

A CLT for capacity inference methods under colored interference

Abla Kammoun¹, Romain Couillet², Jamal Najim¹ and M erouane Debbah²
¹Telecom ParisTech, ²Alcatel-Lucent Chair, Sup elec.

Abstract—In this paper, we address the problem of fast point-to-point channel capacity estimation in the case where the receiver undergoes unknown interference from multiple sources, whereas the channel with the transmitter is perfectly known. For this particular context, we propose a fast estimator for the capacity estimation, and compare its performance with that of the traditional methods. More precisely, we analyse the fluctuations of the traditional and proposed techniques and prove that their behaviors can be approximated by Gaussian random variables for which we derive the variances.

I. INTRODUCTION

As opposed to wired transmissions, communication over the wireless medium is more sensitive to the channel and interference. This is principally attributed to the intrinsic nature of the air medium, which is richly scattered and obviously shared by multiple users. Since the environment is changing fast, it is fundamental for users to rapidly estimate the maximum rate which can be achieved in the communication to other users. This is essential for guaranteeing a reliable data transmission at all time instant. Classical estimation techniques, based on the assumption of a large number of observations, are however usually inappropriate. In order to design improved estimates, the use of the theory of large random matrices, particularly suitable in case the space and time dimensions are of the same magnitude, has been recently pushed forward. As far as the capacity estimation is concerned, several studies have focused on the capacity estimation of MIMO systems, in case of imperfect channel knowledge, but without interference. This scenario has been considered in [1] and [2], where methods based respectively on free probability theory and deterministic equivalents have been proposed.

In this paper, we will consider a different situation where the receiver knows perfectly the channel with the transmitter but does not know anything about the experienced interference. Such a situation can be encountered in multi-cell scenarios, where interference stemming from neighboring cell users changes fast, which is a natural assumption in packet switch transmissions. The estimated capacity can serve first as an upper-bound for the maximum rate that could be achieved. Indeed, this rate cannot be achieved if the channel interference is not exactly estimated and therefore the estimator may serve only as an approximate achievable performance. Another usage is found in the context of cognitive radios where multiple frequency bands are sensed for future transmissions. In this setting, the proposed estimator provides the expected rate performance achievable in each frequency band.

In this paper, we derive the expression of a consistent estimator of the capacity and prove using simulations that it outperforms by far traditional estimators. In addition to consistency, second order statistics of the estimator are an important performance index that measure its reliability. We study thus the fluctuations of the traditional and improved estimators.

The remainder of this paper is divided as follows: In section II, we present the system model and derive the expression of the consistent estimator. In the next two sections, we establish our results regarding the central limit theorem (CLT) for both estimators. Finally, we provide in section V some numerical simulations that support the accuracy of the derived results.

Notations: In the following, boldface lower case symbols represent vectors, capital boldface characters denote matrices (\mathbf{I}_N is the size- N identity matrix). The transpose and Hermitian transpose operators are denoted $(\cdot)^T$ and $(\cdot)^H$, respectively. The symbol ‘ $\xrightarrow{\text{a.s.}}$ ’ denotes almost sure convergence.

II. SYSTEM MODEL AND PROBLEM SETTING

We consider a communication link between two users equipped respectively with N and n_0 antennas and referred to as the *receiver* and the *transmitter*. We assume also the presence of K interferers with n_k antennas each. In practice, this situation is encountered e.g. in the cognitive radio context, where the receiver and transmitter represent two secondary users, whereas the interferers stand for primary users. Note that the secondary users are supposed to adjust their power in order to not incur any interference to the primary network. Figure 1 describes this scenario. Denote by \mathbf{H} and \mathbf{G}_k the $N \times n_0$ and $N \times n_k$ channel matrices between the receiver and, respectively, the transmitter and the k -th interferer. We assume that \mathbf{H} and \mathbf{G}_k remain static during at least M data transmission symbols.

