Asymptotic Capacity
of Multi-User MIMO Communications

Romain Couillet
ST-Ericsson, Supélec
505 Route des Lucioles
06560 Sophia Antipolis, France
Email: romain.couillet@supelec.fr

Mérouane Debbah
Alcatel-Lucent Chair, Supélec
Plateau de Moulon, 3 rue Joliot-Curie
91192 Gif sur Yvette, France
Email: merouane.debbah@supelec.fr

Jack W. Silverstein
Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27695-8205
Email: jack@math.ncsu.edu

Abstract—**This paper introduces two new formulas to derive explicit capacity expressions of a class of communication schemes, which include single-cell multi-user MIMO and single-user MIMO with multi-cell interference. The extension of a classical theorem from Silverstein allows us to assume a channel Kronecker model between the base stations and the cellular terminals, provided that they all embed a large number of antennas. As an introductory example, we study the single-user MIMO setting with multi-cell interference, in the downlink. We provide new asymptotic capacity formulas when single-user decoding of the incoming data or MMSE decoding are used. Simulations are shown to corroborate the theoretical claims, even when the number of transmit/receive antennas is not very large.**

I. INTRODUCTION

In the last decade, while mobile networks were expected to run out of power and frequency resources, Foschini [3] and Telatar [4] introduced the notion of MIMO (multiple input multiple output) systems and predicted a growth of capacity performance of $\min(n_R, n_T)$ times the single antenna capacity for an n_T-antenna transmitter and an n_R-antenna receiver. However, this tremendous multiplexing gain can only be achieved for large SINR (signal-to-interference plus noise ratio) and without signal correlation at the channel ends. In case of correlation due to antenna spacing or to poorly scattering environments, the theoretically achievable gains are still not completely available. Moreover, in present multi-cell wireless mobile networks, neither base stations nor users cooperate; this leaves the device manufacturers with the dilemma of increasing the signal processing capabilities of the transmit/receive units to result into non significant throughput gains when adjacent cells interfere one another. Besides, due to limited computational constraints, suboptimal linear techniques such as MMSE (minimum mean square error) decoding are used at the receiver [5], in place of optimal single-user decoders.

In this work, we derive the per-antenna channel capacity of MMSE receivers against optimal single-user decoders in multi-cell networks, when the number of antennas at the transmitters and receivers is large. The capacity here is defined as the supremum of the achievable rates between a base station and a specific user (in uplink or in downlink) interfered by other cells. We model all transmission channels by the well-spread Kronecker model [6]. Few major contributions propose to study the capacity performance of point-to-point communications with interference. In [9], the authors carry out the performance analysis of TDMA-based networks with inter-cell interference. In [10], a random matrix approach is used to study large CDMA-based networks with inter-cell interference. In the MIMO context, [7] provides an analytic solution to our problem in two-cell networks, using replica methods [8]. These methods are however tedious since they require heavy combinatorial calculus. We propose in the following a more direct approach, based on analytical tools of random matrix theory. Note also that [13] provides a deterministic capacity expression in our context, when the receive covariance matrices $\mathbf{R}_k$ are assumed to share the same eigenspace[1].

Although this specific work is dedicated to the study of point-to-point MIMO systems with multi-cell interference, the potential applications to the mathematical results we introduce cover a larger class of problems, in which channel capacities express as the log determinant of a sum of Gram matrices $\mathbf{X}_i\mathbf{X}_i^{\mathsf{H}}$, where $\mathbf{X}_i$ is a large matrix modelled as Kronecker. For example, beside the uplink/downlink multi-cell single-user MIMO, our results are applicable to single-cell multi-user MIMO communications in the uplink, evaluation of the capacity region of multiple access channels and broadcast channels [2].

The remainder of this work is structured as follows: in Section II, we provide mathematical preliminaries and we introduce a new important theorem, for which we provide a sketch of proof. In Section III, we introduce the system model. In Section IV, the point-to-point capacity of the channel between a base station and a user, interfered by other cells, is derived when single-user decoding or MMSE decoding are performed at the receiver. In Section V, we provide simulation results of the previously derived theoretical formulas. Finally, in Section VI, we give our conclusions.

Notation: In the following, boldface lower-case symbols represent vectors, capital boldface characters denote matrices ($\mathbf{I}_N$ is the $N \times N$ identity matrix). X_{ij} denotes the (i,j) entry of $\mathbf{X}$. The Hermitian transpose is denoted $(\cdot)^{\mathsf{H}}$. The

---

[1]this assumption is too strong for our current study since it boils down to supposing that all entities have the same geometrical antenna pattern and that transmit/receive energy comes from the same solid angle for all users.

operators $\operatorname{tr}\mathbf{X}$, $|\mathbf{X}|$ and $\|\mathbf{X}\|$ represent the trace, determinant and spectral norm of matrix $\mathbf{X}$, respectively. The symbol $x_n \xrightarrow{\text{a.s.}} x$ denotes almost sure convergence of x_n to x. The notation $F^{\mathbf{Y}}$ stands for the empirical distribution of the eigenvalues of the Hermitian matrix $\mathbf{Y}$.

