Time and frequency selective Ricean MIMO capacity: an ergodic operator approach

Walid Hachem, Aris L. Moustakas and Leonid A. Pastur

***Abstract*—From the standpoint of Information theory, a time and frequency selective Ricean ergodic MIMO channel can be represented in the Hilbert space $l^2(\mathbb{Z})$ by a random ergodic self-adjoint operator whose Integrated Density of States (IDS) governs the behavior of the Shannon's mutual information. 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. 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.**

I. Introduction and problem statement

In the landmark papers by Foschini and Gans [1] and by Telatar [2] the great promise of the use of multiple transmit and receive antennas (MIMO) was presented and established. The importance of such MIMO links is based on the fact that parallel data streams emanating from different transmit antennas can be decoded simultaneously from the receive array, making the throughput scale linearly with the number of transmit antennas.

In [2] it was assumed for simplicity that either the channel is completely static or that it is independently drawn at each time step. In the latter case it was shown that the capacity of the system is the ergodic average of the mutual information with respect to the channel realizations. In typical situations however, the channel does vary continuously, albeit perhaps slowly and hence the above result does not obviously apply. To have a more realistic channel description [3], [4] discuss the situation of a so-called block-fading channel. In this case the channel matrix is held constant over a block of length b. In the limit that both b and the number of blocks B within a packet tend to infinity, the previous ergodic result is obtained once again. However, this is also not necessarily a fully realistic model, since the channel tends to vary continuously with significant temporal correlations, as well as frequency correlations in the case of a tap-delay channel.

In this paper we address the calculations of Shannon's mutual information for time- and frequency - correlated MIMO Ricean channels in the limit of large antenna arrays. To do so, we need to deal with matrix sizes that are proportional to the temporal length of the packet, which need to diverge much faster than the size of the antenna array. Hence, one major contribution of our work is the introduction of operator theory, which are effectively infinitely large matrices.

Denoting by T and N the respective numbers of antennas at the transmitter and at the receiver, the $\mathbb{C}^N$-valued signal received at time k is

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

where $\{S(k)\}_{k\in\mathbb{Z}}$ is an independent process with law $\mathcal{CN}(0, I_T)$ representing the input, $\{V(k)\}_{k\in\mathbb{Z}}$ is an independent process with law $\mathcal{CN}(0, I_N)$ representing the noise, and the $\mathbb{C}^{N\times(2L+1)T}$-valued random process $\{\boldsymbol{H}(k) = [H(k,k-L),\ldots,H(k,k+L)]\}_{k\in\mathbb{Z}}$, assumed to be circular Gaussian, stationary, ergodic, and generally non-centered, represents the multipath MIMO channel. It is assumed that $\{S(k)\}$, $\{V(k)\}$, and $\{\boldsymbol{H}(k)\}$ are mutually independent.

To be more specific on the channel statistics, we shall assume that $H(k,\ell) = A(k-\ell) + X(k,\ell)$ where the matrix $\boldsymbol{A} = [A(-L),\ldots,A(L)]$ is deterministic, and where $X(k,\ell) = T^{-1/2}\phi(k-\ell)W(k,\ell)$ with $\phi : \mathbb{Z} \to [0,\infty)$ being a real valued function supported by the set $\{-L,\ldots,L\}$, and $\{W(k,\ell) = [W_{n,t}(k,\ell), n = 0 : N-1, t = 0 : T-1]\}_{k,\ell\in\mathbb{Z}}$ is a complex circular Gaussian centered random field such that

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

where the covariance function γ is summable and satisfies $\gamma(0) = 1$ without generality loss.

The operator $\boldsymbol{A}$ models the deterministic (Ricean) multipath part of the channel. The function ϕ models the random multipath amplitude profile, while the covariance function γ whose "effective support" decreases with the mobile speed is related with the Doppler effect. Since γ is summable, the stationary process $\{\boldsymbol{H}(k)\}$ is ergodic.

