Performance of capacity inference methods under colored interference

Abla Kammoun[1], Romain Couillet[2], Jamal Najim[1] and Mérouane Debbah[2]
[1]*Telecom ParisTech,* [2]*Alcatel-Lucent Chair, Supélec.*

Abstract

In this paper, we address the problem of fast point-to-point channel capacity estimation in the situation where the receiver undergoes unknown colored interference from multiple sources, whereas the channel with the transmitter is perfectly known. We consider the scenario where the number of observations is not sufficient to guarantee high performance of traditional estimators. Using estimation techniques associated to large random matrix theory, we derive an estimator referred to as the G-estimator and compare its performance against the conventional estimator. In particular, we prove that, unlike the conventional estimator, the G-estimator is consistent in the large dimensional setting, its variance going to zero as both space and time dimensions increase simultaneously. We finally complete the analysis by describing its fluctuations: When properly centered and rescaled, the G-estimator satisfies a central limit theorem, hence has Gaussian fluctuations. Simulations are provided which clearly show that the G-estimator outperforms the conventional one; simulations also strongly support the theoretical results even for small system dimensions.

I. Introduction

The use of multiple-input-multiple-output (MIMO) technologies has the potential to achieve high data rates, since several independent channels between the transmitter and the receiver can be exploited. However, the effectiveness of this technology may depend on the conditions of the surrounding environment such as the availability of the channel state information or the presence of colored interference. From a practical point of view, in a fast varying fading channel, it is of fundamental importance for users to rapidly estimate the maximum rate that can be achieved in the communication to other users.

Conventional methods for channel capacity estimation rely on the use of classical estimation techniques which assume a large number of observations. In general, consider θ the parameter we wish to estimate, and M the number of independent and identically distributed observation vectors $\mathbf{y}_1, \cdots, \mathbf{y}_M \in \mathbb{C}^N$. The parameter θ is often a function of the covariance matrix $\mathbf{\Sigma} = \mathbb{E}\left[\mathbf{y}_1\mathbf{y}_1^{\mathsf{H}}\right]$ of the received random process, i.e $\theta = f(\mathbf{\Sigma})$, for some function f. Using the strong law of large numbers, a consistent estimate of the covariance of the random process is simply given by the empirical covariance of $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_M]$, i.e. $\widehat{\mathbf{\Sigma}} \triangleq \frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{y}_i\mathbf{y}_i^{\mathsf{H}}$. Classical estimation methods then consist in using the empirical covariance as a good approximation of $\mathbf{\Sigma}$, thus yielding the estimator $\widehat{\theta}$ of θ, where $\widehat{\theta} = f(\widehat{\mathbf{\Sigma}})$. Such methods provide good performance as long as the number of observations M is very large compared to the vector size N, a situation rarely encountered in wireless communications, especially in fast changing environments.

To address the scenario where the number of observations M is of the same order as the dimension N of each observation, new consistent estimation methods based on large random matrix theory have been proposed in the context of wireless communications. They were initially applied to eigenvector and eigenvalue estimation problems [1], which has given rise to improved subspace estimation techniques [2], [3]. Recently, the use of these methods to estimate performance indexes has spurred the interest of many researchers. In the field of wireless communications, the capacity estimation of MIMO systems under imperfect channel knowledge has been addressed in [4] and [5], where methods based respectively on free probability theory and large random matrix theory have been proposed.

In this paper, we consider a different situation where the receiver perfectly knows the channel with the transmitter but does not a priori know 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. The transmitter-receiver pair then elects the bands achieving the highest rates, for which the exact interference is then inferred for proper transmission at the estimated rate. This approach is much more accurate than the approach consisting only in evaluating the total noise variance in each band and much faster than the approach consisting in evaluating the exact interference matrix for each band.

We specifically derive first a consistent estimator of the ergodic capacity in the case where the channel from the transmitter to the receiver is assumed to be known. In a second step, we study the asymptotic performance of the proposed estimator and compare it with that of the traditional one. In particular, we prove that both estimators converge to Gaussian random variables and identify their theoretical variances.

Notations: In the following, boldface lower case symbols represent vectors, capital boldface characters denote matrices ($\mathbf{I}_N$ is the size-N identity matrix). If $\mathbf{A}$ is a given matrix, $\mathbf{A}^{\mathsf{H}}$ stands for its transconjugate; if $\mathbf{A}$ is square, $\mathrm{tr}(\mathbf{A})$, $\det(\mathbf{A})$ and $\|\mathbf{A}\|$ respectively stand for the trace, the determinant and the spectral norm of $\mathbf{A}$. We say that the variable X has a standard complex Gaussian distribution if $X = U + \mathbf{i}V$ ($\mathbf{i}^2 = -1$) , where U, V are independent real random variables with Gaussian distribution $\mathcal{N}(0, 2^{-1})$. Almost sure convergence will be denoted by $\xrightarrow{\text{a.s.}}$, and convergence in distribution by $\xrightarrow{\mathcal{D}}$. Notation $\mathcal{O}$ will refer to Landau's notation: $u_n = \mathcal{O}(v_n)$ if there exists a bounded sequence K_n such that $u_n = K_n v_n$.

Paper organization: In Section II, we present the system model and formalize mathematically the considered problem. In Section III, we provide first order results for the conventional and the proposed estimator. We show that while the proposed one is consistent with growing N, M, the traditional estimator is asymptotically biased. In Section IV, we study the fluctuations of both estimators: we establish central limit theorems (CLT), hence we

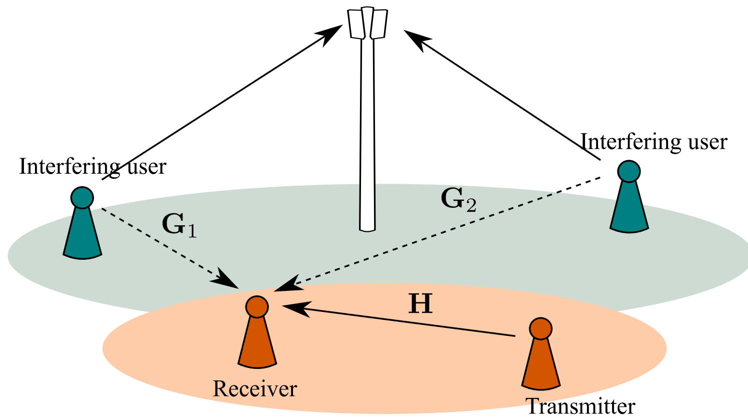

Fig. 1. System model.

prove the Gaussianity of the fluctuations, and we derive the asymptotic variances. Finally, we provide in Section V numerical simulations that support the accuracy of the derived results. Mathematical details are provided in the appendices.

II. System model and problem setting

The system model

Consider a communication link between two users: a transmitter and a receiver equipped with n_0 and N antennas, respectively. Also assume that the communication link is affected by the presence of K interferers with n_k antennas each, $1 \leq k \leq K$. Figure 1 describes this scenario, in the case of two interfering users. Similar to [5], we assume that time is slotted. We denote T the number of time slots and assume that the channel matrices are deterministic and remain constant in every time slot $t \in \{1, \cdots, T\}$. In other words, we assume that within each slot t, the $N \times n_0$ channel matrix $\mathbf{H}_t$ representing the channel between the transmitter and the receiver, and the $N \times n_k$ channel matrix $\mathbf{G}_{t,k}$ standing for the channel between the transmitter and the k-th interferer are deterministic and constant. Denote by M the data transmission periods in each slot. The M concatenated signal vectors received in slot t are gathered in $\overline{\mathbf{Y}}_t \in \mathbb{C}^{N \times M}$ given by:

$$\overline{\mathbf{Y}}_t = \mathbf{H}_t \mathbf{X}_{t,0} + \sum_{k=1}^{K} \mathbf{G}_{t,k} \mathbf{X}_{t,k} + \sigma \mathbf{W}_t,$$

where $\mathbf{X}_{t,0} \in \mathbb{C}^{n \times M}$ is the concatenated matrix of the transmitted signals, $\mathbf{X}_{t,k} \in \mathbb{C}^{n_k \times M}$ represents the interfering signal and $\mathbf{W}_t \in \mathbb{C}^{N \times M}$ stands for the additive noise. Their formal statistical properties are given in the following assumption:

Assumption A1*:* For given t and k where $1 \leq t \leq T$ and $1 \leq k \leq K$, the entries of the matrices $\mathbf{X}_{t,0}$, $\mathbf{X}_{t,k}$ and $\mathbf{W}_t$ are random variables, independent and identically distributed (i.i.d.) with standard complex Gaussian distribution and independent across t, k.

Assuming a perfect decoding of $\mathbf{X}_{t,0}$, initially transmitted at low rate, and a perfect knowledge of the channel matrix $\mathbf{H}_t$, the residual interference to which the receiver has access is given by:

$$\mathbf{Y}_t = \overline{\mathbf{Y}}_t - \mathbf{H}_t \mathbf{X}_{t,0} = \sum_{k=1}^{K} \mathbf{G}_{t,k} \mathbf{X}_{t,k} + \sigma \mathbf{W}_t.$$

This is also the received signal at slot t if no transmissions occurred.

The receiver wants to evaluate the average rate that can be achieved during the T slots, or equivalently by approximating the ergodic capacity (per transmit antenna). Under Assumption **A1**, an approximate of the ergodic capacity is given by:

$$\begin{aligned} C_{\mathrm{erg}} &= \frac{1}{NT} \sum_{t=1}^{T} \left[\log\det\left(\sigma^2 \mathbf{I}_N + \sum_{k=1}^{K} \mathbf{G}_{t,k} \mathbf{G}_{t,k}^{\mathsf{H}} + \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \right) - \log\det\left(\sigma^2 \mathbf{I}_N + \sum_{k=1}^{K} \mathbf{G}_{t,k} \mathbf{G}_{t,k}^{\mathsf{H}} \right) \right] \\ &= \frac{1}{NT} \sum_{t=1}^{T} \left[\log\det\left(\sigma^2 \mathbf{I}_N + \mathbf{G}_t \mathbf{G}_t^{\mathsf{H}} + \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \right) - \log\det\left(\sigma^2 \mathbf{I}_N + \mathbf{G}_t \mathbf{G}_t^{\mathsf{H}} \right) \right] \end{aligned} \tag{1}$$

where

$$\mathbf{G}_t = [\mathbf{G}_{t,1}, \cdots, \mathbf{G}_{t,K}] \in \mathbb{C}^{n \times N} \tag{2}$$

with $n = \sum_{k=1}^{K} n_k$.

In this paper, we address the problem of estimating C_{erg} based on the T successive observations $\mathbf{Y}_1, \ldots, \mathbf{Y}_T$ assuming perfect knowledge of $\mathbf{H}_1, \cdots, \mathbf{H}_T$.

The conventional large-M estimator $\hat{C}_{\mathrm{trad}}$

If the number M of available observations in each slot is very large compared to the channel vector N, the standard estimator $\hat{C}_{\mathrm{trad}}$, hereafter referred to as the large-M estimator, reads:

$$\hat{C}_{\mathrm{trad}} = \frac{1}{NT} \sum_{t=1}^{T} \log\det\left(\frac{1}{M} \mathbf{Y}_t \mathbf{Y}_t^{\mathsf{H}} + \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \right) - \frac{1}{NT} \sum_{t=1}^{T} \log\det\left(\frac{1}{M} \mathbf{Y}_t \mathbf{Y}_t^{\mathsf{H}} \right) . \tag{3}$$

However, in practice, the situation $M \gg N$ is rarely encountered, especially in systems embedded with multiple antennas and under fast fading channel conditions implying that M is of the same order of magnitude as N.

In this case, it can be proved that the large-M estimator is asymptotically biased, hence *not* consistent. The objective of this work is to propose a consistent estimator of C_{erg} when the number of available observations is of the same order (although larger) than N. We will refer to this estimator as the G-estimator in reference to Girko who introduced many estimators [6], [7] in similar contexts and coined these techniques as G-estimation techniques (standing for *general* estimation techniques).

It will be convenient in the sequel to consider the following notation:

$$\hat{C}_{\mathrm{trad}}(y) = \frac{1}{NT} \sum_{t=1}^{T} \log\det\left(\frac{1}{M} \mathbf{Y}_t \mathbf{Y}_t^{\mathsf{H}} + y\, \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \right) - \frac{1}{NT} \sum_{t=1}^{T} \log\det\left(\frac{1}{M} \mathbf{Y}_t \mathbf{Y}_t^{\mathsf{H}} \right) . \tag{4}$$

With this notation at hand, $\hat{C}_{\mathrm{trad}} = \hat{C}_{\mathrm{trad}}(1)$.

The asymptotic regime, remaining assumptions

Recall that $n = \sum_{k=1}^{K} n_k$. The derivation of the G-estimator will be carried out under the following assumptions:

Assumption A2: $M, N, n, n_0 \to +\infty$, and:

$$0 \; < \; \liminf_{M,N\to\infty} \frac{N}{n} \quad \le \quad \limsup_{M,N\to\infty} \frac{N}{n} \; < \; +\infty \ ,$$

$$1 \; < \; \liminf_{M,N\to\infty} \frac{M}{N} \quad \le \quad \limsup_{M,N\to\infty} \frac{M}{N} \; < \; +\infty \ ,$$

$$0 \; < \; \liminf_{N,n_0\to\infty} \frac{n_0}{N} \quad \le \quad \limsup_{N,n_0\to\infty} \frac{n_0}{N} \; < \; +\infty \ .$$

Remark 1: The constraints over N and n simply state that these quantities remain of the same order. The lower bound for the ratio M/N accounts for the fact that that M is larger than N, although of the same order.

In the rest of the paper, this regime will simply be referred to as $M, N, n \to \infty$. We are now in position to formalize the assumptions over the channel matrices:

Assumption A3: Let $t \in \{1, \cdots, T\}$ (T fixed). Consider the family $(\mathbf{G}_t)$ of $N \times n$ matrices and the family $(\mathbf{H}_t)$ of $N \times n_0$ matrices where N, n, n_0 satisfy Assumption **A2**. Then the spectral norms of $\mathbf{G}_t$ and $\mathbf{H}_t$ are uniformly bounded in the sense that:

$$\sup_{1\le t\le T} \sup_{N,n} \|\mathbf{G}_t\| < \infty \ , \qquad \sup_{1\le t\le T} \sup_{N,n_0} \|\mathbf{H}_t\| < \infty \ .$$

Assumption A4: Denote by r_t the rank of $\mathbf{H}_t$. Then

$$0 \; < \; \liminf_{N,n_0\to\infty} \frac{r_t}{N} \; \le \; \limsup_{N,n_0\to\infty} \frac{r_t}{N} \; < \; 1 \ .$$

III. Convergence of the capacity estimators

In this section, we study the asymptotic behaviour of the large-M estimator $\hat{C}_{\mathrm{trad}}$ and prove that under the asymptotic regime **A2**, this estimator is biased. We then build a consistent estimator based on G-estimation techniques. Both results are essentially based on large random matrix theory. Let us first briefly introduce the G-estimation techniques. G-estimation techniques can be roughly classified into two categories. The first one is based on the Stieltjes transform (the definition of which is recalled below) and was taken up by Mestre who developed a framework for eigenvalue and eigenvector estimation issues [1].