II. MATHEMATICAL PRELIMINARIES

Part of this work is dedicated to the introduction of a novel theorem, from which the multi-cell downlink and uplink capacities will be given compact expressions. This theorem generalizes Silverstein and Bai's formula [1] to multiple random matrices with separable variance profiles, i.e. following the Kronecker model, and unfolds as follows,

Theorem 1: Let K, $N \in \mathbb{N}$ be some positive integers. Let

$$\mathbf{B}_N = \sum_{k=1}^{K} \mathbf{R}_k^{\frac{1}{2}} \mathbf{X}_k \mathbf{T}_k \mathbf{X}_k^{\mathsf{H}} \mathbf{R}_k^{\frac{1}{2}} \tag{1}$$

be an $N \times N$ matrix with the following hypothesis for all $k \in \{1, \ldots, K\}$,

1) $\mathbf{T}_k$ is $n_k \times n_k$ Hermitian nonnegative definite, $n_k \in \mathbb{N}^*$, qith unit normalized trace,
2) $\mathbf{R}_k^{\frac{1}{2}}$ is the $N \times N$ Hermitian nonnegative definite square root of the nonnegative definite matrix $\mathbf{R}_k$ with unit normalized trace,
3) $\mathbf{X}_k$ is $N \times n_k$ with i.i.d. complex Gaussian entries with variance $1/n_k$.

For $k \in \{1, \ldots, K\}$, let $c_k = n_k/N$. Also denote, for $z \in \mathbb{C} \setminus \mathbb{R}^+$, $m_N(z) = \frac{1}{N}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}$. Then, as all n_k and N grow large (while K is fixed), with ratio c_k

$$m_N(z) - m_N^{(0)}(z) \xrightarrow{\text{a.s.}} 0 \tag{2}$$

where

$$m_N^{(0)}(z) = \frac{1}{N} \operatorname{tr} \left(\sum_{k=1}^{K} \int \frac{\tau_k dF^{\mathbf{T}_k}(\tau_k)}{1 + \frac{\tau_k}{c_k} e_k(z)} \mathbf{R}_k - z\mathbf{I}_N \right)^{-1} \tag{3}$$

and the set of functions $\{e_i(z)\}$, $i \in \{1, \ldots, K\}$, form the unique solution to the K equations

$$e_i(z) = \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\sum_{k=1}^{K} \int \frac{\tau_k dF^{\mathbf{T}_k}(\tau_k)}{1 + \frac{\tau_k}{c_k} e_k(z)} \mathbf{R}_k - z\mathbf{I}_N \right)^{-1} \tag{4}$$

such that $\operatorname{sgn}(\Im[e_i(z)]) = \operatorname{sgn}(\Im[z])$.

The function $m_N(z)$ is the Stieltjes transform of the random variable with cumulative distribution function $F^{\mathbf{B}_N}$. The complete proof of a more general expression of this theorem is given in an extended version of the present article [2].

Remark 1: This theorem allows us to derive Stieltjes transforms of large matrices *independently* of the realization of the $\mathbf{X}_k$ matrices. In wireless communications, this allows one to characterize the performances of a multi-user or multi-cell communication based only on the transmit and receive correlations $\mathbf{R}_k$ and $\mathbf{T}_k$. This further helps to estimate channel capacity thanks to the Shannon transform,

Theorem 2: Let $\mathbf{B}_N$ be a random Hermitian matrix as defined in Theorem 1 with the additional assumption that, for

all N, n_k, $\operatorname{tr}(\mathbf{T}_k) = n_k$, $\operatorname{tr}(\mathbf{R}_k) = N$, and let $x > 0$. Then, for large N, n_k, $\mathcal{V}(x) - \mathcal{V}^{(0)}(x) \xrightarrow{\text{a.s.}} 0$, where

$$\mathcal{V}(x) = \int \log\left(1 + \frac{b}{x}\right) dF^{\mathbf{B}_N}(b) \tag{5}$$

and

$$\begin{aligned}
\mathcal{V}^{(0)}(x) = &\frac{1}{N} \log \det \left(\mathbf{I}_N + \frac{1}{x} \sum_{k=1}^{K} \mathbf{R}_k \int \frac{\tau_k dF^{\mathbf{T}_k}(\tau_k)}{1 + c_k e_k(-x)\tau_k} \right) \\
&+ \sum_{k=1}^{K} \frac{1}{c_k} \int \log\left(1 + c_k e_k(-x)\tau_k\right) dF^{\mathbf{T}_k}(\tau_k) \\
&+ x \cdot m_N^{(0)}(-x) - 1
\end{aligned} \tag{6}$$

A proof of this result (with less stringent hypothesis) is provided in [2]. From both results, one can derive the capacity formulation of a large range of multi-user/multi-cell network models. In the following, we cast attention on single-cell networks with multi-cell interference.

