a(k)$ (resp. $[J_\ell a](k)$, $[D_\ell a](k)$) is the k^{th} block with size N of the vector a (resp. $J_\ell a$, $D_\ell a$) in the canonical basis. Pursuing, we have

$$\chi_1 \leq \frac{1}{T} \sum_{r,s} |\gamma_T(r)| \phi_T(s)^2 \sum_u \sum_{n,t} \mathbb{E} \left| [H^{M^*}Q^M D_\ell J_{k-\ell} Q^M]_{t,n}(u-s, u) \right|^2.$$

Observe now that the operator $\mathcal{O} = H^{M^*}Q^M D_\ell J_{k-\ell} Q^M Q^{M^*} J_{k-\ell}^* D_\ell Q^{M^*} H^M$ is a finite rank operator such as

$$\text{rank}(\mathcal{O}) \leq N \quad \text{and} \quad \|\mathcal{O}\| \leq \frac{y+|z|}{y^4}.$$

Therefore,

$$\sum_u \sum_{n,t} \left| [H^{M^*}Q^M D_\ell J_{k-\ell} Q^M]_{t,n}(u-s, u) \right|^2 \leq \text{Tr } \mathcal{O} \leq N \frac{y+|z|}{y^4}$$

which shows that

$$\chi_1 \leq \frac{c\sigma^2 g(y+|z|)}{y^4}$$

and we get the same bound for χ_2 . It remains to use the same argument as in the first part of the proof to obtain the bound on $\text{Var } \text{Tr } Q(k, \ell)$.

The derivations leading to the bound on $\text{Var}[H^*Q]_{p,q}(k, \ell)$ are omitted, since they are similar to above.

References

- [1] N. I. Akhiezer and I. M. Glazman. *Theory of linear operators in Hilbert space*. Dover Publications, Inc., New York, 1993. Translated from the Russian and with a preface by Merlynd Nestell, Reprint of the 1961 and 1963 translations, Two volumes bound as one.

- [2] J.G. Andrews, S. Buzzi, W. Choi, S.V. Hanly, A. Lozano, A.C.K. Soong, and J.C. Zhang. What will 5G be? *IEEE J. Select. Areas Commun.*, 32(6):1065–1082, June 2014.
- [3] G. Auer, O. Blume, V. Giannini, I. Godor, M.A. Imran, Y. Jading, E. Katranaras, M. Olsson, D. Sabella, P. Skillermark, and W. Wajda. *D2.3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown*. INFSO-ICT-247733 EARTH, ver. 2.0, 2012.
- [4] E. Biglieri, J. Proakis, and S. Shamai. Fading channels: Information-theoretic and communications aspects. *IEEE Trans. Inform. Theory*, 44(6):2619, Oct. 1998.
- [5] E. Biglieri and G. Taricco. Transmission and reception with multiple antennas: Theoretical foundations. *Commun. Inf. Theory*, 1(2):183–332, October 2004.
- [6] A. Böttcher and B. Silbermann. *Analysis of Toeplitz operators*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, second edition, 2006. Prepared jointly with Alexei Karlovich.
- [7] G. Caire, G. Taricco, and E. Biglieri. Optimum power control over fading channels. *IEEE Trans. on Inform. Theory*, 45(5):1468–1489, Jul 1999.
- [8] G. Calcev, D. Chizhik, B. Goeransson, S. Howard, H. Huang, A. Kogiantis, A.F. Molisch, A.L. Moustakas, D. Reed, and H. Xu. A wideband spatial channel model for system-wide simulations. *IEEE Trans. Veh. Technol.*, 56(2):389, March 2007.
- [9] Y. Chen, S. Zhang, S. Xu, and G.Y. Li. Fundamental trade-offs on green wireless networks. *IEEE Commun. Mag.*, 49(6):30–37, June 2011.
- [10] C.N. Chuah, D.N.C. Tse, J.M. Kahn, and R.A Valenzuela. Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Trans. on Information Theory*, 48(3):637–650, Mar 2002.
- [11] R. Brent Dozier and J. W. Silverstein. On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices. *J. Multivariate Anal.*, 98(4):678–694, 2007.
- [12] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim. On the capacity achieving covariance matrix for Rician MIMO channels: An asymptotic approach. *IEEE Trans. on Information Theory*, 56(3):1048–1069, March 2010.
- [13] P. Almers *et.al.* Survey of channel and radio propagation models for wireless MIMO systems. *EURASIP J. on Wireless Comm. and Net.*, 2007.
- [14] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6:311–335, 1998.
- [15] R. M. Gray. *Entropy and information theory*. Springer, New York, second edition, 2011.
- [16] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur. A new approach for mutual information analysis of large dimensional multi-antenna channels. *IEEE Trans. on Information Theory*, 54(9):3987–4004, Sept 2008.