Let $\bar{\mathbf{y}}_m$ denote the N -dimensional vector received at time m by the receiver then $\bar{\mathbf{y}}_m$ writes:

$$\bar{\mathbf{y}}_m = \mathbf{H}\mathbf{x}_0^{(m)} + \sum_{k=1}^K \mathbf{G}_k \mathbf{x}_k^{(m)} + \sigma \mathbf{w}^{(m)},$$

where $\mathbf{x}_0^{(m)}$ is the signal transmitted at time m by the transmitter, $\mathbf{x}_k^{(m)} \sim \mathcal{CN}(0, \mathbf{I}_{n_k})$ is the signal transmitted at time m by the k -th interferer, and $\mathbf{w}^{(m)}$ is the additive Gaussian noise experienced at time m ($\mathbf{w}^{(m)} \sim \mathcal{CN}(0, \mathbf{I}_N)$). Assuming a perfect decoding of $\mathbf{x}_0^{(m)}$ initially transmitted at low rate, and a perfect knowledge of the channel matrix \mathbf{H} , the residual

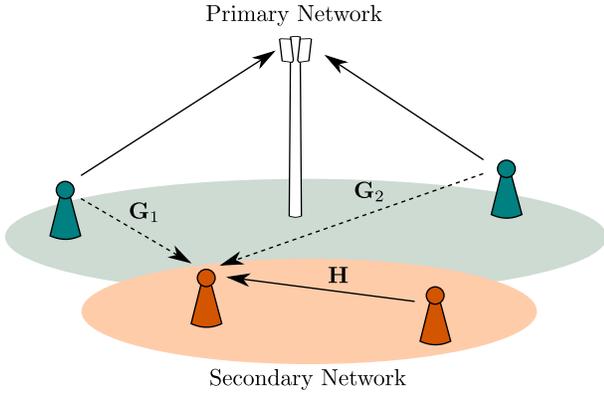


Fig. 1. System model.

interference to which the receiver has access is given by:

$$\mathbf{y}_m = \bar{\mathbf{y}}_m - \mathbf{H}\mathbf{x}_0^{(m)} = \sum_{k=1}^K \mathbf{G}_k \mathbf{x}_k^{(m)} + \sigma \mathbf{w}^{(m)}.$$

This coincides also with the signal received at time m in case of no secondary transmission.

The per-antenna maximum data rate that can be reliably decoded by the receiver is given by:

$$\begin{aligned} C(\sigma^2) &= \frac{1}{N} \log \det \left(\mathbf{I}_N + \left[\sum_{k=1}^K \mathbf{G}_k \mathbf{G}_k^H + \sigma^2 \mathbf{I}_N \right]^{-1} \mathbf{H} \mathbf{H}^H \right) \\ &= \frac{1}{N} \log \det \left(\sigma^2 \mathbf{I}_N + \sum_{k=1}^K \mathbf{G}_k \mathbf{G}_k^H + \mathbf{H} \mathbf{H}^H \right) \\ &\quad - \frac{1}{N} \log \det \left(\sigma^2 \mathbf{I}_N + \sum_{k=1}^K \mathbf{G}_k \mathbf{G}_k^H \right). \end{aligned} \quad (1)$$

Based on restricted number M of successive observations $\mathbf{y}_1, \dots, \mathbf{y}_M$ and perfect knowledge of \mathbf{H} , the receiver needs to infer $C(\sigma^2)$ in an efficient way. By efficient, we mean that the estimation should be accurate and fast, in the sense that it should only require a number of slots of the same order of magnitude as N . Obviously, under this assumption substituting $\sigma^2 \mathbf{I}_N + \sum_{k=1}^K \mathbf{G}_k \mathbf{G}_k^H$ by $\frac{1}{M} \sum_{m=1}^M \mathbf{y}_m \mathbf{y}_m^H$ is not efficient. It can be viewed as a good alternative if the number of slots M is too large with respect to the number of receiving antennas N . In the sequel, we refer to the estimator that assumes $M \gg N$ as the *large- M* estimator. It is given by:

$$\begin{aligned} C_L(\sigma^2) &= \frac{1}{N} \log \det \left(\frac{1}{M} \sum_{m=1}^M \mathbf{y}_m \mathbf{y}_m^H + \mathbf{H} \mathbf{H}^H \right) \\ &\quad - \frac{1}{N} \log \det \left(\frac{1}{M} \sum_{m=1}^M \mathbf{y}_m \mathbf{y}_m^H \right), \end{aligned} \quad (2)$$