III. SYSTEM MODEL

In this work we derive capacity expressions of wireless channels between a multi-antenna transmitter and a multi-antenna receiver, the latter of which is interfered by several multi-antenna transmitters. This scheme is well-suited to multi-cell wireless networks with orthogonal intra-cell and interfering inter-cell transmissions, both in downlink and in uplink. The following scenarios encompass in particular

- multi-cell uplink: the base station of a cell indexed by $i \in \{1, \ldots, K\}$ receives data from one user in this cell[2] and is interfered by $K - 1$ users transmitting on the same physical resource from remote cells indexed by $j \in \{1, \ldots, K\}$, $j \neq i$.
- multi-cell downlink: the user being allocated a given time/frequency resource in a cell indexed by $i \in \{1, \ldots, K\}$ receives data from its dedicated base-station and is interfered by $K - 1$ base stations in neighboring cells indexed by $j \in \{1, \ldots, K\}$, $j \neq i$.

In the following, in order not to confuse both scenarios, only the downlink scheme is considered. However, one must keep in mind that the provided results can easily be adapted to the uplink case.

Consider a wireless mobile network with $K \geq 1$ cells indexed from 1 to K, controlled by *non-physically connected* base stations. On a particular time/frequency resource, each base station serves only one user; therefore the base station and the user of cell j will also be indexed by j. Without loss of generality, we focus our attention on user 1, equipped with $n_{\text{R}} \gg K$ antennas and hereafter referred to as *the user* or *the receiver*. Every base station $j \in \{1, \ldots, K\}$ is equipped with $n_{\text{T}_j} \gg K$ antennas. We additionally denote $c_j = n_{\text{T}_j}/n_{\text{R}}$.

Denote $\mathbf{s}_j \in \mathbb{C}^{n_{\text{T}_j}}$, $\mathsf{E}[\mathbf{s}_j \mathbf{s}_j^{\mathsf{H}}] = \mathbf{I}_{n_{\text{T}_j}}$, the signal transmitted by user j, $\mathbf{y} \in \mathbb{C}^{n_{\text{R}}}$ the signal received by the base station

[2]this user is allocated a given time/frequency resource, which is orthogonal to time/frequency resources of the other users in the cell; e.g. the multi-access protocol is OFDMA.

and $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{n_\mathrm{R}})$ the noise vector received at the base station. The fading MIMO channel between base station j and the user is denoted $\mathbf{H}_j \in \mathbb{C}^{n_\mathrm{R} \times n_{\mathrm{T}_j}}$. Moreover we assume that $\mathbf{H}_j$ has a separable variance profile, i.e. can be decomposed as

$$\mathbf{H}_j = \mathbf{R}_j^{\frac{1}{2}} \mathbf{X}_j \mathbf{T}_j^{\frac{1}{2}} \qquad (7)$$

with $\mathbf{R}_j \in \mathbb{C}^{n_\mathrm{R} \times n_\mathrm{R}}$ the (Hermitian) correlation matrix at the receiver with respect to the channel $\mathbf{H}_j$, $\mathbf{T}_j \in \mathbb{C}^{n_{\mathrm{T}_j} \times n_{\mathrm{T}_j}}$ the correlation matrix at transmitter j and $\mathbf{X}_j \in \mathbb{C}^{n_\mathrm{R} \times n_{\mathrm{T}_j}}$ a random matrix with Gaussian independent entries of variance $1/n_{\mathrm{T}_j}$.

Remark 2: Note that in this model, and contrary to what is often assumed, $\mathbf{R}_j$, the correlation matrix at the receiver, explicitly depends on j. In the uplink scenario, this assumption is of particular relevance in the sense that base stations are usually placed in areas clear of scatterers. In these circumstances, the solid angle from which the signals from user j originate influences the signal correlation at the receive antenna array. Hence the dependence of the receive correlation matrices on j. Note moreover that, in this model, the transmit power assumption $\mathsf{E}[\mathbf{s}_j \mathbf{s}_j^\mathsf{H}] = \mathbf{I}_{n_{\mathrm{T}_j}}$ is not restrictive in the sense that the transmit power correlation of user j can be included into the matrix $\mathbf{T}_j$. However, the Kronecker model has two major drawbacks: (i) the inner matrix $\mathbf{X}_j$ implicitly assumes a high density of scatterers[3] in the communication link and (ii) the correlations on both sides must be inter-independent and independent of the realizations of $\mathbf{X}_j$, which is inaccurate to some extent.