Our purpose is to study the mutual information $I(S;(Y,H))$ between $\{S(k)\}$ and the couple $(\{Y(k)\},\{\boldsymbol{H}(k)\})$, *i.e.*, the one assuming the channel to be perfectly known at the receiver. Writing $H(k,\ell) = 0$ when $|k-\ell| > L$, and defining the $(2n+1)N \times (2n+2L+1)T$ banded matrix $H^n = [H(k,\ell)]_{(k,\ell)=(-n,-n-L)}^{(n,n+L)}$ for large n, it is known [5] that

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

where

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

Walid Hachem is with LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France. Aris Moustakas is with the Department of Physics, National & Capodistrian University of Athens, Athens, Greece. Leonid Pastur is with the Institute of Low Temperature Physics, Kharkiv, Ukraine. The proofs for the theorems in this paper have appeared recently in [16].

One natural way to tackle this problem is to identify the MIMO channel with the random unbounded operator [6] represented by the doubly infinite banded matrix

$$H = \begin{bmatrix} \ddots & \ddots & & 0 \\ H(-1,-1) & H(-1,0) & & \\ H(0,-1) & H(0,0) & H(0,1) & \\ & H(1,0) & H(1,1) & \\ 0 & & \ddots & \ddots \end{bmatrix} \tag{1}$$

and acting on the Hilbert space $l^2(\mathbb{Z})$. Denoting by H^* the adjoint of H, it is easy to show that HH^* is self-adjoint. Moreover, the ergodicity of $\{\boldsymbol{H}(k)\}$ implies that HH^* is itself ergodic in the sense of [6, p. 33] (see Section IV below for more details). In addition, HH^* has a so-called *Integrated Density of States* (IDS). Namely, noticing that $H^n H^{n*}$ is a finite submatrix of HH^*, there exists a deterministic probability measure μ such that

$$\frac{1}{(2n+1)N} \operatorname{Tr} g(H^n H^{n*}) \xrightarrow[n\to\infty]{\text{a.s.}} \int g(\lambda)\, \mu(d\lambda) \tag{2}$$

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

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

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 (ST) of the measure μ, that is the complex analytical function $\boldsymbol{m}_\mu(z) = \int (\lambda - z)^{-1} \mu(d\lambda)$ defined on the upper half-plane $\mathbb{C}_+ = \{z : \Im z > 0\}$. By making profit of the intimate connection between $\boldsymbol{m}_\mu$ and the resolvent $Q(z) = (HH^* - z)^{-1}$ of the operator HH^*, we will produce a probability measure $\boldsymbol{\pi}$ whose ST approximates $\boldsymbol{m}_\mu$. Ultimately, this will lead us to a large T approximation of $I(S; (Y, H))$.

Ergodic operator theory [6] is widely used in the fields of quantum physics and statistical mechanics. It has tight connections with random matrix theory who were explored in [7]. While random matrix theory is a well known tool in the wireless communications literature [8], ergodic operators are much less used in this domain. We however cite [9] who uncovered the relation between the mutual information and the IDS of an operator without much elaborating on the properties of the latter.

II. Main results

We start by providing our working assumptions, indexing the mathematical objects at hand with T to make our convergence results clearer:

1) $0 < \liminf_{T\to\infty} N(T)/T \le \sup_T N(T)/T < \infty$.
2) $0 < \liminf_T \sigma_T^2 \le \sup_T \sigma_T^2 < \infty$ where $\sigma_T^2 = \sum_\ell \phi_T(\ell)^2$.
3) $\sup_T \sum_\ell |\gamma_T(\ell)| < \infty$.
4) $\sup_T \sum_{\ell=-L(T)}^{L(T)} \|A_T(\ell)\| < \infty$, where $\|\cdot\|$ is the spectral norm.