(the subscript L refers here to large- M), Let $n = \sum_{k=1}^K n_k$ and consider the following asymptotic regime: $n \propto M$, and $N \propto M$ as $n, M, N \rightarrow \infty$. Formally, this means:

$$\begin{aligned} N, n, M \rightarrow \infty \text{ with } 1 < \liminf M/N \leq \limsup M/N < \infty \\ 0 < \liminf N/n \leq \limsup N/n < \infty. \end{aligned}$$

In the following, we derive the expression of a consistent estimator of $C(\sigma^2)$, which we refer to as the G-estimator. In particular, we have the following result whose proof can be found in [3]:

Theorem 1: Consider the quantity:

$$\begin{aligned} \hat{C}(\sigma^2) &= \frac{1}{N} \log \det \left(\mathbf{I}_N + y_N \mathbf{H} \mathbf{H}^H \left(\frac{1}{M} \mathbf{Y} \mathbf{Y}^H \right)^{-1} \right) \\ &\quad + \frac{(M-N)}{N} \left[\log \left(\frac{M}{M-N} y_N \right) + 1 \right] - \frac{M}{N} y_N, \end{aligned}$$

where $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M]$, and y_N the solution of :

$$y = \frac{1}{M} \text{tr} y \mathbf{H} \mathbf{H}^H \left(y \mathbf{H} \mathbf{H}^H + \frac{1}{M} \mathbf{Y} \mathbf{Y}^H \right)^{-1} + \frac{M-N}{M},$$

then,

$$\hat{C}(\sigma^2) - C(\sigma^2) \xrightarrow{\text{a.s.}} 0.$$

Let $\mathbf{G} = [\mathbf{G}_1, \dots, \mathbf{G}_K]$. We have:

$$y_N - y_N^* \xrightarrow{\text{a.s.}} 0,$$

where

$$y_N^* = 1 - \frac{1}{M} \text{tr} \left(\mathbf{Z} (\mathbf{H} \mathbf{H}^H + \mathbf{Z})^{-1} \right), \quad (3)$$

and $\mathbf{Z} = \mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I}_N$

III. BIAS AND PERFORMANCE OF THE LARGE- M ESTIMATOR

Obviously, in the asymptotic regime, the large- M estimator is asymptotically biased. Based on previous results [4], [5] about respectively the convergence of the capacity and the *logdet* function of a covariance matrix, we can prove that the large- M estimator converges almost surely to a deterministic quantity, which is different in general from the capacity $C(\sigma^2)$. Apart from the bias, another question regarding the behaviour of the large- M estimator around its asymptotic equivalent is of fundamental importance. This can provide for instance insights about the convergence speed or also about the variance or equivalently the spread of the estimator around its asymptotic equivalent. The objective of this section is to derive the expression of the bias and establish the CLT. Prior to that we need to recall some deterministic quantities, on which our results will depend:

Lemma 1 ([3]): Let $y > 0$. The following assertions hold true:

1) The following functional equation:

$$\kappa(y) = \frac{1}{M} \text{tr} \left(\mathbf{Z} \left(\frac{\mathbf{Z}}{1 + \kappa(y)} + y \mathbf{H} \mathbf{H}^H \right)^{-1} \right)$$

admits a unique positive solution $\kappa(y)$.

2) Denote by $\mathbf{T}(y) = \left(y \mathbf{H} \mathbf{H}^H + \frac{\mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I}_N}{1 + \kappa(y)} \right)^{-1}$ and $\mathbf{Q}(y) = \left(y \mathbf{H} \mathbf{H}^H + \frac{1}{M} \mathbf{Y} \mathbf{Y}^H \right)^{-1}$. Then, for any deterministic matrix $\mathbf{S} \in \mathbb{C}^N$ with uniformly bounded spectral norm, we have:

$$\frac{1}{M} \text{tr} (\mathbf{S} \mathbf{Q}(y)) - \frac{1}{M} \text{tr} (\mathbf{S} \mathbf{T}(y)) \xrightarrow{\text{a.s.}} 0.$$

Although $\mathbf{T}(y)$ approximates $\mathbf{Q}(y)$ in a certain sense, $-\frac{1}{N} \log \det(\mathbf{Q}(1))$ which stands for the first term in (2) is not asymptotically equivalent to $-\frac{1}{N} \log \det(\mathbf{T}(1))$, and in general a scalar functional f of $\mathbf{Q}(1)$ does not necessarily converge to $f(\mathbf{T}(1))$. To find a deterministic equivalent for $-\log \det(\mathbf{Q}(1))$, we will refer to [4]:

Lemma 2: Let

$$V(y) = -\log \det(\mathbf{T}(y)) + M \log(1 + \kappa(y)) - M \frac{\kappa(y)}{1 + \kappa(y)}.$$

Then in the asymptotic regime, we have:

$$-\frac{1}{N} \log \det(\mathbf{Q}(y)) - \frac{1}{N} V(y) \xrightarrow{\text{a.s.}} 0. \quad (4)$$

It remains now to deal with the second term in (2). For this, we recall [5, (1.1)]

Lemma 3: In the asymptotic regime we have:

$$\begin{aligned} & -\frac{1}{N} \log \det(\mathbf{Y}\mathbf{Y}^H) + \log \det(\mathbf{G}\mathbf{G}^H + \sigma^2 \mathbf{I}_N) \\ & + \frac{N-M}{N} \log \left(\frac{M-N}{M} \right) - 1 \xrightarrow{\text{a.s.}} 0. \end{aligned} \quad (5)$$

Combining (4) and (5), we finally have:

Lemma 4: Denote by $\mathcal{V}(y)$ the deterministic quantity given by:

$$\begin{aligned} \mathcal{V}(y) = & -\frac{1}{N} \log \det(\mathbf{T}(y)) + \frac{M}{N} \log(1 + \kappa(y)) - \frac{M}{N} \frac{\kappa(y)}{1 + \kappa(y)} \\ & - \frac{1}{N} \log \det(\mathbf{G}\mathbf{G}^H + \sigma^2 \mathbf{I}_N) + \frac{(M-N)}{N} \log \left(\frac{M-N}{M} \right) + 1. \end{aligned}$$

Then, the large-M estimator $C_L(\sigma^2) - \mathcal{V}(1)$ converges almost surely to zero as $N, M, n \rightarrow \infty$.

Theorem 4 provides an asymptotic equivalent of the large-M estimator. Hereafter, we state our main result in this paper, in which a CLT of

$$C_L(y) = \frac{1}{N} \log \det(y\mathbf{H}\mathbf{H}^H + \mathbf{Y}\mathbf{Y}^H) - \frac{1}{N} \log \det(\mathbf{Y}\mathbf{Y}^H)$$

is derived. As we will see later, this result will also serve to establish the CLT of the G-estimator.

Theorem 2: Let $C_L(y) = \frac{1}{N} \log \det(y\mathbf{H}\mathbf{H}^H + \mathbf{Y}\mathbf{Y}^H) - \frac{1}{N} \log \det(\mathbf{Y}\mathbf{Y}^H)$. Denote $Z_N(y) = N(C_L(y) - \mathcal{V}(1))$. Then, in the asymptotic regime, we have:

$$\frac{1}{\alpha_N(y)} Z_N(y) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1),$$

where $\alpha_N(y)$ is given by (6). In particular, we have:

$$\frac{1}{\alpha_N(1)} (NC_L(\sigma^2) - N\mathcal{V}(1)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Proof: See appendix B. ■

IV. PERFORMANCE OF THE G-ESTIMATOR

As opposed to the large-M estimation method, the improved G-estimator has no closed-form expression, in the sense that key parameter y_N is the solution of an implicit equation that is solved easily through numerical iterations. Establishing the CLT might seem to be more tricky since the randomness comes from both the received matrix \mathbf{Y} and y_N . As a first step,

one can look for the asymptotic behavior of y_N . This is the objective of the following lemma:

Lemma 5: The parameter y_N satisfies in the asymptotic regime the following assertions:

- 1) $\text{var}(y_N) = \mathcal{O}(M^{-2})$,
- 2) $Ey_N = y_N^* + \mathcal{O}(M^{-2})$.

Proof: See appendix A. ■

We are now in position to state the CLT for the G-estimator.

Theorem 3: In the asymptotic regime, the G-estimator satisfies:

$$\frac{N}{\theta_N} (\hat{C}(\sigma^2) - C(\sigma^2)) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1),$$

where θ_N is given by (7) with y_N^* defined in (3).