With the assumptions above, the communication model unfolds

$$\mathbf{y} = \mathbf{H}_1 \mathbf{s}_1 + \sum_{j=2}^{K} \mathbf{H}_j \mathbf{s}_j + \mathbf{n} \qquad (8)$$

where $\mathbf{s}_1$ is the useful signal (from base station 1) and $\mathbf{s}_j$, $j \geq 2$, constitute interfering signals.

IV. MULTI-CELL MIMO CAPACITY

A. Optimal Decoding

If the receiving user considers the signals from the $K-1$ interfering transmitters as pure noise and knows the exact value of the SNR (signal-to-noise ratio) σ^{-2}, then base station 1 can transmit with arbitrarily low decoding error at a per-receive antenna rate $C_\mathrm{SU}(\sigma^2)$ given by

$$C_\mathrm{SU}(\sigma^2) = \frac{1}{n_\mathrm{R}} \log_2 \left| \mathbf{I}_{n_\mathrm{R}} + \frac{1}{\sigma^2} \sum_{j=1}^{K} \mathbf{H}_j \mathbf{H}_j^\mathsf{H} \right|$$
$$- \frac{1}{n_\mathrm{R}} \log_2 \left| \mathbf{I}_{n_\mathrm{R}} + \frac{1}{\sigma^2} \sum_{j=2}^{K} \mathbf{H}_j \mathbf{H}_j^\mathsf{H} \right| \qquad (9)$$

Assume that n_R and the n_{T_i}, $i \in \{1, \dots, K\}$, are large compared to K and such that no eigenvalue of $\mathbf{R}_i$ or $\mathbf{T}_i$ is too

---

[3]the number and distance between scatterers must be of the same order as the number and distance between the transmit and receive antennas.

---

large. As in Theorem 1, we define the function $m^{i,(0)}$ as the approximated Stieltjes transforms of $\sum_{j=i}^{K} \mathbf{H}_j \mathbf{H}_j^\mathsf{H}$, $i \in \{1, 2\}$,

$$m^{i,(0)}(z) = \frac{1}{n_\mathrm{R}} \mathrm{tr} \left(\sum_{k=i}^{K} \int \frac{t_k dF^{\mathbf{T}_k}(t_k)}{1 + \frac{t_k}{c_k} e_k^i(z)} \mathbf{R}_j - z\mathbf{I}_{n_\mathrm{R}} \right)^{-1} \qquad (10)$$

where, for all $i \in \{1, \dots, K\}$, $e_j^i(z)$ is solution of the fixed-point equation

$$e_j^i(z) = \frac{1}{n_\mathrm{R}} \mathrm{tr} \, \mathbf{R}_j \left(\sum_{k=1}^{K} \int \frac{t_k dF^{\mathbf{T}_k}(t_k)}{1 + \frac{t_k}{c_k} e_k^i(z)} \mathbf{R}_k - z\mathbf{I} \right)^{-1} \qquad (11)$$

From Theorem 2, we then have approximately

$$C_\mathrm{SU}(\sigma^2) = \frac{1}{n_\mathrm{R}} \log \det \left(\mathbf{I} + \frac{1}{\sigma^2} \sum_{k=1}^{K} \mathbf{R}_k \int \frac{\tau_k dF^{\mathbf{T}_k}(\tau_k)}{1 + c_k e_k^1(-\sigma^2)\tau_k} \right)$$
$$- \frac{1}{n_\mathrm{R}} \log \det \left(\mathbf{I} + \frac{1}{\sigma^2} \sum_{k=2}^{K} \mathbf{R}_k \int \frac{\tau_k dF^{\mathbf{T}_k}(\tau_k)}{1 + c_k e_k^2(-\sigma^2)\tau_k} \right)$$
$$+ \sum_{k=1}^{K} \frac{1}{c_k} \int \log \left(1 + c_k e_k^1(-\sigma^2)\tau_k \right) dF^{\mathbf{T}_k}(\tau_k)$$
$$- \sum_{k=2}^{K} \frac{1}{c_k} \int \log \left(1 + c_k e_k^2(-\sigma^2)\tau_k \right) dF^{\mathbf{T}_k}(\tau_k)$$
$$+ \sigma^2 \cdot [m^{1,(0)}(-\sigma^2) - m^{2,(0)}(-\sigma^2)] \qquad (12)$$