Let $\mathbb{P}$ be a probability distribution on $\mathbb{R}^+$, then the Stieltjes transform $m(z)$ of $\mathbb{P}$ is defined as

$$m(z) = \int_{\mathbb{R}} \frac{\mathbb{P}(d\lambda)}{\lambda - z} \ , \quad z \in \mathbb{C} \setminus \mathbb{R}^+ \ . \tag{5}$$

For example, the Stieltjes transform $m_{\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}}$ associated to the empirical distribution of the eigenvalues of the Hermitian matrix $\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}$ is simply the normalized trace of the associated resolvent:

$$m_{\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}}(z) = \frac{1}{N} \operatorname{tr}\left(\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}} - z\mathbf{I}_N\right)^{-1} = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\lambda_i - z} \ ,$$

where $\lambda_1, \cdots, \lambda_N$ denotes the eigenvalues of $\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}$. Since their introduction by Marčenko and Pastur in their seminal paper [8], Stieltjes transforms have proved to be a highly efficient tool to study the spectrum of large random matrices. From an estimation point of view, Stieltjes transform are, in the large dimension regime of interest, consistent estimates of well-identified deterministic quantities. Therefore, the approach consists in expressing the parameters of interest as functions of the Stieltjes transform of the eigenvalue distribution of $\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}$. This approach is appropriate as long as we consider estimation of parameters depending either on the eigenvalues or on the eigenvectors of $\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}$, but cannot be used when the dependence is on both of them; it will be illustrated in Lemma 2 below.

The second approach is based on other consistent estimators different from the Stieltjes transform $m_{\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}}(z)$. Details will be provided in Section III-B.

A. The large-M estimator is biased

Recall the definition of the large-M estimator $\hat{C}_{\mathrm{trad}}$ given in (3). Before providing the expression of the asymptotic bias for $\hat{C}_{\mathrm{trad}}$, we shall define some deterministic quantities and also study their properties under the appropriate asymptotic regime $M, N, n \to \infty$.

Lemma 1: Let Assumptions **A1-A4** hold true. Denote $\mathbf{\Gamma}_t = \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N$ and let $y > 0$. Then:

1) The functional equation:
$$\kappa_t(y) = \frac{1}{M}\,\mathrm{tr}\left(\mathbf{\Gamma}_t\left(\frac{\mathbf{\Gamma}_t}{1+\kappa_t(y)} + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\right)^{-1}\right) \tag{6}$$
admits a unique positive solution $\kappa_t(y)$.

Denote by $\mathbf{T}_t(y)$ and $\mathbf{Q}_t(y)$ the following quantities:
$$\mathbf{T}_t(y) = \left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{\Gamma}_t}{1+\kappa_t(y)}\right)^{-1}, \quad \mathbf{Q}_t(y) = \left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right)^{-1}.$$

2) Then, for any deterministic family $(\mathbf{S}_N)$ of $N \times N$ complex matrices with uniformly bounded spectral norm, we have:
$$\frac{1}{M}\,\mathrm{tr}\,\mathbf{S}_N\mathbf{Q}_t(y) - \frac{1}{M}\,\mathrm{tr}\,\mathbf{S}_N\mathbf{T}_t(y) \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0.$$

3) Let
$$V_t(y) = -\log\det(\mathbf{T}_t(y)) + M\log(1+\kappa_t(y)) - M\frac{\kappa_t(y)}{1+\kappa_t(y)},$$
then, the following convergence holds true:
$$-\frac{1}{N}\log\det\mathbf{Q}_t(y) - \frac{1}{N}V_t(y) \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0\,.$$

Proof of Lemma 1 is postponed to Appendix A.

Remark 2: Note that items 2) and 3) provide deterministic equivalents of various random quantities under the asymptotic regime of interest.

In the next lemma, we show how the Stieltjes transform method can be used to compute a consistent estimate of $\frac{1}{N}\sum_{t=1}^{T}\log\det(\sigma^2\mathbf{I}_N + \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}})$. This term only depends on the eigenvalues of $\mathbf{G}_t$ which are not directly observable. The idea underlying G-estimation is to use advanced random matrix theory tools to link the asymptotic non-observable Stieltjes transform of $\mathbf{G}_t$ to that of the observable covariance matrix $\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}$. More precisely, we prove the following:

Lemma 2: Let Assumptions **A1**-**A4** hold true. Then, the following convergence holds true:

$$\frac{1}{N}\log\det(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N)-\frac{1}{N}\log\det(\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}})+\frac{N-M}{N}\log\left(\frac{M-N}{M}\right)-1\xrightarrow[M,N,n\to\infty]{\text{a.s.}}0\ .$$

Proof of Lemma 2 is postponed to Appendix B.

Remark 3: As a consequence of this lemma, it turns out that a consistent estimate of $\frac{1}{N}\log\det(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N)$ is simply the traditionnal large-M estimator (recall that $\frac{1}{M}\mathbb{E}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}=\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N$) up to a term of bias depending on the time and space dimensions.

We now derive the bias of the estimator $\hat{C}_{\text{trad}}$. Prior to that, define the deterministic quantity $\mathcal{V}(y)$ as :

$$\mathcal{V}(y)=-\frac{1}{NT}\sum_{t=1}^{T}\left(\log\det(\mathbf{T}_t(y))+M\log(1+\kappa_t(y))-\frac{M\kappa_t(y)}{1+\kappa_t(y)}-\log\det(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N)\right)$$
$$+\frac{M-N}{N}\log\left(\frac{M-N}{M}\right)+1. \quad (7)$$

where $\kappa_t(y)$ is the unique solution of (6).

Theorem 1 (Bias of the large-M estimator): Let Assumptions **A1**-**A4** hold true. Then,

$$\hat{C}_{\text{trad}}-\mathcal{V}(1)\xrightarrow[M,N,n\to\infty]{\text{a.s.}}0\ .$$

Proof: Gathering items 3) and 4) in Lemma 1 yields the desired result. ■

B. A G-estimator for the capacity

The term $\frac{1}{N}\log\det(\sigma^2\mathbf{I}_N+\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}})$ in the definition of the capacity depends on the eigenvalues of $\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}$. Since matrix $\mathbf{H}_t$ is assumed to be known and to not necessarily share the same eigenvector space as $\mathbf{G}_t$, the capacity depends simultaneously on the eigenvalues and the eigenvectors of the unobservable matrix $\mathbf{G}_t$. Hence, the use of the Stieltjes transform cannot be applied. A similar situation was successfully addressed in [5], by using a novel approach based on deterministic equivalents as developed in [9]. In the sequel, we follow the same approach in [5].

Theorem 2 (a G-estimator for the capacity): Assume that **A1** and **A3** hold true; consider the quantity:

$$\hat{C}_G\ =\ \frac{1}{NT}\sum_{t=1}^{T}\log\det\left(\mathbf{I}_N+\hat{y}_{N,t}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\left(\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right)^{-1}\right)$$
$$+\frac{(M-N)}{N}\left[\log\left(\frac{M}{M-N}\hat{y}_{N,t}\right)+1\right]-\frac{M}{N}\hat{y}_{N,t}\ ,$$

where $\hat{y}_{N,t}$ is the unique real positive solution of the following equation:

$$\hat{y}_{N,t}=\frac{\hat{y}_{N,t}}{M}\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\left(\hat{y}_{N,t}\,\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}+\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right)^{-1}+\frac{M-N}{M}.$$

Then,

$$\hat{C}_G - C_{\rm erg} \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 \ .$$

In the sequel, we will refer to $\hat{C}_G$ as the G-estimator.

Remark 4: Note that $\hat{C}_G$ writes:

$$\hat{C}_G = \hat{C}_{\rm trad}(\hat{y}_{N,t}) + \frac{(M-N)}{N}\left[\log\left(\frac{M}{M-N}\hat{y}_{N,t}\right)+1\right] - \frac{M}{N}\hat{y}_{N,t} \ ,$$

a relation that sheds some light on the difference between $\hat{C}_G$ and $\hat{C}_{\rm trad}$.

In order to prove Theorem 2, it is sufficient to provide a consistent estimate of each quantity in the sum of the expression of the ergodic capacity. Denote by C_t the capacity at time t given by:

$$\begin{aligned} C_t &\triangleq \frac{1}{N}\log\det(\sigma^2\mathbf{I}_N + \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}) - \frac{1}{N}\log\det(\sigma^2\mathbf{I}_N + \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}) \ , \\ &\triangleq C_{t,1} - C_{t,2} \ . \end{aligned}$$

As a consistent estimate $\hat{C}_{t,2}$ of $C_{t,2}$ has already been provided by Lemma 2, it remains to build a consistent estimate for $C_{t,1}$.

The proof of Theorem 2 is postponed to Appendix C. Although technical, this proof is very illustrative on how to build consistent estimators based on deterministic equivalents. We therefore provide below an outline of the proof.

Outline of the proof: The proof is divided into 4 steps:

1) In the first step, we exploit the convergence of parametrized quantities of interest. Denote $f(y) = \frac{1}{N}\log\det(\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}} + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}})$ and recall the definition of $\kappa_t(y)$ as given in Lemma 1-1). By Lemma 1-3), we have:

$$-f(y) + \frac{1}{N}\log\det\left(\frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N}{1+\kappa_t(y)} + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\right) + \frac{M}{N}\log(1+\kappa_t(y)) - \frac{M}{N}\frac{\kappa_t(y)}{1+\kappa_t(y)} \xrightarrow[M,N,n\to\infty]{\text{a.s}} 0 \ .$$

   Clearly, the deterministic quantity to which $f(y)$ converges differs from $C_{t,1}$.

2) In the second step, we find a specific value of y to enforce the desired quantity $C_{t,1}$ to appear: one can readily check that if $y_{N,t}$ is the solution of the following equation:

$$y = \frac{1}{1+\kappa_t(y)} \ , \tag{8}$$

   then one would immediately obtain:

$$C_{t,1} - \left[\frac{1}{N}\log\det\left(\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}} + y_{N,t}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\right) + \frac{M-N}{N}\log(y_{N,t}) + \frac{M}{N}(1-y_{N,t})\right] \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 \ . \tag{9}$$

   Based on the definition of $\kappa_t(y)$, one can prove that there exists a unique positive $y_{N,t}$ solution of (8), given by the following closed-form expression:

$$y_{N,t} = 1 - \frac{1}{M}\,\mathrm{tr}\left[(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)(\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)^{-1}\right] \ . \tag{10}$$

   Unfortunately, the value of $y_{N,t}$ depends upon the unknown matrix $\mathbf{G}_t$.

3) In the third step, we provide a consistent estimator $\hat{y}_{N,t}$ of $y_{N,t}$. Based on an analysis of $\kappa_t(y)$, and on finding a consistent estimate for this quantity, one can prove that there exists a unique positive solution $\hat{y}_{N,t}$ to the following equation:

$$\hat{y}_{N,t} = \frac{1}{M} \operatorname{tr} \hat{y}_{N,t} \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \left(\hat{y}_{N,t} \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} + \frac{1}{M} \mathbf{Y}_t \mathbf{Y}_t^{\mathsf{H}} \right)^{-1} + \frac{M-N}{M}. \tag{11}$$

Moreover, $\hat{y}_{N,t}$ satisfies:

$$\hat{y}_{N,t} - y_{N,t} \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 \ .$$

4) Finally, it remains to check that one can replace $y_{N,t}$ by $\hat{y}_{N,t}$ in the convergence (9). This will immediately yield a consistent estimate $\hat{C}_{t,1}$ for $C_{t,1}$. For the proof of the theorem to be complete, it remains to gather the estimates of $C_{t,1}$ and $C_{t,2}$. This yields :

$$\hat{C}_G \ = \ \frac{1}{T} \sum_{t=1}^{T} \left(\hat{C}_{t,1} - \hat{C}_{t,2} \right) \ ,$$

which is the announced result.

■

IV. Fluctuations for the capacity estimators

We develop in this section fluctuation results for the capacity estimators $\hat{C}_{\text{trad}}$ and $\hat{C}_{\text{G}}$ already introduced. More precisely, we establish CLTs, provide explicit expressions for the variance, and prove that these estimators when correctly centered and rescaled converge in distribution toward a Gaussian random variable.

While the entries of the matrices $\mathbf{X}_t$ and $\mathbf{W}_t$ (cf. Assumption **A1**) could have easily been taken non Gaussian to establish first order results in Section III, the Gaussian property of the entries is a central assumption to establish fluctuation results. This assumption is natural in the current wireless communications context.

The Gaussianity of the entries allows one to use the powerful Gaussian methods adapted along the years to the study of large random matrices by Pastur and co-authors (see e.g. [10] - for application to wireless communication, see [11], etc.). The Gaussian calculus heavily relies (but not exclusively) on the integration by parts formula and the Poincaré-Nash inequality, recalled in Appendix D.

A. Fluctuations of the large-M estimator

In the previous section, we have shown that the large-M estimator is asymptotically biased, in the sense that it converges to a deterministic equivalent which is different from the theoretical ergodic capacity.

In the sequel, we shall study its fluctuations around this deterministic equivalent. We will prove that when properly centered and rescaled, the large-M estimator converges to a standard Gaussian random variable.

This result is an important first step to the study of the fluctuations of the G-estimator.

Theorem 3: Let Assumptions **A1**-**A4** hold true and recall the definition (4) of $\hat{C}_{\text{trad}}(y)$. Then,

1) the sequence of real numbers $(\alpha_N(y))$:

$$\alpha_N(y) = \frac{2\log(M)}{T^2} - \frac{1}{T^2}\sum_{t=1}^{T}\log\left((M-N)\left(M(\kappa_t(y)+1\right)^2 - \mathrm{tr}\left(\frac{\mathbf{I}_N}{\kappa_t(y)+1} + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)^{-1}\right)^{-2}\right)$$

is well-defined. Furthermore:

$$0 \ < \liminf_{M,N,n\to\infty} \alpha_N(y) \ \leq \ \limsup_{M,N,n\to\infty} \alpha_N(y) \ < \ +\infty \ .$$

2) The following convergence holds true:

$$\frac{N}{\alpha_N(y)}\left(\hat{C}_{\mathrm{trad}}(y) - \mathcal{V}(y)\right) \xrightarrow[N,M,n\to\infty]{\mathcal{D}} \mathcal{N}(0,1) \ ,$$

where $\mathcal{V}(y)$ is defined in (7).

Proof: See Appendix D. ■

B. Fluctuations of the G-estimator

As opposed to the large-M estimator, the G-estimator has no closed-form expression, as the $\hat{y}_{N,t}$'s are solutions of implicit equations (easily solved through numerical computations, though). Establishing the CLT might seem more difficult since the randomness comes from both the received matrix $\mathbf{Y}_t$ and the quantity $\hat{y}_{N,t}$.

In the following lemma, we shall prove that the fluctuations of $\hat{y}_{N,t} - y_{N,t}$ are of order $\mathcal{O}(M^{-2})$, a rate which is sufficient, as we will see later, to discard the randomness stemming from $\hat{y}_{N,t}$ in the study of the fluctuations.

Lemma 3: For $t \in \{1, \cdots, T\}$, the following estimates hold true, as $M, N, n \to \infty$:

1) $\mathrm{var}(\hat{y}_{N,t}) = \mathcal{O}(M^{-2})$,
2) $\mathbb{E}\,\hat{y}_{N,t} = y_{N,t} + \mathcal{O}(M^{-2})$.