Proof: Let $\mathcal{C}(y)$ the function defined for $y > 0$ as:

$$\begin{aligned} \mathcal{C}(y) = & \frac{1}{N} \log \det \left(y\mathbf{H}\mathbf{H}^H + \frac{\mathbf{Y}\mathbf{Y}^H}{M} \right) \\ & + \frac{M-N}{N} \left[\log \left(\frac{M}{M-N} y \right) + 1 \right] - \frac{M}{N} y \\ & - \log \det \left(\frac{\mathbf{Y}\mathbf{Y}^H}{M} \right). \end{aligned}$$

Since $\frac{\partial \mathcal{C}}{\partial y} |_{y=y_N} = 0$, a Taylor expansion of $\mathcal{C}(y_N^*)$ around y_N yields:

$$N\mathcal{C}(y_N^*) = N\mathcal{C}(y_N) + \frac{N\partial^2 \mathcal{C}}{\partial y^2} (y_N^* - y_N)^2 + \mathcal{O}(N(y_N^* - y_N)^3).$$

Given that $E(y_N^* - y_N)^2 = \mathcal{O}(M^{-2})$, $N\mathcal{C}(y_N^*) - N\mathcal{C}(y_N)$ converges in probability to zero. By Slutsky theorem, it suffices to establish the CLT theorem for $N\mathcal{C}(y_N^*)$ instead of $N\mathcal{C}(y_N) = N\hat{C}(y_N)$. This is extremely helpful since unlike y_N , y_N^* is deterministic. The result is thus obtained by applying theorem 2 and noticing that $\kappa(y_N^*) + 1 = \frac{1}{y_N^*}$. ■

V. SIMULATION RESULTS

We consider in the simulations the case where a mobile terminal with $N = 4$ antennas received during $M = 15$ slots, data stemming from $n_0 = 4$ antenna secondary users. We assume that the communication link is interfered by $K = 8$ mono-antenna users. Matrices \mathbf{H} , $\mathbf{G}_1, \dots, \mathbf{G}_K$ are arbitrarily chosen and remain constant during the Monte Carlo averaging, while data and noise matrices are randomly chosen standard Gaussian matrices. Fig. 2 represents the theoretical and empirical normalized variance for the G-estimator with respect to $\text{SNR} = \frac{1}{\sigma^2}$. We also display in the same graph the empirical variance of the large-M estimator. We note that the G-estimator exhibits a better performance for all SNR. To assess the Gaussian behavior of both estimators, we represent in fig. 3 and fig. 4, the histogram in case the SNR is set to 10dB. We note a good fit between theoretical and empirical results, although the system dimensions are small.

VI. CONCLUSION

In this paper, we have addressed the question of fast estimation of the capacity in presence of interference. More precisely, we have established the Gaussian behaviour of a recently proposed method using G-estimation theory and of

$$\alpha_N(y) = 2 \log(M) - \log \left((M - N) \left(M(\kappa + 1)^2 - \text{tr} \left(\left(\frac{\mathbf{I}_N}{\kappa + 1} + y \mathbf{H} \mathbf{H}^H (\mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I}_N)^{-1} \right)^{-2} \right) \right) \right) \quad (6)$$

$$\theta_N = 2 \log(My_N^*) - \log \left((M - N) \left(M - \text{tr} \left((\mathbf{I}_N + \mathbf{H} \mathbf{H}^H (\mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I}_N)^{-1} \right)^{-2} \right) \right) \quad (7)$$