However, this capacity expression assumes no specific power allocation at the transmitter. If the transmit covariance matrices $\mathbf{T}_j$ are replaced by $\mathbf{T}_j^{\frac{1}{2}} \mathbf{P}_j \mathbf{T}_j^{\frac{1}{2}}$ with $\mathbf{P}_j$ the signal covariance matrix at transmitter j, we remark that (9) can be rewritten as

$$C_\mathrm{SU}(\sigma^2) = \frac{1}{n_\mathrm{R}} \log \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{A}^{-\frac{1}{2}} \mathbf{R}_1^{\frac{1}{2}} \mathbf{X}_1 \mathbf{T}_1^{\frac{1}{2}} \mathbf{P}_1 \mathbf{T}_1^{\frac{1}{2}} \mathbf{X}_1^\mathsf{H} \mathbf{R}_1^{\frac{1}{2}} \mathbf{A}^{-\frac{1}{2}} \right| \qquad (13)$$

with $\mathbf{A} = \mathbf{I}_{n_\mathrm{R}} + \frac{1}{\sigma^2} \sum_{j>1} \mathbf{H}_j \mathbf{P}_j \mathbf{H}_j^\mathsf{H}$.

From [12], the optimal power allocation for base station 1 consists in aligning the eigenvectors of $\mathbf{P}_1$ to those of $\mathbf{T}_1$ and in choosing the eigenvalues $p_1, \dots, p_{n_{\mathrm{T}_1}}$ as

$$\begin{cases} p_i = 0 & (1/\alpha_i) - 1 \leq \frac{1}{n_{\mathrm{T}_1}} \sum_{l=1}^{n_{\mathrm{T}_1}} (1 - \alpha_l) \\ p_i = \frac{1 - \alpha_i}{\frac{1}{n_{\mathrm{T}_1}} \sum_{l=1}^{n_{\mathrm{T}_1}} (1 - \alpha_l)} & \text{otherwise} \end{cases} \qquad (14)$$

for α_i^{-1} equal to

$$1 + \frac{1}{\sigma^2} \mathbf{h}_i^\mathsf{H} \mathbf{A}^{-\frac{1}{2}} \left[\mathbf{I}_N + \frac{1}{\sigma^2} \mathbf{A}^{-\frac{1}{2}} \mathbf{H}_{-i} \mathbf{P}_1 \mathbf{H}_{-i}^\mathsf{H} \mathbf{A}^{-\frac{1}{2}} \right]^{-1} \mathbf{A}^{-\frac{1}{2}} \mathbf{h}_i \qquad (15)$$

$\mathbf{h}_i$ is the ith column of $\mathbf{H}_1$, and $\mathbf{H}_{-i}$ is $\mathbf{H}_1$ with column i removed.

Denoting $\mathbf{h}_i = \mathbf{R}_1^{\frac{1}{2}} \mathbf{x}_i$, we have $\mathbf{x}_i$ centered Gaussian with covariance $T_{1ii}/n_{\mathrm{T}_1} \mathbf{I}_{n_\mathrm{R}}$ and independent of $\mathbf{R}_1^{\frac{1}{2}} \mathbf{A}^{-\frac{1}{2}} \left[\mathbf{I}_N + \frac{1}{\sigma^2} \mathbf{A}^{-\frac{1}{2}} \mathbf{H}_{-i} \mathbf{P}_1 \mathbf{H}_{-i}^\mathsf{H} \mathbf{A}^{-\frac{1}{2}} \right]^{-1} \mathbf{A}^{-\frac{1}{2}} \mathbf{R}_1^{\frac{1}{2}}$.

Therefore, asymptotically on n_R, from lemmas in [1],

$$\alpha_i = \left(1 + \frac{T_{1_{ii}}}{\sigma^2 n_R} \operatorname{tr} \mathbf{R}_1 \left[\mathbf{A} + \frac{1}{\sigma^2}\mathbf{H}_{-i}\mathbf{P}_1\mathbf{H}_{-i}^{\mathsf{H}}\right]^{-1}\right)^{-1} \quad (16)$$

$$= \left(1 + \frac{T_{1_{ii}}}{\sigma^2 n_R} \operatorname{tr} \mathbf{R}_1 \left[\mathbf{A} + \frac{1}{\sigma^2}\mathbf{H}_1\mathbf{P}_1\mathbf{H}_1^{\mathsf{H}}\right]^{-1}\right)^{-1} \quad (17)$$

$$= \left(1 + \frac{T_{1_{ii}}}{\sigma^2 n_R} \operatorname{tr} \mathbf{R}_1 \left[\mathbf{I} + \frac{1}{\sigma^2}\sum_{j=1}^{K}\mathbf{H}_j\mathbf{P}_j\mathbf{H}_j^{\mathsf{H}}\right]^{-1}\right)^{-1} \quad (18)$$

$$= \left(1 + T_{1_{ii}}e_1^1(-\sigma^2)\right)^{-1} \quad (19)$$

This leads to the power allocation

$$\begin{cases} p_i = 0, \ T_{1_{ii}}e_1^1(-\sigma^2) \le \frac{1}{n_{T_1}}\sum_{l=1}^{n_{T_1}} 1 - \left(1 + T_{1_{ll}}e_1^1(-\sigma^2)\right)^{-1} \\ p_i = \frac{1 - \left(1 + T_{1_{ii}}e_1^1(-\sigma^2)\right)^{-1}}{\frac{1}{n_{T_1}}\sum_{l=1}^{n_{T_1}} 1 - \left(1 + T_{1_{ll}}e_1^1(-\sigma^2)\right)^{-1}} \text{ otherwise} \end{cases} \quad (20)$$