Let us comment these assumptions. The parameter σ_T^2 in Assumption 2 is the part of the received power due to the random part of the channel. The practical interpretation of Assumption 3 is that the so-called coherence time of the channel [10] does not grow with T. Relaxing Assumption 3 would require completely different mathematical tools than those used in this paper. With some extra effort, it is possible to make Assumption 4 less stringent by replacing it with a bound on the Euclidean norms of the rows and columns of the $A(\ell)$. Finally, the Gaussian assumption on the channel elements can be relaxed by using the interpolation tools developed in [11].

We now state our main results.

Theorem 1. *For all $T > 0$, the positive self-adjoint operator HH^* is ergodic and has an IDS defining a probability measure μ_T. The sequence $\{I(S^n; (Y^n, H^n))/(2n+1)\}$ converges as $n \to \infty$, and the limit is $I_T(S; (Y, H)) = N \int \log(1+\lambda)\, \mu_T(d\lambda) < \infty$.*

Let

$$\boldsymbol{\gamma}_T(f) = \sum_k \exp(2\imath\pi k f) \gamma_T(k)$$

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

$$\boldsymbol{A}_T(f) = \sum_k \exp(2\imath\pi k f) A_T(k)$$

be the $N \times T$ Fourier transform of the sequence $\{A_T(k)\}$. Write $(\boldsymbol{A}_T \boldsymbol{A}_T^*)(f) = \boldsymbol{A}_T(f)\boldsymbol{A}_T^*(f)$ and $(\boldsymbol{A}_T^* \boldsymbol{A}_T)(f) = \boldsymbol{A}_T^*(f)\boldsymbol{A}_T(f)$ for compactness. Theorem 2 defines a sequence $\{\boldsymbol{\pi}_T\}$ of probability measures that approximate the μ_T:

Theorem 2. *Let $\boldsymbol{S}_T(f,z)$ and $\widetilde{\boldsymbol{S}}_T(f,z)$ be respectively the $N \times N$ and $T \times T$ matrices*

$$\begin{aligned} \boldsymbol{S}_T(f,z) &= \Big[-z\big(1+\sigma_T^2 \boldsymbol{\gamma}_T(f) \star \tilde{\boldsymbol{\varphi}}_T(f,z)\big) I_N \\ &\quad + \big(1+\sigma_T^2 \boldsymbol{\gamma}_T(-f) \star \boldsymbol{\varphi}_T(f,z)\big)^{-1} (\boldsymbol{A}_T\boldsymbol{A}_T^*)(f)\Big]^{-1}, \\ \widetilde{\boldsymbol{S}}_T(f,z) &= \Big[-z\big(1+\sigma_T^2 \boldsymbol{\gamma}_T(-f) \star \boldsymbol{\varphi}_T(f,z)\big) I_T \\ &\quad + \big(1+\sigma_T^2 \boldsymbol{\gamma}_T(f) \star \tilde{\boldsymbol{\varphi}}_T(f,z)\big)^{-1} (\boldsymbol{A}_T^*\boldsymbol{A}_T)(f)\Big]^{-1} \end{aligned} \tag{3}$$

with

$$\boldsymbol{\gamma}_T(f) \star \tilde{\boldsymbol{\varphi}}_T(f,z) = \int_0^1 \boldsymbol{\gamma}_T(f-u) \tilde{\boldsymbol{\varphi}}_T(u,z)\, du,$$

$$\boldsymbol{\gamma}_T(-f) \star \boldsymbol{\varphi}_T(f,z) = \int_0^1 \boldsymbol{\gamma}_T(u-f) \boldsymbol{\varphi}_T(u,z)\, du.$$

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

$$\boldsymbol{\varphi}_T(f,z) = \frac{\operatorname{Tr} \boldsymbol{S}_T(f,z)}{T}, \ \tilde{\boldsymbol{\varphi}}_T(f,z) = \frac{\operatorname{Tr} \widetilde{\boldsymbol{S}}_T(f,z)}{T}$$

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

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

is the ST of a probability measure $\boldsymbol{\pi}_T$ carried by $[0,\infty)$.