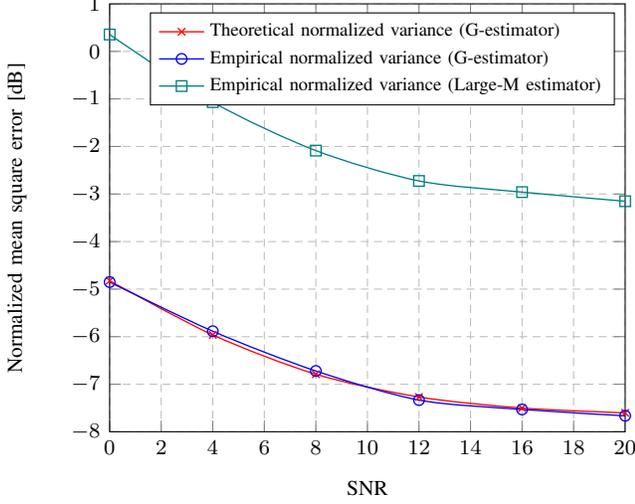


Fig. 2. Empirical and theoretical variances with respect to the SNR

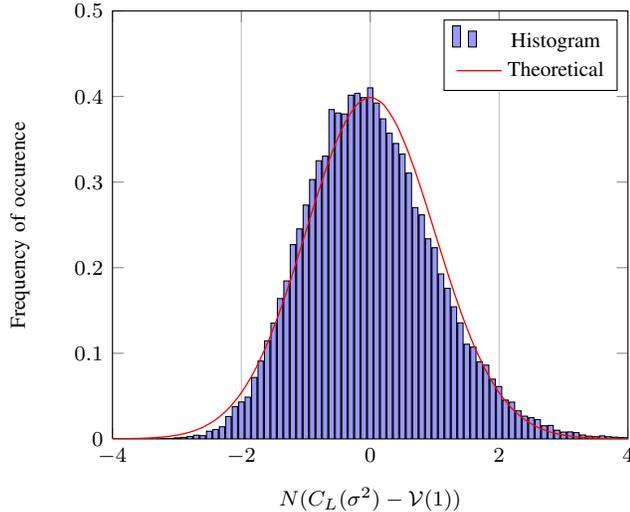


Fig. 3. Histogram of $N(C_L(\sigma^2) - \nu(1))$.

a large-M estimator. We have provided numerical simulations which strongly support the accuracy of our derived results even for usual system dimensions.

APPENDIX A PROOF OF LEMMA 5

1) Denote by $R(y)$ and $f(y)$ the functionals given by:

$$f(y) = \frac{1}{M} \text{tr}(y \mathbf{H} \mathbf{H}^H \mathbf{Q}(y)) + \frac{M - N}{M} - y$$

$$R(y) = -\log \det(\mathbf{Q}(y)) + (M - N) \log(y) - My.$$

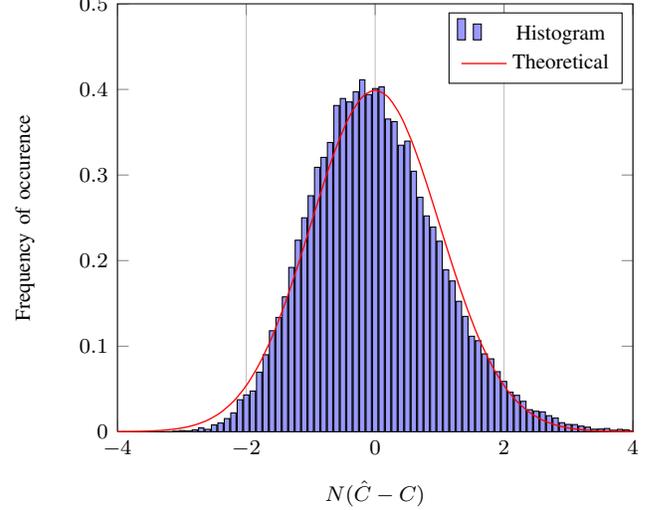


Fig. 4. Histogram of $N(\hat{C} - C)$.

According to the Nash-Poincaré inequality [4], we have:

$$\text{var}(y_N) \leq K \sum_{i=1}^N \sum_{j=1}^M \mathbb{E} \left| \frac{\partial y_N}{\partial Y_{i,j}^*} \right|^2 + \mathbb{E} \left| \frac{\partial y_N}{\partial Y_{i,j}} \right|^2. \quad (8)$$

We will only consider the first sum in the previous inequality. The second sum can be dealt with in the same way. Using the implicit function theorem, we know that if $\frac{\partial f}{\partial y} \neq 0$, $\frac{\partial y_N}{\partial Y_{i,j}^*}$ writes:

$$\frac{\partial y_N}{\partial Y_{i,j}^*} = \frac{\frac{\partial f}{\partial Y_{i,j}^*} |_{y=y_N}}{\frac{\partial f}{\partial y} |_{y=y_N}}. \quad (9)$$