In most cases however, the diagonal entries of $\mathbf{T}_1$ all equal 1, and therefore the optimal power allocation policy is uniform array power allocation.

B. MMSE Decoder

Achieving C_{SU} requires non-linear processing at the receiver, such as MMSE successive interference cancellation. A suboptimal linear technique, the MMSE decoder, is often used instead. The communication model in this case reads

$$\mathbf{y} = \left(\sum_{j=1}^{k}\mathbf{H}_j\mathbf{H}_j^{\mathsf{H}} + \sigma^2\mathbf{I}_{n_R}\right)^{-1}\mathbf{H}_1^{\mathsf{H}}\left(\sum_{j=1}^{k}\mathbf{H}_j\mathbf{s}_j + \mathbf{n}\right) \quad (21)$$

and each entry of $\mathbf{y}$ will be processed individually.

This technique makes it possible to transmit data reliably at any rate inferior to the per-antenna MMSE capacity C_{MMSE},

$$C_{\mathrm{MMSE}}(\sigma^2) = \frac{1}{n_R}\sum_{i=1}^{n_{T_1}}\log_2(1 + \gamma_i) \quad (22)$$

where, denoting $\mathbf{h}_j \in \mathbb{C}^{n_{T_j}}$ the jth column of $\mathbf{H}_1$ and $\mathbf{R}_1^{\frac{1}{2}}\mathbf{x}_j = \mathbf{h}_j$, the SINR γ_i expresses as

$$\gamma_i = \frac{\mathbf{h}_i^{\mathsf{H}}\left(\sum_{j=1}^{K}\mathbf{H}_j\mathbf{H}_j^{\mathsf{H}} + \sigma^2\mathbf{I}_{n_R}\right)^{-1}\mathbf{h}_i}{1 - \mathbf{h}_i^{\mathsf{H}}\left(\sum_{j=1}^{K}\mathbf{H}_j\mathbf{H}_j^{\mathsf{H}} + \sigma^2\mathbf{I}_{n_R}\right)^{-1}\mathbf{h}_i} \quad (23)$$

$$= \mathbf{h}_i^{\mathsf{H}}\left(\sum_{j=1}^{K}\mathbf{H}_j\mathbf{H}_j^{\mathsf{H}} - \mathbf{h}_i\mathbf{h}_i^{\mathsf{H}} + \sigma^2\mathbf{I}_{n_R}\right)^{-1}\mathbf{h}_i \quad (24)$$

$$= \mathbf{x}_i^{\mathsf{H}}\mathbf{R}_i^{\frac{1}{2}}\left(\sum_{j=1}^{K}\mathbf{H}_j\mathbf{H}_j^{\mathsf{H}} - \mathbf{h}_i\mathbf{h}_i^{\mathsf{H}} + \sigma^2\mathbf{I}_{n_R}\right)^{-1}\mathbf{R}_i^{\frac{1}{2}}\mathbf{x}_i \quad (25)$$

where Equation (24) comes from a direct application of the matrix inversion lemma. With these notations, $\mathbf{x}_i$ has i.i.d. complex Gaussian entries with variance $T_{1_{ii}}/n_{T_i}$ and the inner

matrix of the right-hand side of (25) is independent of $\mathbf{x}_i$ (since the entries of $\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}} - \mathbf{h}_i\mathbf{h}_i^{\mathsf{H}}$ are independent of the entries $\mathbf{h}_i$). Applying Lemma 3.1 in [1], for n_{T_i} large, approximately

$$\gamma_i = \frac{T_{1_{ii}}}{n_{T_1}}\operatorname{tr}\mathbf{R}_1\left(\sum_{j=1}^{K}\mathbf{H}_j\mathbf{H}_j^{\mathsf{H}} - \mathbf{h}_i\mathbf{h}_i^{\mathsf{H}} + \sigma^2\mathbf{I}_{n_R}\right)^{-1} \quad (26)$$