The sequences $\{\mu_T\}$ and $\{\boldsymbol{\pi}_T\}$ are tight, and

$$\int g(\lambda)\, \mu_T(d\lambda) \ - \ \int g(\lambda)\, \boldsymbol{\pi}_T(d\lambda) \xrightarrow[T\to\infty]{} 0$$

for any continuous and bounded real function g.

The large T approximation of the mutual information is provided by the following theorem:

Theorem 3. *It holds that $N^{-1} I_T(S;(Y,H)) - \mathcal{I}_T \xrightarrow[T\to\infty]{} 0$ where*

$$\mathcal{I}_T = \int \log(1+\lambda)\, \boldsymbol{\pi}_T(d\lambda),$$

and the integral is given by

$$\begin{aligned}
\mathcal{I}_T = & \\
& \frac{1}{N} \int_0^1 \log\det\Big((1 + \sigma_T^2 \boldsymbol{\gamma}_T(f) \star \tilde{\boldsymbol{\varphi}}_T(f,-1)) I_N \\
& \qquad + \frac{(\boldsymbol{A}_T \boldsymbol{A}_T^*)(f)}{1 + \sigma_T^2 \boldsymbol{\gamma}_T(-f) \star \boldsymbol{\varphi}_T(f,-1)}\Big)\, df \\
& + \frac{T}{N} \int_0^1 \log(1 + \sigma_T^2 \boldsymbol{\gamma}_T(-f) \star \boldsymbol{\varphi}_T(f,-1))\, df \\
& - \frac{T}{N} \int_{[0,1]^2} \sigma_T^2 \boldsymbol{\gamma}_T(f-v) \tilde{\boldsymbol{\varphi}}_T(v,-1) \boldsymbol{\varphi}_T(f,-1)\, dv\, df.
\end{aligned}$$

These results call for some remarks. We first observe that the variance profile represented by the function ϕ_T^2 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 can also show that if the channel is centered, then $\boldsymbol{\pi}_T$ is the Marchenko-Pastur distribution who coincides with the deterministic equivalent of the spectral measure of ZZ^* where Z is a $N \times T$ matrix with independent centered elements having the common variance σ_T^2/T. Finally, for small coherence times, *i.e.*, when $\boldsymbol{\gamma}_T(f)$ becomes close to the constant function equal to one, the approximation $\mathcal{I}_T$ becomes close to the one provided by the so called Information plus Noise model of [12].

III. 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 [13], which, in the time domain has the following correlation form:

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

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 $\boldsymbol{A}_T(f)$. 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

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

where $\theta_k = k\pi/L_{tot}$, for $k = -L, \ldots, L$ and $L_{tot} = 2L+1$. The exponential dependence on the delay spread ξ has been seen experimentally [14], [15]. Then $\boldsymbol{A}_T(f)$ follows from

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

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

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

in various cases. In Fig. 1 we plot the mutual information for different normalized values of f_d. We see that the dependence of the mutual information on 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 $(\boldsymbol{A}_T \boldsymbol{A}_T^*)(f)$ 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.

In Fig. 2 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.