Proof: See Appendix E. ■

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

Theorem 4: Let Assumptions **A1**-**A3** hold true. Then,

$$\frac{N}{\theta_N}(\hat{C}_G - C_{\mathrm{erg}}) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0,1),$$

where θ_N given by:

$$\theta_N = \frac{1}{T^2}\sum_{t=1}^{T} 2\log(My_{N,t}) - \log\left((M-N)\left(M - \mathrm{tr}\left(\left(\mathbf{I}_N + \mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)^{-1}\right)^{-2}\right)\right)\right) \qquad (12)$$

is well-defined and satisfies

$$0 \ < \liminf_{M,N,n\to\infty} \theta_N \ \leq \ \limsup_{M,N,n\to\infty} \theta_N \ < \ +\infty \ .$$

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

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

Then $\hat{C} = \frac{1}{T}\sum_{t=1}^{T} \mathcal{C}_t(\hat{y}_{N,t})$. Since all the random variables $(\mathcal{C}_t(\hat{y}_{N,t}),\ 1 \le t \le T)$ are independent, it is sufficient to prove a CLT for $\mathcal{C}_t(\hat{y}_{N,t})$, for a given $t \in \{1, \cdots, T\}$. In order to handle the randomness of $\hat{y}_{N,t}$, we shall perform a Taylor expansion of $\mathcal{C}_t$ around $\hat{y}_{N,t}$. Recall the following differentiation formula:

$$\frac{d}{dx}\log\det A(x) = \operatorname{tr} A'(x)A^{-1}(x)$$

(see for instance [12, Section 15]). A direct application of this formula, together with the mere definition of $\hat{y}_{N,t}$ yields:

$$\frac{d\,\mathcal{C}_t}{d\,y}(\hat{y}_{N,t}) = 0\ .$$

Hence, the Taylor expansion writes:

$$N\mathcal{C}_t(y_{N,t}) = N\mathcal{C}_t(\hat{y}_{N,t}) + N\frac{(y_{N,t} - \hat{y}_{N,t})^2}{2} \times \frac{d^2\mathcal{C}_t}{dy^2}(\hat{y}_{N,t}) + N\,\frac{(y_{N,t} - \hat{y}_{N,t})^3}{6} \times \frac{d^3\mathcal{C}_t}{dy^3}(\xi_{N,t})\ , \tag{13}$$

where $\xi_{N,t}$ lies between $y_{N,t}$ and $\hat{y}_{N,t}$. The mere definition (11) of $\hat{y}_{N,t}$ yields:

$$\frac{M-N}{M} \le \hat{y}_{N,t} \le 1 + \frac{M-N}{M}\ .$$

In particular, $\hat{y}_{N,t}$ uniformly belongs to a fixed compact interval, so does $y_{N,t}$ for similar reasons. One can easily prove that the second and third derivatives of $\mathcal{C}_t(y)$ are uniformly bounded on the union of these intervals. This result combined with the fact that $N\mathbb{E}(\hat{y}_{N,t} - y_{N,t})^2 = \mathcal{O}(M^{-1})$ implies that the last two terms in the right hand side (r.h.s.) of (13) converge to zero in probability. By Slutsky's Theorem [13], it suffices to establish the CLT for $N\mathcal{C}(y_{N,t})$ instead of $N\mathcal{C}(\hat{y}_{N,t}) = N\hat{C}(\hat{y}_{N,t})$. This is extremely helpful since unlike $\hat{y}_{N,t}$ whih is random, $y_{N,t}$ is deterministic. The result is thus obtained by applying Theorem 3 and noticing that $\kappa(y_{N,t}) + 1 = \frac{1}{y_{N,t}}$. ∎

V. Simulations

In the simulations, we consider the case where a mobile terminal with $N = 4$ antennas receives during $M = 15$ slots, data stemming from an $n_0 = 4$ antenna secondary transmitter. We assume that the communication link is interfered by $K = 8$ mono-antenna users. For each $t \in \{1, \cdots, T\}$, matrices $\mathbf{H}_t$ and $\mathbf{G}_t$ are randomly chosen as standard Gaussian matrices and remain constant during the Monte Carlo averaging. In a first experiment we set T to 10 and represent in Fig. 2 the theoretical and empirical normalized variances for the G-estimator with respect to $\mathrm{SNR} = \frac{1}{\sigma^2}$. We also display in the same graph the empirical variance of the large-M estimator. We note that the G-estimator exhibits better performance for all SNR range. We study in a second experiment the effect of T when the SNR is set to 10 dB. Fig. 3 represents the obtained results. We note that since the large-M estimator is biased, its mean square error does not significantly decrease with T and remains almost unchanged, whereas the G-estimator exhibits a low variance which drops linearly with T. Finally, to assess the Gaussian behaviour of both estimators, we represent in Fig. 4 and Fig. 5 their corresponding histograms. We note a good fit between theoretical and empirical results although the system dimensions are small.

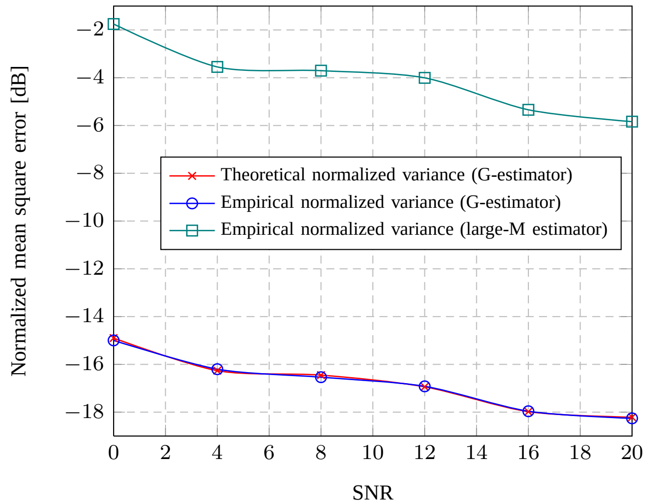

Fig. 2. Empirical and theoretical variances with respect to the SNR.

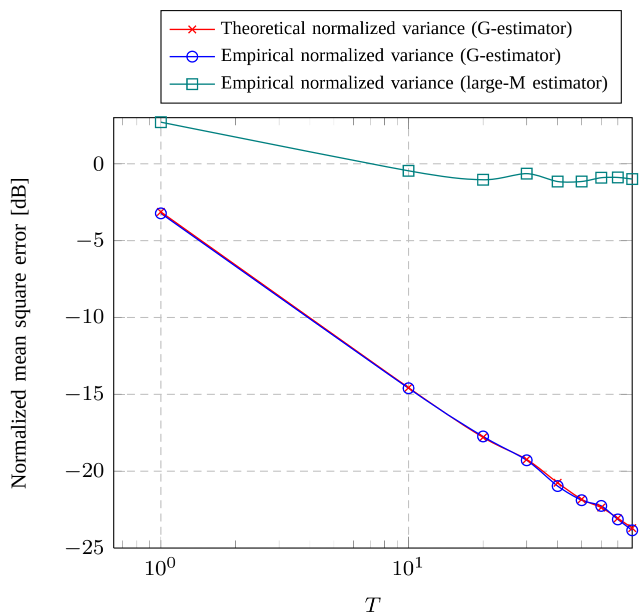

Fig. 3. Empirical and theoretical variances with respect to the SNR.

VI. Conclusion

In this paper, we have proposed a novel G-estimator for fast estimation of the ergodic capacity in presence of unknown interference in the case where the number of available observations is of the same order as the dimension of each observation. In particular, we have shown that the conventional estimator, based on the replacement of the unknown covariance matrix of the observations by the empirical covariance matrix, is biased. Based on large random matrix theory, we have introduced a novel G-estimator which is unbiased and consistent. We then have

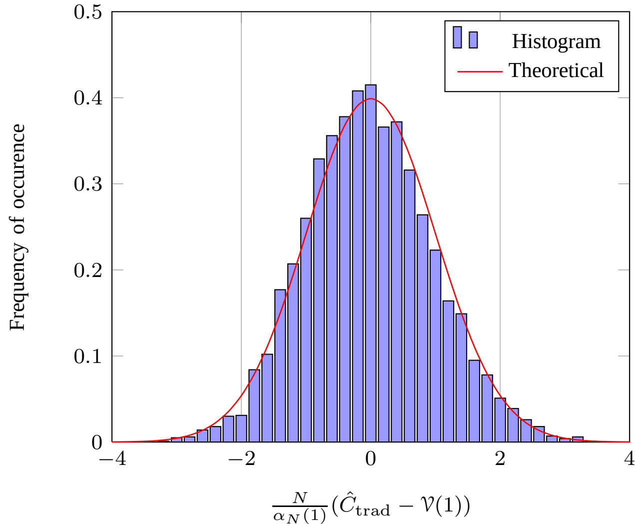

Fig. 4. Histogram of $\frac{N}{\alpha_N(1)}(\hat{C}_{\mathrm{trad}} - \mathcal{V}(1))$.

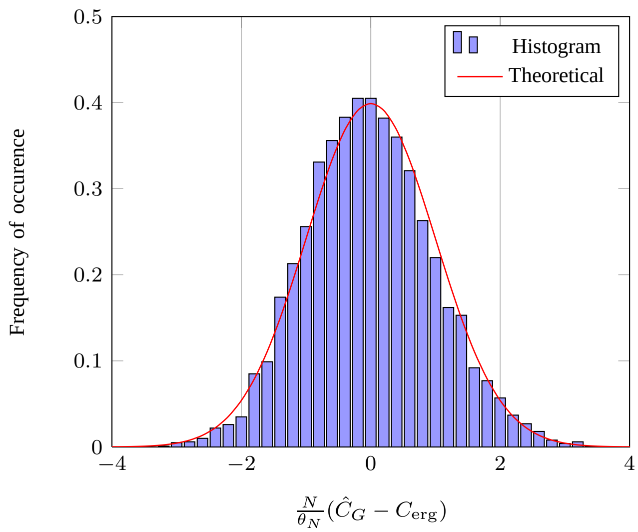

Fig. 5. Histogram of $\frac{N}{\theta_N}(\hat{C}_G - C_{\mathrm{erg}})$.

studied the fluctuations of the two estimators and established CLTs for both of them. Numerical simulations have been provided and strongly support the accuracy of our derived results even for usual system dimensions.

ACKNOWLEDGMENT

The authors would like to thank Jakob Hoydis for useful discussions.

APPENDIX A
PROOF OF LEMMA 1

Define for $\rho \geq 0$:

$$\mathbf{Q}_t(\rho, y) = \left(\rho \mathbf{I}_N + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right)^{-1} ,$$

$$g_t(\rho, y) = \frac{1}{N}\log\det\left(\rho \mathbf{I}_N + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right)^{-1} .$$

Denote by $\mathbf{X}_t = \left[\mathbf{X}_{t,1}^{\mathsf{H}}, \cdots, \mathbf{X}_{t,K}^{\mathsf{H}}\right]^{\mathsf{H}}$, and $\mathbf{Z}_t = \left[\mathbf{W}_t^{\mathsf{H}} \;\; \mathbf{X}_t^{\mathsf{H}}\right]^{\mathsf{H}}$ then $\mathbf{Y}_t = [\sigma\mathbf{I}_N \;\; \mathbf{G}_t]\,\mathbf{Z}_t$. Denote by $[\sigma\mathbf{I}_N \;\; \mathbf{G}_t] = \mathbf{U}_t\mathbf{\Sigma}_t\mathbf{V}_t^{\mathsf{H}}$ the singular value decomposition of $[\sigma\mathbf{I}_N \;\; \mathbf{G}_t]$ where $\mathbf{\Sigma}_t = \left[\mathbf{D}_t^{\frac{1}{2}} \; \mathbf{0}_{N\times n}\right]$, $\mathbf{D}_t$ being the diagonal matrix of eigenvalues of $\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N$; in particular, $\mathbf{D}_t$'s entries are nonnegative and bounded away from zero. Let $\widetilde{\mathbf{Z}}_t = \mathbf{V}_t^{\mathsf{H}}\left[\mathbf{W}_t^{\mathsf{H}} \;\; \mathbf{X}_t^{\mathsf{H}}\right]^{\mathsf{H}}$. Since the entries of $\mathbf{Z}_t$ are i.i.d. and Gaussian, $\widetilde{\mathbf{Z}}_t$ has the same entry distribution as $\mathbf{Z}_t$. Writing $\widetilde{\mathbf{Z}}_t = \left[\widetilde{\mathbf{W}}_t^{\mathsf{H}} \;\; \widetilde{\mathbf{X}}_t^{\mathsf{H}}\right]^{\mathsf{H}}$, $g_t(\rho, y)$ becomes:

$$\begin{aligned} g_t(\rho, y) &= \frac{1}{N}\log\det\left(\rho\mathbf{I}_N + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{1}{M}\mathbf{U}_t\mathbf{D}_t^{\frac{1}{2}}\widetilde{\mathbf{W}}_t\widetilde{\mathbf{W}}_t^{\mathsf{H}}\mathbf{D}_t^{\frac{1}{2}}\mathbf{U}_t^{\mathsf{H}}\right) , \\ &= \frac{1}{N}\log\det\left(\rho\mathbf{I}_N + y\mathbf{U}_t^{\mathsf{H}}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{U}_t + \frac{1}{M}\mathbf{D}_t^{\frac{1}{2}}\widetilde{\mathbf{W}}_t\widetilde{\mathbf{W}}_t^{\mathsf{H}}\mathbf{D}_t^{\frac{1}{2}}\right) . \end{aligned}$$

Obviously, we have $-\frac{1}{N}\log\det(\mathbf{Q}_t(y)) = g_t(0, y)$ and $\frac{1}{M}\operatorname{tr}\mathbf{Q}_t(y) = \frac{1}{M}\operatorname{tr}\mathbf{Q}_t(0, y)$. Deterministic equivalents for $g_t(\rho, y)$ and $\mathbf{Q}_t(\rho, y)$ have been derived in [9] and are recalled in the lemma below.

Lemma 4 (cf. [9]): Let $\rho > 0$.

1) Denote by $\mathbf{\Gamma}_t = \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N$ and let $y > 0$. The following functional equation:

$$\kappa_t(\rho, y) = \frac{1}{M}\operatorname{tr}\left(\mathbf{\Gamma}_t\left(\rho\mathbf{I}_N + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{\Gamma}_t}{1 + \kappa_t(\rho, y)}\right)^{-1}\right)$$

admits a unique positive solution $\kappa_t(\rho, y)$.

2) Define

$$\mathbf{T}_t(\rho, y) = \left(\rho\mathbf{I}_N + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{\Gamma}_t}{1 + \kappa_t(\rho, y)}\right)^{-1} .$$

Then, for any sequence of deterministic matrices $\mathbf{S}_N \in \mathbb{C}^{N\times N}$ with uniformly bounded spectral norm:

$$\frac{1}{M}\operatorname{tr}\mathbf{S}_N\mathbf{Q}_t(\rho, y) - \frac{1}{M}\operatorname{tr}\mathbf{S}_N\mathbf{T}_t(\rho, y) \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 .$$

In particular, setting $\mathbf{S}_N = \mathbf{\Gamma}_t$, we get:

$$\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\rho, y) - \kappa_t(\rho, y) \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 .$$

3) Let

$$V_t(\rho, y) = -\frac{1}{N}\log\det(\mathbf{T}_t(\rho, y)) + \frac{M}{N}\log(1 + \kappa_t(\rho, y)) - \frac{M}{N}\frac{\kappa_t(\rho, y)}{1 + \kappa_t(\rho, y)} ,$$

then

$$g(\rho, y) - V_t(\rho, y) \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 .$$

The general idea of the proof of Lemma 1 is to transfer these determinitic equivalents to the case $\rho \searrow 0$; we will proceed by taking advantage from from the fact that all the diagonal elements of $\mathbf{D}_t$ are positive and uniformly bounded away from zero.