As will be shown later, to conclude that $\text{var}(y_N) = \mathcal{O}(M^{-2})$, we need that $\left| \frac{\partial f}{\partial y} |_{y=y_N} \right|$ be lower bounded away from zero, which is a much stronger requirement than $\frac{\partial f}{\partial y} \neq 0$. This can be proved by noticing that $\frac{\partial R}{\partial y} = \frac{Mf}{y}$. Hence

$$\frac{\partial^2 R}{\partial y^2} |_{y=y_N} = \frac{M \frac{\partial f}{\partial y} |_{y=y_N}}{y_N}. \quad (10)$$

On the other hand, straightforward calculations lead to $\left| \frac{\partial^2 R}{\partial y^2} |_{y=y_N} \right| \geq \frac{M-N}{y_N^2}$ which, plugged into (10), yields:

$$\left| \frac{\partial f}{\partial y} \right| \geq \frac{M - N}{My_N} \geq \liminf_N \frac{M - N}{M}. \quad (11)$$

Therefore,

$$\begin{aligned} \sum_{i=1}^N \sum_{j=1}^M \mathbb{E} \left| \frac{\partial y_N}{\partial Y_{i,j}^*} \right|^2 &\leq \frac{K}{M^4} \sum_{i=1}^N \sum_{j=1}^M | [y_N \mathbf{Q} \mathbf{H} \mathbf{H}^H \mathbf{Q} \mathbf{Y}]_{i,j} |^2 \\ &\leq \frac{K}{M^3} \text{tr} \left(\mathbf{Q} \mathbf{H} \mathbf{H}^H \mathbf{Q} \frac{\mathbf{Y} \mathbf{Y}^*}{M} \mathbf{Q} \mathbf{H} \mathbf{H}^H \mathbf{Q} \right) \\ &\leq \frac{K}{M^2}. \end{aligned}$$

To prove 2), we will resort to the resolvent identity which states:

$$\mathbf{Q}(a) - \mathbf{Q}(b) = (b - a) \mathbf{Q}(a) \mathbf{H} \mathbf{H}^H \mathbf{Q}(b). \quad (12)$$

Using (12), we obtain:

$$\begin{aligned} y_N &= \frac{1}{M} (y_N - \mathbb{E} y_N) \text{tr} \mathbf{H} \mathbf{H}^H \mathbf{Q}(y_N) \\ &+ \frac{1}{M} \text{tr} \mathbb{E}(y_N) \mathbf{H} \mathbf{H}^H \mathbf{Q}(y_N) + \frac{M - N}{M} \\ &= \frac{1}{M} (y_N - \mathbb{E} y_N) \mathbf{H} \mathbf{H}^H \mathbf{Q}(y_N) \\ &- \frac{1}{M} \text{tr} (y_N - \mathbb{E} y_N)^2 \mathbf{H} \mathbf{H}^H \mathbf{Q}(y_N) \mathbf{H} \mathbf{H}^H \mathbf{Q}(\mathbb{E} y_N) \\ &+ \frac{1}{M} \text{tr} \mathbb{E}(y_N) \mathbf{H} \mathbf{H}^H \mathbf{Q}(\mathbb{E} y_N) \\ &- \frac{1}{M} \text{tr} \mathbb{E}(y_N) (y_N - \mathbb{E}(y_N)) \mathbf{H} \mathbf{H}^H \mathbf{Q}(y_N) \mathbf{H} \mathbf{H}^H \mathbf{Q}(\mathbb{E}(y_N)) \\ &+ \frac{M - N}{M} \\ &\stackrel{(a)}{=} \frac{1}{M} (y_N - \mathbb{E} y_N) \text{tr} \mathbf{H} \mathbf{H}^H \mathbf{T}(\mathbb{E}(y_N)) \\ &+ \frac{1}{M} \mathbb{E}(y_N) \mathbf{H} \mathbf{H}^H \mathbf{T}(\mathbb{E}(y_N)) \\ &- \mathbb{E}(y_N) (y_N - \mathbb{E} y_N) \mathbb{E} \left[\frac{1}{M} \text{tr} \mathbf{H} \mathbf{H}^H \mathbf{Q}(y_N) \mathbf{H} \mathbf{H}^H \mathbf{Q}(\mathbb{E}(y_N)) \right] \\ &+ \frac{M - N}{M} + \epsilon. \end{aligned}$$