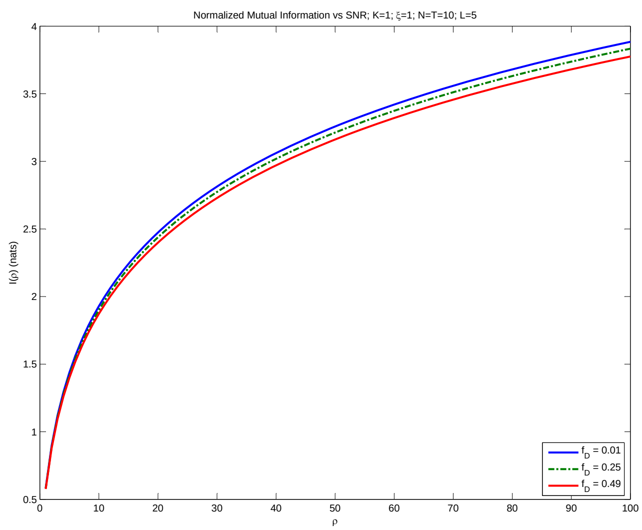

Figure 1. Mutual Information versus ρ for various f_D

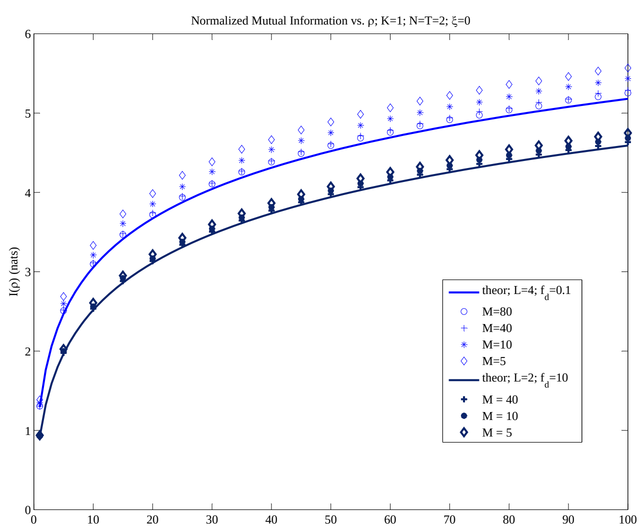

Figure 2. Mutual Information versus ρ for various L and block sizes

IV. MAIN PROOF IDEAS

We now sketch the main ideas of the proofs. For a comprehensive treatment, the interested reader is referred to the long version [16] of this paper. We first focus on the channel representation as an ergodic operator. Assume for the moment that $N = T = 1$. Let us redenote the process $\{\boldsymbol{H}(k)\}$ and the operator H as $\{\boldsymbol{H}(\omega, k)\}$ and $H(\omega)$ respectively where ω is an elementary event on the probability space Ω. We know from our channel model that the shift $B : \Omega \to \Omega$ characterized by the equation $\boldsymbol{H}(B\omega, k) = \boldsymbol{H}(\omega, k+1)$ is ergodic. On the other hand, the observation of (1) clearly shows that $H(B\omega) = UH(\omega)U^{-1}$ where U is the shift operator $Ua = \sum_k \alpha_{k+1} e_k$ for $a = \sum_k \alpha_k e_k$, and e_k is the k^{th} canonical basis vector of $l^2(\mathbb{Z})$. Since U is unitary, the operator H is then also ergodic in the sense of [6, p. 33]. Moreover, the operator HH^* who is self-adjoint is also ergodic, since $[HH^*](B\omega) = U[HH^*](\omega)U^{-1}$. A key consequence of the ergodicity of HH^* is that the expectation of the (k, ℓ) element of the resolvent $Q(z) = (HH^* - z)^{-1}, z \in \mathbb{C}_+$ depends on $k - \ell$ only. We denote $\mathbb{E}Q(z, k-\ell)$ such an element.

Another key fact related with the banded nature of H is that HH^* admits an IDS (see (2)) whose ST will then coincide with $\mathbb{E}Q(z, 0)$ (see *e.g.* [6, Th. II.4.8]).

All these properties of HH^* can be shown to still hold true when N or T is > 1. In particular, representing $Q(z)$ as an infinite matrix with $N \times N$ blocks at the same positions as those who obviously appear in the construction of HH^*, the expectation of the (k, ℓ) block will depend on $k - \ell$ only and will be denoted as $\mathbb{E}Q_T(z, k-\ell)$. Furthermore, the IDS exists, but its ST $\boldsymbol{m}_{\mu_T}(z)$ will now coincide with $N^{-1} \operatorname{Tr} \mathbb{E}Q_T(z, 0)$.