From Lemma 2.1 in [11], the rank 1 perturbation $(-\mathbf{h}_i\mathbf{h}_i^{\mathsf{H}})$ does not affect asymptotically the trace in (26). And therefore, approximately,

$$\gamma_i = \frac{T_{1_{ii}}}{n_{T_1}}\operatorname{tr}\mathbf{R}_1\left(\sum_{j=1}^{K}\mathbf{H}_j\mathbf{H}_j^{\mathsf{H}} + \sigma^2\mathbf{I}_{n_R}\right)^{-1} \quad (27)$$

Noting that $e_1(z)$ in Section IV-A corresponds to the normalized trace in Equation (27) (this is shown precisely in the proof of Theorem 1 [2]), we finally have the compact expression for C_{MMSE},

$$C_{\mathrm{MMSE}}(\sigma^2) = \frac{1}{n_R}\sum_{i=1}^{n_{T_1}}\log_2\left(1 + \frac{1}{c_1}T_{1_{ii}}e_1(-\sigma^2)\right) \quad (28)$$

In practice, when no power allocation strategy is applied, $T_{1_{ii}} = P$ the average power per transmit symbol, and the capacity becomes $C_{\mathrm{MMSE}} = c_1 \cdot \log_2(1 + \frac{P}{c_1}e_1(-\sigma^2))$.

V. SIMULATION AND RESULTS

In the following, we apply the results (12) and (28) to the downlink of a two-cell network. The capacity analyzed here is the achievable rate on the link between base station 1 and the user, the latter of which is interfered by base station 2. The relative power of the signal received from user 2 is on average Γ times that of user 1. Both base stations 1 and 2 are equipped with linear arrays of n_T antennas and the user with a linear array of n_R antennas. The correlation matrices $\mathbf{T}_i$ at the transmission and $\mathbf{R}_i$ at the reception, $i \in \{1, 2\}$, are modeled thanks to a generalization of Jake's model including solid angles of transmit/receive power, i.e. for instance,

$$T_{i_{ab}} = \int_{\theta_{\min}^{(i)}}^{\theta_{\max}^{(i)}}\exp\left(2\pi \cdot i \cdot \frac{d_{ab}^{T_i}}{\lambda}\cos(\theta)\right)d\theta \quad (29)$$

with $d_{ab}^{T_i}$ the distances between antennas indexed by $a, b \in \{1, \ldots, n_{T_i}\}$ for transmitter i, $(\theta_{\min}^{(i)}, \theta_{\max}^{(i)})$ the angles over which useful power (i.e. power that will be received by user 1) is transmitted, and λ the wavelength.

In Figure 1, we took $n_R = 16$, $\Gamma = 0.25$ and we consider optimal single-user decoding at the receiver. For every realization of $\mathbf{T}_i$, $\mathbf{R}_i$, 1000 channel realizations are processed to produce the simulated ergodic capacity and compared to the theoretical capacity (28). Those capacities are then averaged over 100 realizations of $\mathbf{T}_i$, $\mathbf{R}_i$, varying in the random choice of $\theta_{\min}^{(i)}$ and $\theta_{\max}^{(i)}$ with constraint $\theta_{\max}^{(i)} - \theta_{\min}^{(i)} = \pi/2$, while $d_{ab}^{T_i} = 10\lambda|a - b|$ at the transmitters, $d_{ab}^{R} = 2\lambda|a - b|$ at the receiver. The SNR ranges from -5 dB to 30 dB, and $n_T \in \{8, 16\}$. We observe here that Monte-Carlo simulations perfectly match the capacity obtained from Equation (12).
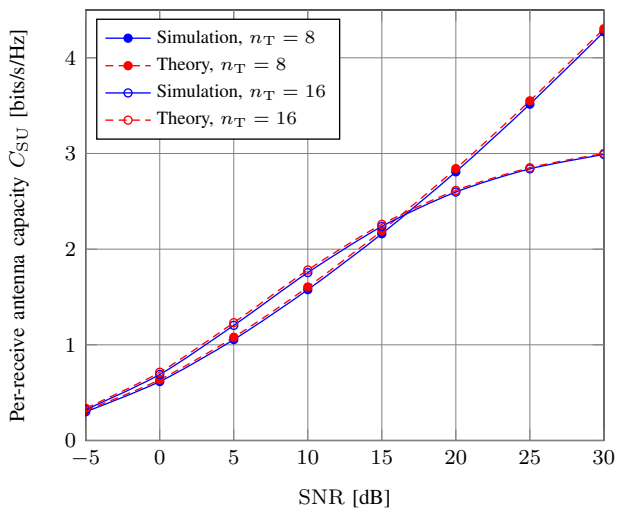
Fig. 1. Capacity of point-to-point MIMO in two-cell uplink, optimal single-user decoding, $n_R = 16$, $n_T \in \{8, 16\}$, $\Gamma = 25\%$.