We first prove the existence and uniqueness of $\kappa_t(y)$. Consider the function f defined on $[0,\infty[$ by:

$$f : x \mapsto x - \frac{1}{M}\operatorname{tr}\mathbf{D}_t\left(y\mathbf{U}_t^{\mathsf{H}}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{U}_t + \frac{\mathbf{D}_t}{1+x}\right)^{-1} .$$

An easy computation yields the derivative of f with respect to x:

$$f'(x) = 1 - \frac{1}{M}\operatorname{tr}\mathbf{D}_t\left(y\mathbf{U}_t^{\mathsf{H}}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{U}_t + \frac{\mathbf{D}_t}{1+x}\right)^{-1}\frac{\mathbf{D}_t}{(1+x)^2}\left(y\mathbf{U}_t^{\mathsf{H}}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{U}_t + \frac{\mathbf{D}_t}{1+x}\right)^{-1}$$

which is obviously always positive. Function f is thus always increasing and thus establishes a bijection from $[0,\infty[$ to $[f(0),+\infty[$. Since $f(0)$ is negative, we conclude that f has a single zero. This proves the existence and uniqueness of $\kappa_t(y)$. It remains to extend the asymptotic convergence results to the case $\rho = 0$.

In the sequel, we only prove item 2) for $\mathbf{S}_N = \mathbf{D}_N$ as it captures the key arguments of the proof; the extension to general sequences $(\mathbf{S}_N)$ will then be straightforward. Write $\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y) - \kappa_t(y)$ as:

$$\begin{aligned}\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y) - \kappa_t(y) &= \frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y) - \frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\epsilon,y)\\ &\quad + \frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\epsilon,y) - \kappa_t(\epsilon,y)\\ &\quad + \kappa_t(\epsilon,y) - \kappa_t(y) \ ,\end{aligned}$$

where $\epsilon > 0$. We now handle sequentially each of the differences of the r.h.s. of the previous decomposition. We first prove that there exists a fixed constant $K > 0$ (which only depends on $\limsup NM^{-1}$) such that for every $\epsilon > 0$, there exists N_1 (which depends on the realization and hence is random) such that for every $N \geq N_1$, we have:

$$\left|\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y) - \frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\epsilon,y)\right| \leq \frac{\epsilon}{K} \ . \tag{14}$$

This can be proved by noting that from the resolvent identity, we have:

$$\begin{aligned}\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y) - \frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\epsilon,y) &= \frac{\epsilon}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(0,y)\mathbf{Q}_t(\epsilon,y) \ ,\\ &\leq \frac{\epsilon}{M}\operatorname{tr}\mathbf{\Gamma}_t\left\|\left(\frac{1}{M}\mathbf{D}_t^{\frac{1}{2}}\widetilde{\mathbf{W}}\widetilde{\mathbf{W}}^{\mathsf{H}}\mathbf{D}_t^{\frac{1}{2}}\right)^{-1}\right\|^2 .\end{aligned}$$

Recall that $\widetilde{\mathbf{W}}_t$ is a $N \times M$ matrix and that by Assumption **A2**, $\limsup_{M,N} NM^{-1} < 1$. Therefore the spectrum of $\widetilde{\mathbf{W}}_t\widetilde{\mathbf{W}}_t^{\mathsf{H}}$ is almost surely eventually bounded away from zero[1]. In particular, there exists a constant K such that eventually, we have $\left\|\left(\frac{1}{M}\mathbf{D}_t^{\frac{1}{2}}\widetilde{\mathbf{W}}\widetilde{\mathbf{W}}^{\mathsf{H}}\mathbf{D}_t^{\frac{1}{2}}\right)^{-1}\right\|^2 \leq K^{-1}$, hence:

$$\exists N_1,\ \forall N \geq N_1, \quad \left|\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y) - \frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\epsilon,y)\right| \leq \frac{\epsilon}{K} \ .$$

[1]Recall that if $\lim NM^{-1} = c < 1$, then the smallest eigenvalue $\lambda_{\min}(\widetilde{\mathbf{W}}_t\widetilde{\mathbf{W}}_t^{\mathsf{H}})$ converges to $(1-\sqrt{c})^2 > 0$; it remains to argue on subsequences to conclude in the case where $\limsup_{M,N} NM^{-1} < 1$.

The second step consists in proving that for some constant $\widetilde{K}$ (depending on $\limsup NM^{-1}$) there exists N_2 (depending on the realization) such that for all $N \geq N_2$:

$$|\kappa_t(\epsilon, y) - \kappa_t(y)| \leq \widetilde{K}\epsilon \ . \tag{15}$$

The proof of (17) relies on the following identity:

$$\kappa_t(y) - \kappa_t(\epsilon, y) = \epsilon\alpha_N + \beta_N\left(\kappa_t(y) - \kappa_t(\epsilon, y)\right) \ , \tag{16}$$

where

$$\begin{aligned}
\alpha_N &= \frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{T}_t(\epsilon, y)\mathbf{\Gamma}_t\mathbf{T}_t(y) \ , \\
\beta_N &= \frac{1}{M}\operatorname{tr}\left(\frac{\mathbf{\Gamma}_t\mathbf{T}_t(\epsilon, y)\mathbf{\Gamma}_t\mathbf{T}_t(y)}{(1+\kappa_t(y))(1+\kappa_t(\epsilon, y))}\right) \ .
\end{aligned}$$

It is clear that $\beta_N < 1$ and one can prove that there exists $\widetilde{K} > 0$ such that $\limsup \alpha_N < \widetilde{K}$. In fact, α_N satisfies:

$$\alpha_N \leq \frac{N}{M}\|\mathbf{\Gamma}_t\|^2\left\|\mathbf{\Gamma}_t^{-1}\right\|^2(1+\kappa_t(y))(1+\kappa_t(\epsilon, y)) \ . \tag{17}$$

One can prove that $\kappa_t(y)$ and $\kappa_t(\epsilon, y)$ are lower than $\frac{N}{M(1-N/M)}$. In fact, $\kappa_t(y)$ writes:

$$\begin{aligned}
\kappa_t(y) &= \frac{N(1+\kappa_t(y))}{M} - \frac{(1+\kappa_t(y))}{M}\operatorname{tr}\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{\Gamma}_t}{1+\kappa_t(y)}\right)^{-1}\right) \ , \\
&= \frac{N}{M(1-\frac{N}{M})} - \frac{(1+\kappa_t(y))}{M(1-\frac{N}{M})}\operatorname{tr}\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{\Gamma}_t}{1+\kappa_t(y)}\right)^{-1}\right) \ , \\
&\leq \frac{N}{M(1-\frac{N}{M})} \ .
\end{aligned}$$

Similar arguments hold for $\kappa_t(\epsilon, y)$, thus proving that $\limsup \alpha_N \leq \widetilde{K}$. From (16), we conclude that there exists N_3 such that for all $N \geq N_3$,

$$|\kappa_t(\epsilon, y) - \kappa_t(y)| \leq \widetilde{K}\epsilon \ .$$

We are now in position to prove the almost sure convergence of $\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y) - \kappa_t(y)$. Consider the constants K and $\widetilde{K}$ as defined previously and let $\epsilon > 0$. According to (14), there exists N_1 such that:

$$\forall N \geq N_1 \ , \quad \left|\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y) - \frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\epsilon, y)\right| \leq \frac{\epsilon}{K} \ .$$

Using the almost sure convergence result of $\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\epsilon, y)$ stated in Lemma 4, there exists N_2 such that:

$$\forall N \geq N_2 \ , \quad \left|\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(\epsilon, y) - \kappa_t(\epsilon, y)\right| \leq \epsilon \ .$$

Finally from (15), there exists N_3 such that for all $N \geq N_3$:

$$|\kappa_t(\epsilon, y) - \kappa_t(y)| \leq \widetilde{K}\epsilon \ .$$

Combining all these results, we have, for $N \geq \max(N_1, N_2, N_3)$:

$$\left|\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}(y) - \kappa_t(y)\right| \leq \epsilon\left(\frac{1}{K} + 1 + \widetilde{K}\right) \ ,$$

hence proving that:

$$\frac{1}{M}\operatorname{tr}\mathbf{\Gamma}_t\mathbf{Q}_t(y)-\kappa_t(y)\xrightarrow[M,N,n\to\infty]{\text{a.s.}}0\ ,$$

which is the desired result.

APPENDIX B
PROOF OF LEMMA 2

Using the same eigenvalue decomposition as in Appendix A, we can prove that $\mathbf{Y}_t=\mathbf{U}_t\mathbf{D}_t^{\frac{1}{2}}\widetilde{\mathbf{W}}_t$ where $\widetilde{\mathbf{W}}_t$ is a $N\times M$ standard Gaussian matrix, and where $\mathbf{D}_t$ is a diagonal matrix with the same eigenvalues as $\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N$. In the sequel, if $\mathbf{A}$ is a $p\times p$ hermitian matrix, denote by $F^{\mathbf{A}}$ the empirical distribution of its eigenvalues, i.e. $F^{\mathbf{A}}=\frac{1}{p}\sum_{i=1}^p\delta_{\lambda_i(\mathbf{A})}$, and by $m_{\mathbf{A}}$ the associated Stieltjes transform.

Denote by $m_{\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t}(z)$ the Stieltjes transform corresponding to the empirical eigenvalue distribution of $\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t$, i.e.,

$$m_{\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t}(z)=\frac{1}{M}\operatorname{tr}\left(\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t-z\mathbf{I}_M\right)^{-1}\ .$$

Notice that $m_{\mathbf{D}_t}(z)=m_{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}(z-\sigma^2)$. Using this fact, and the result in [14], on can easily prove that $m_{\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t}$ satisfies:

$$\forall z\in\mathbb{C}\setminus\mathbb{R}^+\ ,\qquad m_{\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t}(z)-\underline{m}(z)\xrightarrow[M,N,n\to\infty]{\text{a.s.}}0\ ,$$

where $\underline{m}(z)$ is the unique Stieltjes transform of a probability distribution $\underline{F}$, solution of the following functional equation:

$$\underline{m}(z)=\left(-z+\frac{N}{M}\int\frac{\lambda+\sigma^2}{1+(\lambda+\sigma^2)\underline{m}(z)}d\,F^{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}(\lambda)\right)^{-1}\ . \tag{18}$$

Moreover, $\underline{m}(z)$ is analytical on $\mathbb{C}^+=\{z\in\mathbb{C},\Im(z)>0\}$ where $\Im(z)$ stands for the imaginary part of $z\in\mathbb{C}$. Using (18), one can prove that $m_{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}(z)$ satisfies:

$$m_{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}\left(-\frac{1}{\underline{m}(z)}-\sigma^2\right)=\underline{m}(z)(1-\frac{M}{N})-\frac{M}{N}\,z\,\underline{m}^2(z)\ . \tag{19}$$

The link between the unobservable Stieltjes transform $m_{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}$ and the deterministic equivalent $\underline{m}(z)$ being established, it remains to express $N^{-1}\log\det(\mathbf{I}_N+\sigma^{-2}\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}})$ in terms of $m_{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}$, which follows easily by differentiation:

$$\frac{\partial}{\partial\sigma^2}\frac{1}{N}\log\det\left(\mathbf{I}_N+\frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}{\sigma^2}\right)=\frac{1}{N}\operatorname{tr}\left(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}-\frac{1}{\sigma^2}\ .$$

Hence:

$$\begin{aligned}\frac{1}{N}\log\det\left(\mathbf{I}_N+\frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}{\sigma^2}\right)&=\int_{\sigma^2}^{+\infty}\frac{1}{t}-\frac{1}{N}\operatorname{tr}\left(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\frac{1}{t}\mathbf{I}_N\right)^{-1}dt\ ,\\&=\int_0^{\frac{1}{\sigma^2}}\frac{1}{t}-\frac{1}{t^2}m_{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}\left(-\frac{1}{t}\right)dt\ .\end{aligned} \tag{20}$$

We shall now perform a change of variables within the integral in order to substitute $\underline{m}$ for $m_{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}$ with the help of (19). It has been proved in [15] that $\underline{m}(z)$ is continuous and increasing on $\mathbb{R}_-^*$; in particular, the application

$$u\mapsto\left(\frac{1}{\underline{m}(u)}+\sigma^2\right)^{-1}$$

establishes a bijection from $\mathbb{R}_-^*$ to $(0, 1/\sigma^2)$. Considering the change of variable $\frac{1}{t} = \frac{1}{\underline{m}(u)} + \sigma^2$, (20) writes:

$$\begin{aligned}
&\frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}{\sigma^2}\right) \\
&= \int_{-\infty}^0 \left[\frac{1}{\underline{m}(u)} + \sigma^2 - \left(\frac{1}{\underline{m}(u)} + \sigma^2\right)^2 m_{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}\left(-\frac{1}{\underline{m}(u)} - \sigma^2\right)\right]\frac{\underline{m}'(u)}{(1+\sigma^2\underline{m}(u))^2}\,du \\
&= \int_{-\infty}^0 \left[\frac{\underline{m}'(u)}{\underline{m}(u)(1+\sigma^2\underline{m}(u))} - \left(1 - \frac{M}{N}\right)\frac{\underline{m}'(u)}{\underline{m}} + \frac{M}{N}\,u\,\underline{m}'(u)\right]du \\
&= \int_{-\infty}^0 \left[\frac{M}{N}\frac{\underline{m}'(u)}{\underline{m}(u)} - \frac{\sigma^2\underline{m}'(u)}{1+\sigma^2\underline{m}(u)} + \frac{M}{N}u\underline{m}'(u)\right]du.
\end{aligned}$$

We shall now compute this integral, denoted by I in the sequel. Write $I = \lim_{\substack{x\to-\infty \\ y\to 0}} I_{x,y}$ where

$$I_{x,y} = \int_x^y \left[\frac{M}{N}\frac{\underline{m}'(u)}{\underline{m}(u)} - \frac{\sigma^2\underline{m}'(u)}{1+\sigma^2\underline{m}(u)} + \frac{M}{N}u\underline{m}'(u)\right]\,du\ .$$

Straightforward computations yield:

$$I_{x,y} = \log\left|\frac{(\underline{m}(y))^{\frac{M}{N}}}{1+\sigma^2\underline{m}(y)}\right| - \log\left|\frac{(\underline{m}(x))^{\frac{M}{N}}}{1+\sigma^2\underline{m}(x)}\right| + \frac{M}{N}y\underline{m}(y) - \frac{M}{N}x\underline{m}(x) - \int_x^y \frac{M}{N}\underline{m}(u)du\ . \tag{21}$$