Theorem 1 follows from these results, the fact that $\log(1+\lambda)$ is unbounded unlike the function g in (2) being easily manageable.

The purpose of Theorems 2 and 3 is to render the integral in the statement of Theorem 1 more informative by resorting to the asymptotic regime $T \to \infty$. The approach follows the lines of [17]. Specifically, we look for a sequence $\boldsymbol{m}_{\boldsymbol{\pi}_T}(z)$ of ST of probability measures $\boldsymbol{\pi}_T$ such that

$$m_{\boldsymbol{\pi}_T}(z) - \frac{\operatorname{Tr} \mathbb{E}Q_T(z,0)}{N} \xrightarrow[T\to\infty]{} 0, \quad z \in \mathbb{C}_+.$$

To that end, we rely on two basic tools that are frequently used in the close field of random matrix theory: an integration by parts formula for evaluating the expectations of functionals of random Gaussian vectors (already used in [17]), and the Poincaré-Nash inequality for the variance controls (see [11] for further details about these tools and their applications in random matrix theory).

By adapting these tools to the infinite dimensional context, we get after some calculations

$$
\begin{aligned}
&\mathbb{E}Q(z,k) \\
&= -z^{-1}I(k) - \sigma^2 \sum_r \gamma(r) \Big[\frac{\operatorname{Tr} \mathbb{E}\widetilde{Q}(z,r)}{T} \Big] \mathbb{E}Q(z,k-r) \\
&\quad + z^{-1}\mathbb{E}[AH^*Q](z,k) + E_k(z), \\
&\mathbb{E}[AH^*Q](z,k) \\
&= -\sigma^2 \sum_r \gamma(-r) \Big[\frac{\operatorname{Tr} \mathbb{E}Q(z,r)}{T} \Big] \mathbb{E}[AH^*Q](z,k-r) \\
&\quad + \mathbb{E}[AA^*Q](z,k) + E'_k(z) \quad \text{for all } k \in \mathbb{Z},
\end{aligned} \tag{4}
$$

where $\widetilde{Q}(z) = (H^*H - z)^{-1}$, and $\mathbb{E}\widetilde{Q}(z,r)$, $\mathbb{E}[AH^*Q](z,r)$ and $\mathbb{E}[AA^*Q](z,r)$ are the respective expectations of the $(k, k-r)$ block of $\widetilde{Q}(z)$, $AH^*Q(z)$ and $AA^*Q(z)$ for all $k \in \mathbb{Z}$. The terms $E_k(z)$ and $E'_k(z)$ are perturbation terms who become negligible for large T. We also have similar equations for the blocks of $\widetilde{Q}(z)$.

Now, identifying the function $\boldsymbol{S}_T(\cdot, z)$ defined in (3) with a multiplication operator on the Hilbert space $\mathcal{L}^2([0,1] \to \mathbb{C}^N)$, and letting $\mathcal{F}_T$ be the operator who sends $\boldsymbol{g} \in \mathcal{L}^2([0,1] \to \mathbb{C}^N)$, to the sequence of its Fourier coefficients in $l^2(\mathbb{Z})$, the operator

$$S_T(z) = \mathcal{F}_T \boldsymbol{S}_T(\cdot, z) \mathcal{F}_T^*$$

is a bounded block-convolution operator. It turns out that the blocks $S_T(z,r)$ who lie on the block-diagonal r of $S_T(z)$ satisfy an infinite system of equations similar to (4), but without the perturbations. The same remark holds for the diagonal blocks of $\widetilde{S}_T(z)$. From these considerations, we can show that $N^{-1} \operatorname{Tr} S_T(z,0) = \int_0^1 N^{-1} \operatorname{Tr} \boldsymbol{S}_T(f,z)\, df$ is on the one hand the ST of a probability measure $\boldsymbol{\pi}_T$, and on the other hand, approximates the ST of μ_T in the sense that $N^{-1} \operatorname{Tr} S_T(z,0) - N^{-1} \operatorname{Tr} \mathbb{E}Q_T(z,0) \to_T 0$. Theorem 2 follows.