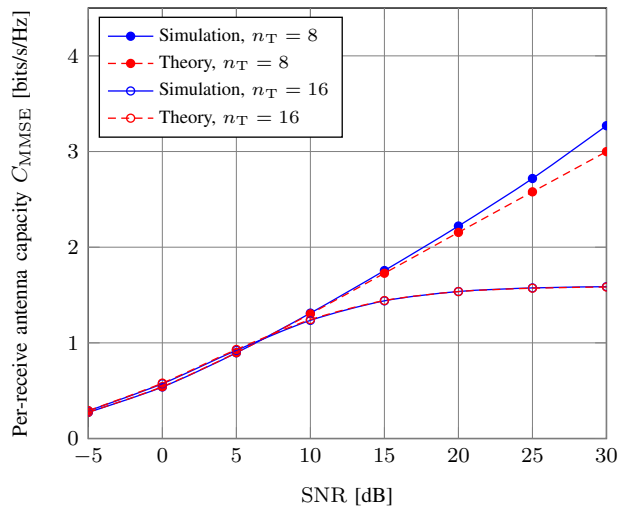

Fig. 2. Capacity of point-to-point MIMO in two-cell uplink, with MMSE decoder, $n_R = 16$, $n_T \in \{8, 16\}$, $\Gamma = 25\%$.

In Figure 2, with the same assumptions as previously, we apply MMSE decoding at the base station. Here, a slight difference is observed in the high SNR regime between theory and practice. This was somehow expected, since the large n_R approximations in Silverstein's lemmas [1] are very loose for σ^2 close to $\mathbb{R}^-$ in the sense of the Euclidean distance. To cope with this gap, many more antennas must be used. We also observe a significant difference in performance between optimum and linear MMSE decoders, especially in the high SNR region. Therefore, in wireless networks, when interfering cells are treated as Gaussian correlated noise at the cell-edge, i.e. where the interference is maximum, the MMSE decoder provides tremendous performance loss.

VI. Conclusion

In this paper, we introduced an important theorem relating the Stieltjes transform of a class of large matrices to a deterministic approximate. Based on this formula, we provided compact expressions for the optimal and MMSE-decoder capacities of point-to-point MIMO with inter-cell interference, for random channel matrices with separable variance profile, both in downlink and in uplink. The simulations show perfect match with the theoretical formulas in the low-SNR region, even if fewer antennas are used at the transmitters and receivers. As for the high SNR region, a large number of antennas must be used to reach an accurate match between theory and Monte-Carlo simulations.

VII. Acknowledgment

References

[1] J. Silverstein and Z. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, issue 2, pp. 175–192, 1995

[2] R. Couillet, M. Debbah and J. W. Silverstein, "A deterministic equivalent approach for the capacity analysis of multi-user MIMO channels," *submitted to IEEE Trans. on Information Theory*.

[3] G. Foschini and M. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," Wireless Personal Communications, vol. 6, no. 3, pp. 311-335, 1998.

[4] E. Telatar, "Capacity of multi-antenna Gaussian channels," European transactions on telecommunications, vol. 10, no. 6, pp. 585-595, 1999.

[5] A. Paulraj, R. Nabar and D. Gore, "Introduction to Space-Time Wireless Communications," Cambridge University Press, 2003.

[6] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity Scaling in MIMO Wireless Systems under Correlated Fading," IEEE Trans. on Information Theory, pp. 637650, 2002.

[7] A. Moustakas, S. Simon and A. 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, vol. 49, no. 10, 2003.

[8] D. Guo and S. Verdú, "Multiuser Detection and Statistical Physics," Communications on Information and Network Security, Kluwer Academic Publishers, 2003.

[9] S. Shamai and A. Wyner, "Information-Theoretic Considerations for Symmetric, Cellular, Multiple-Access Fading Channels - Part I," IEEE Trans. on Information Theory, vol. 43, no. 6, 1997.

[10] B. Zaidel, S. Shamai and S. Verdú, "Multicell Uplink Spectral Efficiency of Coded DS-CDMA With Random Signatures," IEEE Journal on Selected Areas in Communications, vol. 19, no. 8, 2001.

[11] Z. Bai and J. Silverstein, "On the signal-to-interference-ratio of CDMA systems in wireless communications," Annals of Applied Probability vol. 17 no. 1, pp. 81-101, 2007.

[12] A.M. Tulino, A. Lozano and S. Verdu, "Impact of antenna correlation on the capacity of multiantenna channels," IEEE Trans. on Information Theory, vol. 51, no. 7, pp. 2491-2509, 2005.

[13] M. J. M. Peacock, I. B. Collings and M. L. Honig, "Eigenvalue distributions of sums and products of large random matrices via incremental matrix expansions," IEEE Trans. on Information Theory, vol. 54, no. 5, pp. 2123-2138, 2008.