As our objective is to compute the limit of $I_{x,y}$ as $x \to -\infty$ and $y \to 0$, we need to obtain equivalents for $\underline{m}$ at 0 and $-\infty$. A direct application of the dominated convergence theorem yields:

$$\underline{m}(x) \underset{x\to-\infty}{\sim} -\frac{1}{x}\ .$$

Recall that $\underline{F}$ is the probability distribution associated to $\underline{m}$. Then, $\underline{F}(\{0\}) = M^{-1}(M-N)$. Although this property is not easy to write down properly, it is quite intuitive if one sees $\underline{F}$ as close to $F^{\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t}$ (the empirical distribution of the eigenvalues of $\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t$) which clearly satisfies $F^{\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t}(\{0\}) = M^{-1}(M-N)$ by Assumption **A2**: This assumption implies in fact that zero is an eigenvalue of $\mathbf{Y}_t^{\mathsf{H}}\mathbf{Y}_t$ of order $M-N$. Hence,

$$\underline{m}(y) \underset{y\to 0}{\sim} -\frac{M-N}{My}\ .$$

Using these relations, we can derive equivalents for the first four terms in the right-hand side of (21). In particular, we obtain:

$$\log\left|\frac{(\underline{m}(y))^{\frac{M}{N}}}{1+\sigma^2\underline{m}(y)}\right| \underset{y\to 0}{\sim} \left(\frac{M}{N} - 1\right)\log\left(\frac{M-N}{M}\right) - \log(\sigma^2) + \left(1 - \frac{M}{N}\right)\log|y|\ , \tag{22}$$

$$-\log\left|\frac{(\underline{m}(x))^{\frac{M}{N}}}{1+\sigma^2\underline{m}(x)}\right| \underset{x\to-\infty}{\sim} \frac{M}{N}\log|x|\ , \tag{23}$$

$$\frac{M}{N}\,y\,\underline{m}(y) \underset{y\to 0}{\sim} -\left(\frac{M}{N} - 1\right)\ , \tag{24}$$

$$-\frac{M}{N}\,x\,\underline{m}(x) \underset{x\to-\infty}{\sim} \frac{M}{N}\ . \tag{25}$$

Let us now handle the last term in the. of (21). Denote by F the probability distribution defined by

$$\underline{F}(dx) = \frac{(M-N)}{M}\delta_0(dx) + \frac{N}{M}F(dx)\ .$$

If m is the Stietjes transform associated to F, then:

$$m(z) = \frac{M}{N}\underline{m}(z) + \frac{(M-N)}{N}\frac{1}{z} \ .$$

Note in particular that $m_{\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}} - m \to 0$, hence that F is a deterministic approximation of $F^{\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}}$, the empirical distribution of the eigenvalues of $\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}$. Now,

$$\begin{aligned} \int_x^y \frac{M}{N}\underline{m}(u)du &= \int_x^y \int \frac{dF(t)}{t-u}du - \frac{M-N}{Nu}du \ , \\ &= \int (-\log|t-y| + \log|t-x|)dF(t) + \frac{M-N}{N}(\log|x| - \log|y|) \ . \end{aligned} \tag{26}$$

Using the dominated convergence theorem, one can provethat the r.h.s. of (26) is equivalent to:

$$\int_x^y \frac{M}{N}\underline{m}(u)du \underset{\substack{x\to-\infty \\ y\to 0}}{\sim} -\int \log(t)dF(t) + \frac{M}{N}\log|x| - \frac{M-N}{N}\log|y| \ . \tag{27}$$

Plugging (22), (23), (24), (25) and (27) into (21) yields:

$$\lim_{\substack{x\to-\infty \\ y\to 0}} I_{x,y} = \frac{M-N}{N}\log\left(\frac{M-N}{M}\right) - \log\sigma^2 + \int \log(t)dF(t).$$

Since the spectrum of $\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}$ is almost surely eventually bounded away from zero and upper-bounded, uniformly along N, we have:

$$\frac{1}{N}\sum_{i=1}^N \log(\lambda_i) - \int \log(t)dF(t) \xrightarrow[M,N,n\to+\infty]{\text{a.s.}} 0$$

where $(\lambda_i, 1 \le i \le N)$ are the eigenvalues of $\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}$. A consistent estimator of $\frac{1}{N}\log\det(\sigma^2\mathbf{I}_N + \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}})$ is thus given by:

$$\begin{aligned} C_1 &= \frac{M-N}{N}\log\left(\frac{M-N}{M}\right) + 1 + \frac{1}{N}\sum_{i=1}^N \log(\lambda_i) \\ &= \frac{M-N}{N}\log\left(\frac{M-N}{M}\right) + 1 + \frac{1}{N}\log\det\left(\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right) \ , \end{aligned}$$

which concludes the proof.

APPENDIX C
PROOF OF THEOREM 2

As previously mentionned, the proof of Theorem 2 relies on the existence of a consistent estimate for

$$C_{t,1} = \frac{1}{N}\log\det(\sigma^2\mathbf{I}_N + \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}) \ .$$

Denote by $f(y)$ the parametrized quantity:

$$f(y) = \frac{1}{N}\log\det(\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}} + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}) \ .$$

Then by Lemma 1-3), we obtain:

$$-f(y) + \frac{1}{N}\log\det\left(\frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N}{1+\kappa_t(y)} + y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\right) + \frac{M}{N}\log(1+\kappa_t(y)) - \frac{M}{N}\frac{\kappa_t(y)}{1+\kappa_t(y)} \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 \ . \tag{28}$$

Obviously, if y is replaced by $y_{N,t}$, solution of:

$$y_{N,t} = \frac{1}{1+\kappa_t(y_{N,t})} \ , \tag{29}$$

then the term $C_{t,1}$ appears in (28). The existence and uniqueness of $y_{N,t}$ immediatly follows from the fact that the function g defined as:

$$g : x \mapsto (1+x)\frac{1}{M}\operatorname{tr}(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)(\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)^{-1}$$

is a contraction. Moreover, straightforward computations yield:

$$y_{N,t} = 1 - \frac{1}{M}\operatorname{tr}(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)(\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)^{-1} \ . \tag{30}$$

Unfortunately, $y_{N,t}$ depends on the unobservable matrix $\mathbf{G}_t$. One need therefore to provide a consistent estimate $\hat{y}_{N,t}$ of $y_{N,t}$. In order to proceed, we shall study the asymptotics of $\kappa_t(y)$. By Lemma 1-2), we have:

$$\frac{y}{M}\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(y) - \frac{y}{M}\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{T}_t(y) \xrightarrow[M,N,n\to\infty]{a.s.} 0 \ . \tag{31}$$

On the other hand, we have:

$$\begin{aligned}
\frac{y}{M}\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{T}_t(y) &= \frac{1}{M}\operatorname{tr} y\,\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N}{1+\kappa_t(y)}\right)^{-1} \ , \\
&= \frac{N}{M} - \frac{1}{M(\kappa_t(y)+1)}\operatorname{tr}\left((\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N}{1+\kappa_t(y)}\right)^{-1}\right) \ , \\
&= \frac{N}{M} - \frac{\kappa_t(y)}{1+\kappa_t(y)} \ , \\
&= \frac{N}{M} - 1 + \frac{1}{1+\kappa_t(y)} \ .
\end{aligned} \tag{32}$$

Substituting (32) into (31), we obtain:

$$\frac{1}{M}\operatorname{tr} y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(y) - \frac{N}{M} + 1 - \frac{1}{\kappa_t(y)+1} \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 \ . \tag{33}$$

Intuitively, a consistent estimate of $\hat{y}_{N,t}$ of $y_{N,t}$ should satisfy $\hat{y}_{N,t} = M^{-1}\hat{y}_{N,t}\ \operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\hat{y}_{N,t})$. This intuition is confirmed by the following lemma:

Lemma 5: There exists a unique positive solution $\hat{y}_{N,t}$ to the equation:

$$\frac{\hat{y}_{N,t}}{M}\operatorname{tr}\mathbf{Q}_t(\hat{y}_{N,t}) - \frac{N}{M} + 1 - \hat{y}_{N,t} = 0 \ .$$

Moreover, the following convergence holds true:

$$\hat{y}_{N,t} - y_{N,t} \xrightarrow[M,N,n\to\infty]{a.s.} 0 \ ,$$

where $y_{N,t}$ is defined by (29) (see also (30)).

Proof: The existence and uniqueness of $\hat{y}_{N,t}$ follows from the fact that $y :\mapsto \frac{1}{M}\operatorname{tr}\mathbf{Q}_t(y) - \frac{N}{M} + 1$ is a contraction. Moreover, using Assumption **A2**, it is straightforward to check that $\hat{y}_{N,t}$ is eventually lower than 1. Using (33), we get that:

$$\frac{y_{N,t}}{M}\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(y_{N,t}) - \frac{N}{M} + 1 - y_{N,t} \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0 \ .$$

Beware that in (33), the convergence holds true for a fixed y while $y_{N,t}$ depends upon N. A way to circumvent this issue is to merge $y_{N,t}$ into $\mathbf{H}_t$ and to consider the slightly different model based on $\widetilde{\mathbf{H}}_t = \sqrt{y_{N,t}}\mathbf{H}_t$.

Therefore, the mere definition of $\hat{y}_{N,t}$ and the previous convergence yield:

$$\frac{\hat{y}_{N,t}}{M}\operatorname{tr}(\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\hat{y}_{N,t})) - \hat{y}_{N,t} + y_{N,t} - \frac{y_{N,t}}{M}\operatorname{tr}(\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(y_{N,t})) \xrightarrow[M,N,n\to\infty]{\text{a.s.}} 0.$$

It can be easily proved that $h_N : y \mapsto \frac{y}{M}\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(y)$ is a contraction on $\mathbb{R}^+$, i.e. that there exists $0 \le k_N \le 1$ such that:

$$|h_N(x) - h_N(y)| \le k_N|x-y|,$$

whenever $x, y \ge 0$; moreover, due to Assumption **A2**, $\limsup k_N < 1$. On the other hand, we have:

$$\begin{aligned}
|y_{N,t} - \hat{y}_{N,t}| &= |y_{N,t} - \hat{y}_{N,t} - h_N(y_{N,t}) + h_N(\hat{y}_{N,t}) - h_N(\hat{y}_{N,t}) + h_N(y_{N,t})| \ , \\
&\le |y_{N,t} - \hat{y}_{N,t} - h_N(y_{N,t}) + h_N(\hat{y}_{N,t})| + |h_N(\hat{y}_{N,t}) - h_N(y_{N,t})| \ , \\
&\le |y_{N,t} - \hat{y}_{N,t} - h_N(y_{N,t}) + h_N(\hat{y}_{N,t})| + k_N|\hat{y}_{N,t} - y_{N,t}| \ .
\end{aligned}$$

Hence, we get:

$$0 \le (1-k_N)|\hat{y}_{N,t} - y_{N,t}| \le |y_{N,t} - \hat{y}_{N,t} - h(y_{N,t}) + h(\hat{y}_{N,t})| \ .$$

Since the r.h.s. converges to zero, $y_{N,t} - \hat{y}_{N,t}$ converges also to zero almost surely. ■

With the help of Lemma 5, the following convergences can be easily verified:

$$\begin{aligned}
\frac{1}{N}\log\det(\mathbf{Q}_t(y_{N,t})) - \frac{1}{N}\log\det(\mathbf{Q}_t(\hat{y}_{N,t})) \quad &\xrightarrow[M,N,n\to\infty]{\text{a.s.}} \quad 0 \ , \\
\kappa(\hat{y}_{N,t}) - \kappa(y_{N,t}) \quad &\xrightarrow[M,N,n\to\infty]{a.s.} \quad 0 \ .
\end{aligned}$$

Therefore:

$$-f(\hat{y}_{N,t}) + \frac{1}{N}\log\det(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N + \mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}) - \frac{M-N}{N}\log(\hat{y}_{N,t}) - \frac{M}{N}(1-\hat{y}_{N,t}) \xrightarrow[M,N,n\to\infty]{a.s.} 0 \ ,$$

which in turn implies that:

$$C_{t,1} - \log\det(\hat{y}_{N,t}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}) - \frac{M-N}{N}\log(\hat{y}_{N,t}) - \frac{M}{N}(1-\hat{y}_{N,t}) \xrightarrow[M,N,n\to\infty]{a.s.} 0 \ .$$

Using this estimate of $C_{t,1}$ together with the estimate of $C_{t,2}$ as provided in Lemma 2 immediatly yields a consistent estimate for $C_t(\sigma^2) = C_{t,1} - C_{t,2}$, and the theorem is proved.

APPENDIX D
PROOF OF THEOREM 3

The proof of theorem 3 relies on the tools used in [11] suitable for dealing with Gaussian random variables. Recall that $\hat{C}_{\text{trad}}(y)$ is given by:

$$\hat{C}_{\text{trad}}(y) = \frac{1}{NT}\sum_{t=1}^{T}\log\det\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right) - \log\det\left(\frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right),$$

where $\mathbf{Y}_t = [\sigma \mathbf{I}_N \ \ \mathbf{G}_t] \begin{bmatrix} \mathbf{W}_t \\ \mathbf{X}_t \end{bmatrix}$ and $\mathbf{X}_t = \left[\mathbf{X}_{t,1}^{\mathsf{H}}, \ldots, \mathbf{X}_{t,K}^{\mathsf{H}}\right]^{\mathsf{H}}$. Similarly, as in Appendix A and Appendix B, we can prove that $\mathbf{Y}_t = \mathbf{U}_t \mathbf{D}_t^{\frac{1}{2}} \widetilde{\mathbf{W}}_t$ where $\widetilde{\mathbf{W}}_t$ is a $N \times M$ standard Gaussian matrix, and $\mathbf{D}_t$ is the $N \times N$ diagonal matrix containing the eigenvalues of $\mathbf{G}_t \mathbf{G}_t^{\mathsf{H}} + \sigma^2 \mathbf{I}_N$. Then, $\hat{C}_{\rm trad}(y)$ becomes:

$$\begin{aligned}
\hat{C}_{\rm trad}(y) &= \frac{1}{NT} \sum_{t=1}^{T} \log\det(y \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} + \frac{1}{M} \mathbf{U}_t \mathbf{D}_t^{\frac{1}{2}} \widetilde{\mathbf{W}}_t \widetilde{\mathbf{W}}_t^{\mathsf{H}} \mathbf{D}_t^{\frac{1}{2}} \mathbf{U}_t^{\mathsf{H}}) - \log\det(\frac{1}{M} \mathbf{D}_t^{\frac{1}{2}} \widetilde{\mathbf{W}}_t \widetilde{\mathbf{W}}_t^{\mathsf{H}} \mathbf{D}_t^{\frac{1}{2}}), \\
&= \frac{1}{NT} \sum_{t=1}^{T} \log\det(y \mathbf{D}_t^{-\frac{1}{2}} \mathbf{U}_t^{\mathsf{H}} \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \mathbf{U}_t \mathbf{D}_t^{-\frac{1}{2}} + \frac{1}{M} \widetilde{\mathbf{W}}_t \widetilde{\mathbf{W}}_t^{\mathsf{H}}) - \log\det(\frac{1}{M} \widetilde{\mathbf{W}}_t \widetilde{\mathbf{W}}_t^{\mathsf{H}}), \\
&= \frac{1}{NT} \sum_{t=1}^{T} \log\det\left(y \mathbf{D}_t^{-\frac{1}{2}} \mathbf{U}_t^{\mathsf{H}} \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \mathbf{U}_t \mathbf{D}_t^{-\frac{1}{2}} \left(\frac{1}{M} \widetilde{\mathbf{W}}_t \widetilde{\mathbf{W}}_t^{\mathsf{H}} \right)^{-1} + \mathbf{I}_N \right).
\end{aligned}$$