To prove Theorem 3, it is well known that we need to find an expression for an antiderivative of the ST $\boldsymbol{m}_{\boldsymbol{\pi}_T}$ in order to obtain an expression for $\mathcal{I}_T$ (see *e.g.* [8]). Fortunately, the system (3) turns out to be formally close to the system described in [18, Th. 2.4]. The expression of $\mathcal{I}_T$ is obtained by similar means as in that paper.

V. Conclusion

We close this article with two remarks. First, our problem could have been solved with large random matrix techniques instead of random operator ones. In the former case, one would need to work on the rather complicated matrix model $H^n H^{n*}$, making $n \to \infty$ faster than T, N, and dealing with the border effects. Working with random operators is much simpler and more natural, since this approach naturally fits the ergodic nature of the channel. Second, other asymptotic regimes can be dealt with using the techniques of this paper (see also [17]). For instance, we could have fixed N and T and made L tend to infinity.

References

[1] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[2] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Transactions on Telecommunications and Related Technologies*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[3] G. Caire, G. Taricco, and E. Biglieri, “Optimum power control over fading channels,” *IEEE Trans. on Inform. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul 1999.

[4] L. H. Ozarow, S. Shamai, and A. D. Wyner, “Information theoretic considerations for cellular mobile radio,” *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[5] R. M. Gray, *Entropy and information theory*, 2nd ed. Springer, New York, 2011.

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

[7] L. A. Pastur, “On connections between the theory of random operators and the theory of random matrices,” *Algebra i Analiz*, vol. 23, no. 1, pp. 169–199, 2011.

[8] A. Tulino and S. Verdú, “Random matrix theory and wireless communications,” in *Foundations and Trends in Communications and Information Theory*. Now Publishers, Jun. 2004, vol. 1, pp. 1–182.

[9] V. Kafedziski, “Capacity of time varying single and multiple antenna channels with ISI using random ergodic operators,” in *Proc. IEEE ISIT*, 2002, pp. 479–.

[10] E. Biglieri, J. Proakis, and S. Shamai, “Fading channels: Information-theoretic and communications aspects,” *IEEE Trans. Inform. Theory*, vol. 44, no. 6, p. 2619, Oct. 1998.

[11] L. A. Pastur and M. Shcherbina, *Eigenvalue distribution of large random matrices*, ser. Mathematical Surveys and Monographs. Providence, RI: American Mathematical Society, 2011, vol. 171.

[12] R. B. Dozier and J. W. Silverstein, “On the empirical distribution of eigenvalues of large dimensional information-plus-noise-type matrices,” *J. Multivariate Anal.*, vol. 98, no. 4, pp. 678–694, 2007.

[13] W. C. Jakes, *Microwave Mobile Communications*. IEEE Press, 1994.

[14] K. Pedersen, P. Mogensen, and B. Fleury, “A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments,” *IEEE Trans. Veh. Technol.*, vol. 49, no. 2, p. 437, Mar. 2000.

[15] G. Calcev, D. Chizhik, B. Goeransson, S. Howard, H. Huang, A. Kogiantis, A. Molisch, A. Moustakas, D. Reed, and H. Xu, “A wideband spatial channel model for system-wide simulations,” *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, p. 389, Mar. 2007.

[16] W. Hachem, A. Moustakas, and L. Pastur, “The Shannon’s mutual information of a multiple antenna time and frequency dependent channel: An ergodic operator approach,” *J. Math. Phys.*, vol. 56, no. 11, pp. 113 501, 29, 2015.

[17] 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.*, vol. 153, no. 3, pp. 605–646, 1993.

[18] W. Hachem, P. Loubaton, and J. Najim, “Deterministic equivalents for certain functionals of large random matrices,” *Ann. Appl. Probab.*, vol. 17, no. 3, pp. 875–930, 2007.