where ϵ verifies $\mathbb{E}(\epsilon) = \mathcal{O}(M^{-2})$. Note that equality (a) follows from the fact that $\text{var}(y_N)^2 = \mathcal{O}(M^{-2})$ and $\text{var}(\frac{1}{M} \text{tr} \mathbf{H} \mathbf{H}^H \mathbf{Q}(y_N) \mathbf{H} \mathbf{H}^H \mathbf{Q}(\mathbb{E}(y_N))) = \mathcal{O}(M^{-2})$. Therefore:

$$\begin{aligned} \mathbb{E}(y_N) &= \frac{1}{M} \mathbb{E}(y_N) \text{tr} \mathbf{H} \mathbf{H}^H \mathbf{T}(\mathbb{E}(y_N)) + \frac{M - N}{M} + \mathcal{O}(M^{-2}) \\ &= 1 - \frac{1}{M(1 + \kappa(\mathbb{E}(y_N)))} \text{tr}((\mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I}_N) \mathbf{T}(\mathbb{E}(y_N))) \\ &+ \mathcal{O}(M^{-2}) \\ &= 1 - \frac{\kappa(\mathbb{E}(y_N))}{1 + \kappa(\mathbb{E}(y_N))} + \mathcal{O}(M^{-2}) \\ &= \frac{1}{1 + \kappa(\mathbb{E}(y_N))} + \mathcal{O}(M^{-2}). \end{aligned}$$

Since y_N^* is the unique solution satisfying:

$$y_N^* = \frac{1}{1 + \kappa(y_N^*)},$$

we obtain:

$$\mathbb{E}(y_N) = y_N^* + \mathcal{O}(M^{-2}).$$

APPENDIX B PROOF OF THEOREM 2

The proof of theorem 2 relies on the tools used in [6], suitable for dealing with Gaussian random variables. Using an appropriate changing variable, one can prove :

$$\begin{aligned} C_L(y) &= \log \det(\mathbf{I}_r + \frac{\mathbf{D}_r^{\frac{1}{2}} \mathbf{W} \mathbf{W}^H \mathbf{D}_r^{\frac{1}{2}}}{M}) - \log \det(\frac{\mathbf{W} \mathbf{W}^H}{M}) \\ &- \log \det(\mathbf{D}_r). \end{aligned}$$

where r is the rank of \mathbf{H} , \mathbf{W} is a $r \times M - N + r$ standard Gaussian matrix and \mathbf{D}_r is a diagonal matrix depending on \mathbf{G}_k , \mathbf{H} and y . We retrieve thus the same model as in [6], with the slight difference that $C_L(y)$ has an extra random term equal to $\log \det(\frac{\mathbf{W} \mathbf{W}^H}{M})$. This has no effect on the applicability of the method, and one can get the desired result by following the same lines of [6].

REFERENCES

- [1] Ø. Ryan and M. Debbah, "Channel Capacity Estimation using Free Probability Theory," *IEEE Trans. Signal Process.*, vol. 56, no. 11, Nov. 2008.
- [2] P. Vallet and P. Loubaton, "A G-estimator of the MIMO channel ergodic capacity," in *Proc. IEEE International Symposium on Information Theory (ISIT'09)*, Seoul, Korea, Jun. 2009, pp. 774–778.
- [3] A. Kammoun, R. Couillet, J. Najim, and M. Debbah, "Performance of Capacity Inference Methods under Colored Interference," 2011. [Online]. Available: <http://arxiv.org/abs/1105.5305>
- [4] W. Hachem, P. Loubaton, and J. Najim, "Deterministic Equivalents for Certain Functionals of Large Random Matrices," *Annals of Applied Probability*, vol. 17, no. 3, pp. 875–930, 2007.
- [5] Z. D. Bai and J. W. Silverstein, "CLT of linear spectral statistics of large dimensional sample covariance matrices," *Annals of Probability*, vol. 32, no. 1A, pp. 553–605, 2004.
- [6] W. Hachem, O. Khorunzhy, P. Loubaton, J. Najim, and L. A. Pastur, "A new approach for mutual information analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, 2008.