Denote by $\mathbf{D}_t^{-\frac{1}{2}} \mathbf{U}_t^{\mathsf{H}} \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \mathbf{U}_t \mathbf{D}_t^{-\frac{1}{2}} = \widetilde{\mathbf{U}}_t \mathbf{\Lambda}_t \widetilde{\mathbf{U}}_t^{\mathsf{H}}$ be the eigenvalue decomposition of $\mathbf{D}_t^{-\frac{1}{2}} \mathbf{U}_t^{\mathsf{H}} \mathbf{H}_t \mathbf{H}_t^{\mathsf{H}} \mathbf{U}_t \mathbf{D}_t^{-\frac{1}{2}}$. Since r is the rank of $\mathbf{H}_t \mathbf{H}_t^{\mathsf{H}}$, matrix $\mathbf{\Lambda}_t$ has exactly r non zero entries which we denote by $(\Lambda_{i,t}, 1 \leq i \leq r)$. We get then:

$$\hat{C}_{\rm trad}(y) = \frac{1}{NT} \sum_{t=1}^{T} \log\det\left(y \mathbf{\Lambda}_t \left(\frac{1}{M} \widetilde{\mathbf{W}}_t \widetilde{\mathbf{W}}_t^{\mathsf{H}} \right)^{-1} + \mathbf{I}_N \right).$$

Let $\mathbf{\Lambda}_{r,t} = \mathrm{diag}\left(\lambda_{1,t}, \ldots, \lambda_{r,t}\right)$. Then using theorem 3.2.11 in [16], we can prove that $\hat{C}_{\rm trad}(y)$ can be written as:

$$\hat{C}_{\rm trad}(y) = \frac{1}{NT} \sum_{t=1}^{T} \log\det\left(y \mathbf{\Lambda}_{r,t} \left(\frac{1}{M} \widetilde{\mathbf{W}}_{r,t} \widetilde{\mathbf{W}}_{r,t}^{\mathsf{H}} \right)^{-1} + \mathbf{I}_N \right),$$

where $\widetilde{\mathbf{W}}_{r,t}$ is a $r \times M - N + r$ standard Gaussian matrix. Let $\mathbf{M} = \frac{(M-N+r)}{My} \mathbf{\Lambda}_{r,t}^{-1}$, we finally get:

$$\begin{aligned}
\hat{C}_{\rm trad}(y) &= \frac{1}{NT} \sum_{t=1}^{T} \log\det\left(\frac{1}{M-N+r} \mathbf{M}^{\frac{1}{2}} \widetilde{\mathbf{W}}_{r,t} \widetilde{\mathbf{W}}_{r,t}^{\mathsf{H}} \mathbf{M}^{\frac{1}{2}} + \mathbf{I}_r \right) - \log\det\left(\mathbf{M}\right) - \log\det\left(\frac{1}{M-N+r} \widetilde{\mathbf{W}}_{r,t} \widetilde{\mathbf{W}}_{r,t}^{\mathsf{H}} \right) \\
&\triangleq \sum_{t=1}^{T} \hat{C}_{\rm trad,t}(y).
\end{aligned}$$

Let $s = M - N + r$. By Assumption **A4**, we have:

$$0 < \liminf \frac{s}{r} \leq \limsup \frac{s}{r} < +\infty.$$

Moreover, matrix $\mathbf{M}$ satisfies:

$$\sup \|\mathbf{M}\| < \infty \text{ and } \inf \frac{1}{s} \operatorname{tr} \mathbf{M} > 0.$$

We retrieve then the same model as in [11], with the slight difference that $\hat{C}_{\rm trad,t}(y)$ has an extra random term $\log\det\left(\frac{1}{M} \widetilde{\mathbf{W}}_{r,t} \widetilde{\mathbf{W}}_{r,t}^{\mathsf{H}}\right)$. As we will see next, this has no impact on the applicability of the method and one can get the desired result by following the same lines of [11]. In particular, we consider to prove a CLT for the functional $\log\det(\frac{z}{s} \mathbf{M}^{\frac{1}{2}} \widetilde{\mathbf{W}} \widetilde{\mathbf{W}}^{\mathsf{H}} \mathbf{M}^{\frac{1}{2}} + \mathbf{I}_r) - \log\det(\frac{1}{s} \mathbf{M}^{\frac{1}{2}} \widetilde{\mathbf{W}} \widetilde{\mathbf{W}}^{\mathsf{H}} \mathbf{M}^{\frac{1}{2}})$ where $z > 0$. The expression of the variance for this CLT will depend on some deterministic quantities which we recall hereafter.

A. Notations

Let $\mathbf{Z} = \mathbf{M}^{\frac{1}{2}}\widetilde{\mathbf{W}}$ and define the resolvent matrix $\mathbf{S}(z)$ by:

$$\mathbf{S}(z) = \left(\frac{z}{s}\mathbf{M}^{\frac{1}{2}}\widetilde{\mathbf{W}}\widetilde{\mathbf{W}}^{\mathsf{H}}\mathbf{M}^{\frac{1}{2}} + \mathbf{I}_r\right)^{-1} = \left(\frac{z}{s}\mathbf{Z}\mathbf{Z}^{\mathsf{H}} + \mathbf{I}_r\right)^{-1},$$

Let also $I_s(z)$ be given by:

$$I_s(z) = \log\det\left(\frac{z}{s}\mathbf{M}^{\frac{1}{2}}\widetilde{\mathbf{W}}\widetilde{\mathbf{W}}^{\mathsf{H}}\mathbf{M}^{\frac{1}{2}} + \mathbf{I}_r\right) = -\log\det\mathbf{S}(z).$$

We introduce the following intermediate quantities:

$$\beta(z) = \frac{1}{s}\operatorname{tr}\mathbf{M}\mathbf{S}, \quad \alpha(z) = \frac{1}{s}\operatorname{tr}\mathbf{M}\mathbb{E}\mathbf{S}, \quad \text{and} \quad \overset{o}{\beta} = \beta - \alpha.$$

Matrix $\widetilde{\mathbf{R}}(z)$ is a $s \times s$ diagonal matrix defined by:

$$\widetilde{\mathbf{R}}(z) = \tilde{r}\mathbf{I}_s,$$

where $\tilde{r} = \frac{1}{1+z\alpha(z)}$. We also define $\mathbf{R}(z)$ the $r \times r$ matrix given by:

$$\mathbf{R}(z) = \left(\mathbf{I}_r + z\tilde{r}(z)\mathbf{M}\right)^{-1}.$$

We also define $\delta(z)$ as the unique positive solution of the following equation:

$$\delta(z) = \frac{1}{s}\operatorname{tr}\mathbf{M}\left(\mathbf{I}_r + \frac{z}{1+z\delta(z)}\mathbf{M}\right)^{-1},$$

where the existence and uniqueness of $\delta(z)$ have already been proven in [11]. Let $\mathbf{\Xi}$ and $\widetilde{\mathbf{\Xi}}$ the $r \times r$ and $s \times s$ diagonal matrices defined by:

$$\mathbf{\Xi} = \left(\mathbf{I}_r + \frac{z}{1+z\delta(z)}\mathbf{M}\right)^{-1} \text{ and } \widetilde{\mathbf{\Xi}} = \frac{1}{1+z\delta(z)}\mathbf{I}_s$$

Define also γ, $\tilde{\delta}(z)$ and $\tilde{\gamma}$ as $\gamma = \frac{1}{s}\operatorname{tr}\mathbf{M}^2\mathbf{\Xi}^2$, $\tilde{\delta}(z) = \frac{1}{1+z\delta(z)}$ and $\widetilde{\gamma} = \frac{1}{(1+z\delta(z))^2}$.

B. Mathematical tools

We recall here the mathematical tools that will be used to establish theorem 3.

1) Differentiation formulas:

$$\frac{\partial S_{p,q}}{\partial Z_{i,j}} = -\frac{z}{s}\left[\mathbf{Z}^{\mathsf{H}}\mathbf{S}\right]_{j,q} S_{p,i},$$

$$\frac{\partial S_{p,q}}{\partial Z_{i,j}^*} = -\frac{z}{s}\left[\mathbf{S}\mathbf{Z}\right]_{p,j} S_{i,q},$$

$$\frac{\partial I_s(z)}{\partial Z_{i,j}^*} = \frac{z}{s}\left[\mathbf{S}\mathbf{Z}\right]_{i,j},$$

$$\frac{\partial \log\det(\frac{1}{s}\mathbf{Z}\mathbf{Z}^{\mathsf{H}})}{\partial Z_{i,j}^*} = \left[\left(\mathbf{Z}\mathbf{Z}^{\mathsf{H}}\right)^{-1}\mathbf{Z}\right]_{i,j}.$$

2) Integration by parts formula for Gaussian functionals: Denote by Φ be a $\mathcal{C}^1$ complex function polynomially bounded with its derivatives, then

$$\mathbb{E}\left[Z_{i,j}\Phi(\mathbf{Z})\right] = m_i \mathbb{E}\left[\frac{\partial \Phi(\mathbf{Z})}{\partial Z_{i,j}^*}\right].$$

where m_i is the i-th diagonal element of $\mathbf{M}$.

3) Poincaré-Nash inequality: The variance of $\Phi(\mathbf{Z})$ can be upper-bounded as:

$$\operatorname{var}(\Phi(\mathbf{Z})) \leq \sum_{i=1}^{r}\sum_{j=1}^{s} m_i \mathbb{E}\left[\left|\frac{\partial \Phi(\mathbf{Z})}{\partial Z_{i,j}}\right|^2 + \left|\frac{\partial \Phi(\mathbf{Z})}{\partial Z_{i,j}^*}\right|^2\right].$$

4) Deterministic approximations of some functionals:

Proposition 1: Let $\mathbf{A}_r$ and $\mathbf{B}_r$ be two sequences of respectively $r\times r$ and $s\times s$ diagonal deterministic matrices with uniformly bounded spectral norm. Assume that assumptions **A1-A4** hold true. Then, the following holds true:

$$\frac{1}{s}\operatorname{tr}\mathbf{A}_r\mathbf{R} = \frac{1}{s}\operatorname{tr}\mathbf{A}_r\mathbf{\Xi} + \mathcal{O}\left(s^{-2}\right), \quad \tilde{r} = \tilde{\delta} + \mathcal{O}\left(s^{-2}\right) \quad \text{and} \quad \mathbb{E}\frac{1}{s}\operatorname{tr}\mathbf{A}_r\mathbf{H} = \frac{1}{s}\operatorname{tr}\mathbf{A}_r\mathbf{\Xi} + \mathcal{O}\left(s^{-2}\right).$$

Proposition 2: Let $\mathbf{A}_r$, $\mathbf{B}_r$ and $\mathbf{C}_r$ be three sequences of $r\times r$, $s\times s$ and $r\times r$ diagonal deterministic matrices whose spectral norm are uniformly bounded in r. Consider the following:

$$\Phi(\mathbf{Z}) = \frac{1}{s}\operatorname{tr}\left(\mathbf{A}_r\mathbf{S}\frac{\mathbf{Z}\mathbf{B}_r\mathbf{Z}^{\mathsf{H}}}{s}\right), \quad \Psi(\mathbf{Z}) = \frac{1}{s}\operatorname{tr}\left(\mathbf{A}_r\mathbf{SMS}\frac{\mathbf{Z}\mathbf{B}_r\mathbf{Z}^{\mathsf{H}}}{s}\right),$$

and assume that **A1-A4** hold true. Then,

a) The following estimations hold true: $\operatorname{var}(\Phi(\mathbf{Z})), \operatorname{var}(\Psi(\mathbf{Z})), \operatorname{var}(\beta)$ are $\mathcal{O}\left(s^{-2}\right)$.

b) The following approximations hold true:

$$\mathbb{E}\left[\Phi(\mathbf{Z})\right] = \tilde{\delta}\frac{1}{s}\operatorname{tr}\mathbf{A}_r\mathbf{M}\mathbf{\Xi} + \mathcal{O}\left(s^{-2}\right), \tag{34}$$

$$\mathbb{E}\left[\Psi(\mathbf{Z})\right] = \frac{1}{1-z^2\gamma\tilde{\gamma}}\left(\tilde{\delta}\frac{1}{s}\operatorname{tr}\mathbf{B}_r\frac{1}{s}\operatorname{tr}(\mathbf{A}_r\mathbf{M}^2\mathbf{\Xi}^2) - z\gamma\tilde{\gamma}\frac{1}{s}\operatorname{tr}\mathbf{B}_r\frac{1}{s}\operatorname{tr}\mathbf{A}_r\mathbf{M}\mathbf{\Xi}\right) + \mathcal{O}\left(s^{-2}\right), \tag{35}$$

$$\mathbb{E}\left[\frac{1}{s}\operatorname{tr}\mathbf{MSMS}\right] = \frac{\gamma}{1-z^2\gamma\tilde{\gamma}} + \mathcal{O}\left(s^{-2}\right). \tag{36}$$

C. Central limit theorem

All the notations being defined, we are now in position to show the CLT. We recall that our objective is to study the fluctuations of $\hat{C}_{\mathrm{trad}}(y) = \sum_{t=1}^{T}\hat{C}_{\mathrm{trad},t}(y)$. Since $\left(\hat{C}_{\mathrm{trad},t}(y), t=1,\cdots,T\right)$ are independent, it suffices to consider the CLT for $\hat{C}_{\mathrm{trad},t}(y)$, for $t\in\{1,\cdots,T\}$. We consider thus the random quantity $I_s(z) - \log\det\left(\frac{1}{s}\mathbf{Z}\mathbf{Z}^{\mathsf{H}}\right)$. Before getting into the proof details, we shall first recall the CLT of $g(\mathbf{Z}) = -\log\det(\frac{1}{s}\mathbf{Z}\mathbf{Z}^{\mathsf{H}})$ whose proof can be found in [17]. Indeed, it is shown that:

$$\frac{-1}{\log(1-\frac{r}{s})}\left(-\log\det(\frac{1}{s}\mathbf{Z}\mathbf{Z}^{\mathsf{H}}) - b_s\right) \xrightarrow[N,M,n\to\infty]{\mathcal{D}} \mathcal{N}(0,1).$$

where $b_s = -r\left[\left(1-\frac{s}{r}\right)\log\left(1-\frac{r}{s}\right)-1\right]$. Like in [11], define $\Psi_s(u,z) = \mathbb{E}\left[e^{\jmath u(I_s(z)-V_s(z)+g(\mathbf{Z})-b_s)}\right]$, where $V_s(z)$ is the deterministic equivalent defined by:

$$V_s(z) = s\log\left(1+z\delta(z)\right)+\log\det\left(\mathbf{I}_r+\frac{z}{1+z\delta(z)}\mathbf{M}\right)-sz\delta(z)\tilde{\delta}(z),$$

and verifying:

$$\frac{1}{s}\left(I_s(z)-V_s(z)\right)\xrightarrow[r,s\to\infty]{\text{a.s}}0.$$

The principle of the proof is to establish a differential equation verified by $\Psi_s(u,z)$. Writing the derivative of $\Psi_s(u,z)$ with respect to z, we get:

$$\frac{\partial\Psi_s}{\partial z} = \mathbb{E}\left[\jmath u\frac{\partial I_s(z)}{\partial z}e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]e^{-\jmath uV_s(z)-\jmath ub_s}-\jmath u\frac{\partial V_s(z)}{\partial z}\Psi_s(u,z). \tag{37}$$