- [17] W. Hachem, Ph. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. *Ann. Appl. Probab.*, 17(3):875–930, 2007.
- [18] W. Hachem, Ph. Loubaton, J. Najim, and P. Vallet. On bilinear forms based on the resolvent of large random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(1):36–63, 2013.
- [19] W. C. Jakes. *Microwave Mobile Communications*. IEEE Press, 1994.
- [20] V. Kafedziski. Capacity of time varying single and multiple antenna channels with ISI using random ergodic operators. In *Proceedings of the 2002 IEEE International Symposium on Information Theory*, pages 479–, 2002.
- [21] V. Kafedziski. Rate distortion of stationary and nonstationary vector gaussian sources. In *IEEE workshop on Statistical Signal Processing*, pages 1054–1059. IEEE, 2005.
- [22] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, Berlin, second edition, 1976. Grundlehren der Mathematischen Wissenschaften, Band 132.
- [23] A.M. Khorunzhy and L.A. Pastur. Limits of infinite interaction radius, dimensionality and the number of components for random operators with off-diagonal randomness. *Comm. Math. Phys.*, 153(3):605–646, 1993.
- [24] M.G. Krein and A.A. Nudelman. *The Markov Moment Problem and Extremal Problems*. American Mathematical Society, Providence, Rhode Island, 1997.
- [25] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186 – 195, February 2014.
- [26] X. Mestre, J.R. Fonollosa, and A Pages-Zamora. Capacity of MIMO channels: asymptotic evaluation under correlated fading. *IEEE J. on Selected Areas in Communications*, 21(5):829–838, June 2003.
- [27] A.L. Moustakas and S.H. Simon. Random matrix theory of multi-antenna communications: the Ricean channel. *J. of Phys. A: Math. Gen.*, 38(49):10859, 2005.
- [28] A.L. Moustakas, S.H. Simon, and A.M. Sengupta. MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis. *IEEE Trans. on Information Theory*, 49(10):2545–2561, Oct 2003.
- [29] L. Pastur and A. Figotin. *Spectra of random and almost-periodic operators*, volume 297 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1992.
- [30] L. A. Pastur. On connections between the theory of random operators and the theory of random matrices. *Algebra i Analiz*, 23(1):169–199, 2011.
- [31] L. A. Pastur and M. Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.

- [32] K.I. Pedersen, P.E. Mogensen, and B.H. Fleury. A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments. *IEEE Trans. Veh. Technol.*, 49(2):437, March 2000.
- [33] Qualcomm. The 1000x data challenge. Technical report.
- [34] M. Reed and B. Simon. *Methods of modern mathematical physics. I*. Academic Press Inc., New York, second edition, 1980. Functional analysis.
- [35] F. Rusek, D. Persson, B.K. Lau, E.G. Larsson, T.L. Marzetta, O. Edfors, and F. Tufvesson. Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, 30(1):40 – 60, Jan 2013.
- [36] G. Taricco. Asymptotic mutual information statistics of separately correlated Rician fading MIMO channels. *IEEE Trans. on Information Theory*, 54(8):3490–3504, Aug 2008.
- [37] I.E. Telatar. Capacity of Multi-Antenna Gaussian Channels. Published in *European Transactions on Telecommunications*, 10(6):585–595, Nov/Dec 1999. Technical Memorandum, Bell Laboratories, Lucent Technologies, October 1995.
- [38] A.M. Tulino, A Lozano, and S. Verdu. Impact of antenna correlation on the capacity of multiantenna channels. *IEEE Trans. on Inform. Theory*, 51(7):2491–2509, July 2005.