On the other hand, we have:

$$\begin{aligned}\mathbb{E}\left[\frac{\partial I_s(z)}{\partial z}e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right] &= \mathbb{E}\left[\operatorname{tr}\left(\frac{\mathbf{SZZ}^{\mathsf{H}}}{s}\right)e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]\\ &= \frac{1}{s}\sum_{p,i=1}^{r}\sum_{j=1}^{s}\mathbb{E}\left[Z_{i,j}S_{p,i}Z_{p,j}^*e^{\jmath uI(z)+\jmath ug(\mathbf{Z})}\right].\end{aligned}$$

Applying the integration by part formula, we get:

$$\begin{aligned}\mathbb{E}\left[Z_{i,j}S_{p,i}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right] &= \mathbb{E}\left[m_i\frac{\partial}{\partial Z_{i,j}^*}\left[S_{p,i}Z_{p,j}^*e^{\jmath uI(z)+\jmath ug(\mathbf{Z})}\right]\right]\\ &= \mathbb{E}\left[m_iS_{p,i}\delta(p-i)e^{\jmath uI(z)+\jmath ug(\mathbf{Z})}\right]\\ &\quad -\frac{z}{s}\mathbb{E}\left[\left[\mathbf{SZ}\right]_{p,j}m_iS_{i,i}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]\\ &\quad +\frac{\jmath uz}{n}\mathbb{E}\left[m_iS_{p,i}Z_{p,j}^*\left[\mathbf{SZ}\right]_{i,j}e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]\\ &\quad +\mathbb{E}\left[\jmath um_iS_{p,i}Z_{p,j}^*\frac{\partial g(\mathbf{Z})}{\partial Z_{i,j}^*}e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right].\end{aligned}$$

After summing over index i, we obtain:

$$\begin{aligned}\mathbb{E}\left[\left[\mathbf{SZ}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right] &= \mathbb{E}\left[m_pS_{p,p}e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]\\ &\quad -\frac{z}{s}\mathbb{E}\left[\operatorname{tr}(\mathbf{MS})\left[\mathbf{SZ}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]\\ &\quad +\frac{\jmath zu}{s}\mathbb{E}\left[\left[\mathbf{SMSZ}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]\\ &\quad -\jmath u\mathbb{E}\left[\left[\mathbf{SM}\left(\mathbf{ZZ}^{\mathsf{H}}\right)^{-1}\mathbf{Z}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right].\end{aligned} \tag{38}$$

Recall the relation $\beta=\frac{1}{s}\operatorname{tr}\mathbf{MS}$ and $\overset{o}{\beta}=\beta-\alpha$ where $\alpha=\frac{1}{s}\operatorname{tr}\mathbf{M}\mathbb{E}\mathbf{S}$. Plugging the relation $\beta=\alpha+\overset{o}{\beta}$ into (39), we get:

$$\begin{aligned}\mathbb{E}\left[\left[\mathbf{SZ}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right] &= \mathbb{E}\left[m_pS_{p,p}e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]-z\mathbb{E}\left[\overset{o}{\beta}\left[\mathbf{SZ}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]\\ &\quad -z\alpha\mathbb{E}\left[\left[\mathbf{SZ}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]+\frac{\jmath zu}{s}\mathbb{E}\left[\left[\mathbf{SMSZ}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right]\\ &\quad -\jmath u\mathbb{E}\left[\left[\mathbf{SM}\left(\mathbf{ZZ}^{\mathsf{H}}\right)^{-1}\mathbf{Z}\right]_{p,j}Z_{p,j}^*e^{\jmath uI_s(z)+\jmath ug(\mathbf{Z})}\right].\end{aligned} \tag{39}$$

Hence, solving this equation with respect to $\mathbb{E}\left[[\mathbf{SZ}]_{p,j}\, Z_{p,j}^* e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]$ and using the fact that $\tilde{r}=\frac{1}{1+z\alpha}$, we get:

$$\begin{aligned}\mathbb{E}\left[[\mathbf{SZ}]_{p,j}\, Z_{p,j}^* e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= \mathbb{E}\left[m_p\tilde{r}S_{p,p}e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - z\mathbb{E}\left[\overset{o}{\beta}\,\tilde{r}\,[\mathbf{SZ}]_{p,j}\, Z_{p,j}^* e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &+\frac{z}{s}\mathbb{E}\left[\jmath u\tilde{r}\,[\mathbf{SMSZ}]_{p,j}\, Z_{p,j}^* e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &-\jmath u\mathbb{E}\left[\tilde{r}\left[\mathbf{SM}\left(\mathbf{ZZ}^{\mathsf{H}}\right)^{-1}\mathbf{Z}\right]_{p,j} Z_{p,j}^* e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]. \end{aligned} \tag{40}$$

Using the relation $S_{p,p}=1-\frac{z}{s}\left[\mathbf{SZZ}^{\mathsf{H}}\right]_{p,p}$, we get after summing with respect to j,

$$\begin{aligned}\mathbb{E}\left[\left[\frac{\mathbf{SZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= \mathbb{E}\left[m_p\tilde{r}e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - zm_p\tilde{r}\left[\left[\frac{\mathbf{SZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &- z\mathbb{E}\left[\overset{o}{\beta}\,\tilde{r}\left[\frac{\mathbf{SZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]+\frac{\jmath u z}{s}\mathbb{E}\left[\tilde{r}\left[\mathbf{SMS}\frac{\mathbf{ZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &-\jmath u\mathbb{E}\left[\tilde{r}\left[\frac{\mathbf{SM}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right].\end{aligned}$$

Using the relation $r_p=\frac{1}{1+z\tilde{r}m_p}$, we get:

$$\begin{aligned}\mathbb{E}\left[\left[\frac{\mathbf{SZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= \mathbb{E}\left[m_p r_p\tilde{r}e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - z\mathbb{E}\left[\overset{o}{\beta}\,\tilde{r}r_p\left[\frac{\mathbf{SZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &+\frac{\jmath u z}{s}\mathbb{E}\left[\tilde{r}r_p\left[\mathbf{SMS}\frac{\mathbf{ZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]-\jmath u\mathbb{E}\left[\tilde{r}r_p\left[\frac{\mathbf{SM}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right].\end{aligned}$$

Summing over p, we finally get:

$$\begin{aligned}\mathbb{E}\left[\operatorname{tr}\left(\frac{\mathbf{SZZ}^{\mathsf{H}}}{s}\right) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= \tilde{r}\operatorname{tr}(\mathbf{MR})\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - z\mathbb{E}\left[\overset{o}{\beta}\,\tilde{r}\operatorname{tr}\left(\mathbf{RS}\frac{\mathbf{ZZ}^{\mathsf{H}}}{s}\right) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &+\frac{z}{s}\jmath u\mathbb{E}\left[\tilde{r}\operatorname{tr}\left(\mathbf{RSMS}\frac{\mathbf{ZZ}^{\mathsf{H}}}{n}\right) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &-\jmath u\tilde{r}\mathbb{E}\left[\operatorname{tr}\left(\frac{\mathbf{RSM}}{s}\right) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &=\chi_1+\chi_2+\chi_3+\chi_4.\end{aligned}$$

It remains thus to deal with the terms $(\chi_i, 1\leq i\leq 4)$. Using proposition 1, we have:

$$\chi_1=\tilde{r}\operatorname{tr}\mathbf{MR}\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]=s\delta\tilde{\delta}\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]+\mathcal{O}\left(s^{-1}\right). \tag{41}$$

To deal with χ_3, we apply the results of proposition 2-b, with $\mathbf{A}_r=\mathbf{R}$ and $\mathbf{B}_r=\mathbf{I}$. In this case, χ_3 writes as : $\chi_3=z\jmath u\tilde{r}\mathbb{E}\Psi(\mathbf{Z})e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}$. Using Cauchy-Schwartz inequality, we get:

$$\left|\mathbb{E}\left(\Psi(\mathbf{Z})e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right)-\mathbb{E}e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\mathbb{E}\left(\Psi(\mathbf{Z})\right)\right|\leq\sqrt{\mathbb{E}\left[\left|\overset{o}{\Psi}(\mathbf{Z})\right|^2\right]},$$

where $\overset{o}{\Psi}(\mathbf{Z})=\Psi(\mathbf{Z})-\mathbb{E}\left(\Psi(\mathbf{Z})\right)$. Therefore,

$$\chi_3=\frac{z\jmath u\tilde{\delta}}{1-z^2\gamma\tilde{\gamma}}\left[\tilde{\delta}\frac{1}{n}\operatorname{tr}(\mathbf{M}^2\Xi^3)-\frac{z\gamma\tilde{\gamma}}{s}\operatorname{tr}\mathbf{M}\Xi^2\right]\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]+\mathcal{O}\left(s^{-1}\right). \tag{42}$$

The term χ_2 can be dealt with in the same way, thus proving:

$$\chi_2 = -z\mathbb{E}\left[\overset{o}{\beta}\, e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\tilde{\gamma}\,\mathrm{tr}(\mathbf{M}\Xi^2)+\mathcal{O}\left(s^{-1}\right). \tag{43}$$

Since $\mathrm{tr}(\mathbf{M}\mathbf{\Xi}^2)$ is of order s, we shall expand $\mathbb{E}\left[\overset{o}{\beta}\, e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]$ to at least the order s^{-1}, and thus $\overset{o}{\beta}$ and $\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]$ cannot be separated in the same way as above.

Indeed, we shall first take the sum over j in (40), thus yielding:

$$\begin{aligned}\mathbb{E}\left[\left[\mathbf{SZZ}^{\mathsf{H}}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= \mathbb{E}\left[s m_p \tilde{r} S_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - z\mathbb{E}\left[\overset{o}{\beta}\,\tilde{r}\left[\mathbf{SZZ}^{\mathsf{H}}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &+\frac{z}{s}\mathbb{E}\left[\jmath u\tilde{r}\left[\mathbf{SMSZZ}^{\mathsf{H}}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - \jmath u\mathbb{E}\left[\tilde{r}\left[\mathbf{SM}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right].\end{aligned} \tag{44}$$

Using the fact that:

$$\frac{z}{s}\left[\left[\mathbf{SZZ}^{\mathsf{H}}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] = \mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - \mathbb{E}\left[S_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right],$$

(44) becomes:

$$\begin{aligned}\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - \mathbb{E}\left[S_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= z\mathbb{E}\left[m_p\tilde{r} S_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - z^2\mathbb{E}\left[\overset{o}{\beta}\,\tilde{r}\left[\frac{\mathbf{SZZ}^{\mathsf{H}}}{s}\right] e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &+\frac{z^2}{s}\mathbb{E}\left[\jmath u\tilde{r}\left[\mathbf{SMS}\frac{\mathbf{ZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - \frac{\jmath u z}{s}\mathbb{E}\left[\tilde{r}\left[\mathbf{SM}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right].\end{aligned} \tag{45}$$

Solving $\mathbb{E}\left[S_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]$ in (45) and using the relation $r_p = \frac{1}{1+z m_p \tilde{r}}$, we obtain:

$$\begin{aligned}\mathbb{E}\left[S_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= \mathbb{E}\left[r_p e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \frac{z^2}{s}\mathbb{E}\left[\overset{o}{\beta}\, r_p\tilde{r}\left[\mathbf{SZZ}^{\mathsf{H}}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &-\frac{z^2}{s}\mathbb{E}\left[\jmath u\tilde{r} r_p\left[\mathbf{SMS}\frac{\mathbf{ZZ}^{\mathsf{H}}}{s}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \frac{\jmath u z}{s}\mathbb{E}\left[\tilde{r} r_p\left[\mathbf{SM}\right]_{p,p} e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right].\end{aligned} \tag{46}$$

Multiplying both sides in (46) by m_p and summing over p, we get:

$$\begin{aligned}\mathbb{E}\left[\overset{o}{\beta}\, e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= \mathbb{E}\left[\frac{1}{s}\,\mathrm{tr}(\mathbf{MR}-\mathbf{M}\mathbb{E}\mathbf{S}) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \frac{z^2}{s}\mathbb{E}\left[\overset{o}{\beta}\,\frac{\tilde{r}}{s}\,\mathrm{tr}(\mathbf{MRSZZ}^{\mathsf{H}}) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &-\frac{z^2}{s}\mathbb{E}\left[\jmath u\tilde{r}\frac{1}{s}\,\mathrm{tr}(\mathbf{MRSMS}\frac{\mathbf{ZZ}^{\mathsf{H}}}{s}) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \frac{\jmath u z}{s^2}\tilde{r}\mathbb{E}\left[\mathrm{tr}\left(\mathbf{RMS}\right) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right].\end{aligned}$$

Using the approximating expressions in proposition 2, we get:

$$\begin{aligned}\mathbb{E}\left[\overset{o}{\beta}\, e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= z^2\gamma\tilde{\gamma}\mathbb{E}\left[\overset{o}{\beta}\, e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] - \frac{z^2\tilde{\delta}\jmath u}{s(1-z^2\gamma\tilde{\gamma})}\left(\tilde{\delta}\frac{1}{s}\,\mathrm{tr}(\mathbf{M}^3\Xi^3) - z\gamma^2\tilde{\gamma}\right)\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &+\frac{\jmath u z}{s^2}\tilde{r}\mathbb{E}\left[\mathrm{tr}(\mathbf{MRSM}) e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \mathcal{O}\left(s^{-2}\right).\end{aligned}$$

Hence,

$$\begin{aligned}\mathbb{E}\left[\overset{o}{\beta}\, e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] &= -\frac{z^2\jmath u}{s(1-z^2\gamma\tilde{\gamma})^2}\left(\tilde{\gamma}\frac{1}{s}\,\mathrm{tr}(\mathbf{M}^3\Xi^3) - z\gamma^2\tilde{\delta}^3\right)\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right]\\ &+\frac{\jmath u z\tilde{\delta}\gamma}{s(1-z^2\gamma\tilde{\gamma})}\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \mathcal{O}\left(s^{-2}\right).\end{aligned} \tag{47}$$

Plugging (47) into (43), we get:

$$\chi_2 = \frac{z^3 \jmath u \tilde{\gamma}}{s(1-z^2\gamma\tilde{\gamma})^2}\left(\tilde{\gamma}\frac{1}{s}\operatorname{tr}(\mathbf{M}^3\mathbf{\Xi}^3) - z\gamma^2\tilde{\delta}^3\right)\operatorname{tr}(\mathbf{M}\mathbf{\Xi}^2)\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] \tag{48}$$

$$-\frac{\jmath u z^2\gamma\tilde{\delta}^3}{(1-z^2\gamma\tilde{\gamma})}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right)\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \mathcal{O}\left(s^{-1}\right). \tag{49}$$

Finally, it remains to deal with χ_4. Using proposition 1, we get:

$$\chi_4 = -\frac{\jmath u \tilde{\delta}}{s}\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right)\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \mathcal{O}\left(s^{-1}\right). \tag{50}$$

From (41), (42), (49) and (50), we then have:

$$\mathbb{E}\left[\operatorname{tr}\left(\frac{\mathbf{S}\mathbf{Z}\mathbf{Z}^{\mathsf{H}}}{s}\right)e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] = \left[s\delta\tilde{\delta} + \frac{z^3\jmath u\tilde{\gamma}^2}{s(1-z^2\gamma\tilde{\gamma})^2}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}^3\mathbf{\Xi}^3\right)\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right) + \frac{z\jmath u\tilde{\gamma}}{1-z^2\gamma\tilde{\gamma}}\frac{1}{s}\operatorname{tr}(\mathbf{M}^2\mathbf{\Xi}^3)\right.$$
$$\left. - \frac{z^2\jmath u\gamma\tilde{\delta}^3}{(1-z^2\gamma\tilde{\gamma})^2}\frac{1}{s}\operatorname{tr}(\mathbf{M}\mathbf{\Xi}^2) - \frac{\jmath u\tilde{\delta}}{1-z^2\gamma\tilde{\gamma}}\frac{1}{s}\operatorname{tr}\mathbf{M}\mathbf{\Xi}^2\right]\times\mathbb{E}\left[e^{\jmath u I_s(z)+\jmath u g(\mathbf{Z})}\right] + \mathcal{O}\left(s^{-1}\right). \tag{51}$$

Hence $\Psi_s(u,z)$ satisfies:

$$\frac{\partial\Psi_s}{\partial z} = \left[\frac{-u^2z^3\tilde{\gamma}^2}{(1-z^2\gamma\tilde{\gamma})^2}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}^3\mathbf{\Xi}^3\right)\frac{1}{s}\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right) - \frac{u^2z\tilde{\gamma}}{(1-z^2\gamma\tilde{\gamma})}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}^2\mathbf{\Xi}^3\right)\right.$$
$$\left.+\frac{u^2z^2\gamma\tilde{\delta}^3}{(1-z^2\gamma\tilde{\gamma})^2}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right) + \frac{u^2\tilde{\delta}}{1-z^2\gamma\tilde{\gamma}}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right)\right]\Psi_s(u,z) + \mathcal{O}\left(s^{-1}\right).$$

Following the same lines as in [11], one can prove that:

$$-\frac{d\log\left(1-z^2\gamma\tilde{\gamma}\right)}{dz} = \frac{1}{1-z^2\gamma\tilde{\gamma}}\left(-\frac{z^2\gamma\tilde{\delta}^3\frac{1}{s}\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right)}{1-z^2\gamma\tilde{\gamma}} + z\tilde{\gamma}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}^2\mathbf{\Xi}^3\right) + \frac{z^3\tilde{\gamma}^2\frac{1}{s}\operatorname{tr}\left(\mathbf{M}^3\mathbf{\Xi}^3\right)\frac{1}{s}\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right)}{1-z^2\gamma\tilde{\gamma}}\right). \tag{52}$$

Moreover, from the system of equations (54) in [11], one can find that:

$$\frac{1}{2}\frac{d\log\tilde{\gamma}}{dz} = -\frac{\tilde{\delta}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}\mathbf{\Xi}^2\right)}{1-z^2\gamma\tilde{\gamma}}. \tag{53}$$

Using (52) and (53), we finally get:

$$\frac{\partial\Psi_s}{\partial z} = -\frac{u^2}{2}\left[-\frac{d}{dz}\log(1-z^2\gamma\tilde{\gamma}) + \frac{d\log\tilde{\gamma}}{dz}\right]\Psi_s(u,z) + \mathcal{O}\left(s^{-1}\right).$$

Let $\sigma_T^2 = -\log\left(1-z^2\gamma\tilde{\gamma}\right) + \log\tilde{\gamma}$ and $K_s(u,z) = \Psi_s(u,z)\exp\left(\frac{u^2\sigma_T^2}{2}\right)$. Therefore, $K_s(u,z)$ satisfies:

$$\frac{\partial K_s}{\partial z} = \epsilon(s,z)\exp\left(\frac{u^2\sigma_T^2}{2}\right),$$

where $\epsilon(s,z) = \mathcal{O}\left(s^{-1}\right)$. On the other hand, we have:

$$K_s(u,z) = \mathbb{E}\left[e^{\jmath u\left(-\log\det\left(\frac{1}{s}\mathbf{Z}\mathbf{Z}^{\mathsf{H}} - b_s\right)\right)}\right].$$

Hence,

$$K_s(u,z) = K_s(u,0) + \int_0^z \epsilon_s(u,x)dx$$
$$= e^{\frac{u^2\log(1-\frac{r}{s})}{2}} + \mathcal{O}\left(s^{-1}\right).$$

The characteristic function $\Psi_s(u,z)$ can be thus approximated as:

$$\Psi_s(u,z) = \exp\left(-\frac{u^2\sigma_T^2}{2} + \frac{u^2\log(1-\frac{r}{s})}{2}\right) + \mathcal{O}\left(s^{-1}\right). \tag{54}$$

The characteristic function satisfies the same equation as in [11]. The single difference is that the variance $\alpha_{N,t}(y)$ given by:

$$\alpha_{N,t}(y) = -\log\left(\frac{1-\gamma\tilde{\gamma}}{\tilde{\gamma}}\right) - \log(1-\frac{r}{s}) \tag{55}$$

has two additive terms accounting for the variance of $g(\mathbf{Z})$ and the correlation between $g(\mathbf{Z})$ and $I_s(z)$. The CLT can be thus established by using the same arguments in [11], provided that we show that $\liminf \alpha_{N,t}(y) > 0$. For that, we need only to prove that:

$$\liminf \frac{1-z^2\gamma\tilde{\gamma}}{\tilde{\gamma}} > 0.$$

Deriving $\tilde{\delta}$ with respect to z, one can easily see that:

$$\frac{1-z^2\gamma\tilde{\gamma}}{\tilde{\gamma}} = -\frac{1}{\frac{d\tilde{\delta}}{dz}}\frac{1}{s}\operatorname{tr}\left(\mathbf{M}\boldsymbol{\Xi}^2\right).$$

It has been shown in [11, eq.(67)] that $-\frac{d\tilde{\delta}}{dz}$ satisfies:

$$0 < -\frac{d\tilde{\delta}}{dz} < \frac{r}{s}\lambda_{\max,t},$$

where $\lambda_{\max} = \max\left(\lambda_{1,t},\cdots,\lambda_{r,t}\right)$. This fact combined with $\liminf \frac{1}{s}\operatorname{tr}\left(\mathbf{M}\boldsymbol{\Xi}^2\right)$ implies that $\liminf \alpha_{N,t}(y) > 0$. It remains thus to express the variance $\alpha_{N,t}(y)$ using the original notations. One can easily show that:

$$\begin{aligned}
\delta &= \frac{1}{M-N+r}\operatorname{tr}\left(\frac{My}{M-N+r}\mathbf{D}_t^{-\frac{1}{2}}\mathbf{U}_t^{\mathsf{H}}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{U}_t\mathbf{D}_t^{-\frac{1}{2}} + \frac{\mathbf{I}_N}{1+\delta}\right)^{-1} - \frac{(N-r)(1+\delta)}{M-N+r} \\
&= \frac{1}{M-N+r}\operatorname{tr}\left(\left(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)\left(\frac{My}{M-N+r}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N}{1+\delta}\right)^{-1}\right) - \frac{(N-r)(1+\delta)}{M-N+r} \\
&= \frac{1}{M}\operatorname{tr}\left(\left(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{(M-N+r)(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N)}{M(1+\delta)}\right)^{-1}\right) - \frac{(N-r)(1+\delta)}{M-N+r}.
\end{aligned} \tag{56}$$

Then, from (56), we can prove that $\frac{M(\delta+1)}{M-N+r} - 1$ is solution in x of:

$$x = \frac{1}{M}\operatorname{tr}\left(\left(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}}{1+x}\right)^{-1}\right). \tag{57}$$

Since κ_t is the unique solution of (57), we have:

$$\frac{M(\delta+1)}{M-N+r} - 1 = \kappa_t,$$

or equivalently:

$$\tilde{\delta} = \frac{1}{1+\delta} = \frac{M}{(M-N+r)(\kappa_t+1)}.$$

Therefore:

$$\tilde{\gamma} = \tilde{\delta}^2 = \frac{M^2}{(M-N+r)^2(\kappa_t+1)^2}. \tag{58}$$

In the same way, one can prove that γ can be expressed in terms of the original notations as:

$$\gamma = \frac{(M-N+r)}{M^2}\,\mathrm{tr}\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\left(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}+\frac{\mathbf{I}_N}{\kappa_t+1}\right)^{-2}-\frac{(\kappa_t+1)^2(N-r)(M-N+r)}{M^2}. \quad (59)$$

Substituting (59) and (58) into (55), $\alpha_{N,t}(y)$ becomes

$$\alpha_{N,t}(y) = \log M^2 - \log\left((M-N)\left(M(\kappa_t+1)^2-\mathrm{tr}\left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\left(\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}}+\sigma^2\mathbf{I}_N\right)^{-1}+\frac{\mathbf{I}_N}{\kappa_t+1}\right)^{-2}\right)\right).$$

APPENDIX E
PROOF OF THEOREM 3

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

$$f(y) = \frac{1}{M}\,\mathrm{tr}(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(y)) + \frac{M-N}{M} - y$$
$$R(y) = -\log\det(\mathbf{Q}_t(y)) + (M-N)\log(y) - My.$$

where $\mathbf{Q}_t(y) = \left(y\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}} + \frac{1}{M}\mathbf{Y}_t\mathbf{Y}_t^{\mathsf{H}}\right)^{-1}$. According to Poincaré-Nash inequality, we have:

$$\mathrm{var}(\hat{y}_{N,t}) \le K\sum_{i=1}^{N}\sum_{j=1}^{M}\left[\mathbb{E}\left|\frac{\partial \hat{y}_{N,t}}{\partial Y_{i,j}^*}\right|^2 + \mathbb{E}\left|\frac{\partial \hat{y}_{N,t}}{\partial Y_{i,j}}\right|^2\right] . \quad (60)$$

We only deal with the first sum in the previous inequality; the second one can be handled similarly. By the implicit function theorem, if $\frac{\partial f}{\partial y} \neq 0$ then $\frac{\partial \hat{y}_{N,t}}{\partial Y_{i,j}^*}$ writes:

$$\frac{\partial \hat{y}_{N,t}}{\partial Y_{i,j}^*} = \frac{\frac{\partial f}{\partial Y_{i,j}^*}(\hat{y}_{N,t})}{\frac{\partial f}{\partial y}(\hat{y}_{N,t})} . \quad (61)$$

As will be shown later, to conclude that $\mathrm{var}(\hat{y}_{N,t}) = \mathcal{O}(M^{-2})$, we need to establish that $\left|\frac{\partial f}{\partial y}(\hat{y}_{N,t})\right|$ is 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

$$\frac{\partial^2 R}{\partial y^2}(\hat{y}_{N,t}) = \frac{M\frac{\partial f}{\partial y}(\hat{y}_{N,t})}{\hat{y}_{N,t}} . \quad (62)$$

On the other hand, one can prove by straightforward calculations that $\left|\frac{\partial^2 R}{\partial y^2}(\hat{y}_{N,t})\right| \ge \frac{M-N}{\hat{y}_{N,t}^2}$ which, plugged into (62), yields:

$$\left|\frac{\partial f}{\partial y}\right| \ge \frac{M-N}{M\hat{y}_{N,t}} , \quad (63)$$

which is eventually uniformily lower bounded away from 0 due to Assumption **A2** and to the fact that $\hat{y}_{N,t} \le 1$ by mere definition. Therefore,

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

To prove 2), we rely on the resolvent identity which states:

$$\mathbf{Q}_t(a) - \mathbf{Q}_t(b) = (b-a)\mathbf{Q}_t(a)\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(b)\ . \tag{64}$$

Using (64), we obtain:

$$\begin{aligned}
\hat{y}_{N,t} &= \frac{1}{M}(\hat{y}_{N,t} - \mathbb{E}\hat{y}_{N,t}) \operatorname{tr} \mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\hat{y}_{N,t}) + \frac{1}{M}\operatorname{tr} \mathbb{E}(\hat{y}_{N,t})\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\hat{y}_{N,t}) + \frac{M-N}{M}\ ,\\
&= \frac{1}{M}(\hat{y}_{N,t} - \mathbb{E}\hat{y}_{N,t})\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\mathbb{E}\hat{y}_{N,t}) - \frac{1}{M}\operatorname{tr}(\hat{y}_{N,t} - \mathbb{E}\hat{y}_{N,t})^2\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\hat{y}_{N,t})\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\mathbb{E}\hat{y}_{N,t})\\
&\quad + \frac{1}{M}\operatorname{tr}\mathbb{E}(\hat{y}_{N,t})\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\mathbb{E}\hat{y}_{N,t}) - \frac{1}{M}\operatorname{tr}\mathbb{E}(\hat{y}_{N,t})(\hat{y}_{N,t} - \mathbb{E}(\hat{y}_{N,t}))\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\hat{y}_{N,t})\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\mathbb{E}(\hat{y}_{N,t})) + \frac{M-N}{M}\ ,\\
&\overset{(a)}{=} \frac{1}{M}(\hat{y}_{N,t} - \mathbb{E}\hat{y}_{N,t}) \operatorname{tr} \mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{T}(\mathbb{E}(\hat{y}_{N,t})) + \frac{1}{M}\mathbb{E}(\hat{y}_{N,t})\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{T}(\mathbb{E}(\hat{y}_{N,t}))\\
&\quad - \mathbb{E}(\hat{y}_{N,t})(\hat{y}_{N,t} - \mathbb{E}\hat{y}_{N,t})\mathbb{E}\left[\frac{1}{M}\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\hat{y}_{N,t})\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\mathbb{E}(\hat{y}_{N,t}))\right] + \frac{M-N}{M} + \varepsilon\ ,
\end{aligned}$$

where ε satisfies $\mathbb{E}(\varepsilon) = \mathcal{O}(M^{-2})$. Note that equality (a) follows from the fact that

$$\operatorname{var}(\hat{y}_{N,t}) = \mathcal{O}\left(\frac{1}{M^2}\right) \quad \text{and} \quad \operatorname{var}\left(\frac{1}{M}\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\hat{y}_{N,t})\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{Q}_t(\mathbb{E}(\hat{y}_{N,t}))\right) = \mathcal{O}\left(\frac{1}{M^2}\right)\ .$$

Both estimates can be established with the help of Poincaré-Nash inequality. Therefore:

$$\begin{aligned}
\mathbb{E}(\hat{y}_{N,t}) &= \frac{1}{M}\mathbb{E}(\hat{y}_{N,t})\operatorname{tr}\mathbf{H}_t\mathbf{H}_t^{\mathsf{H}}\mathbf{T}_t(\mathbb{E}(\hat{y}_{N,t})) + \frac{M-N}{M} + \mathcal{O}(M^{-2})\\
&= 1 - \frac{1}{M(1+\kappa(\mathbb{E}(\hat{y}_{N,t}))}\operatorname{tr}((\mathbf{G}_t\mathbf{G}_t^{\mathsf{H}} + \sigma^2\mathbf{I}_N)\mathbf{T}_t(\mathbb{E}(\hat{y}_{N,t}))) + \mathcal{O}(M^{-2})\\
&= 1 - \frac{\kappa(\mathbb{E}(\hat{y}_{N,t}))}{1+\kappa(\mathbb{E}(\hat{y}_{N,t}))} + \mathcal{O}(M^{-2})\\
&= \frac{1}{1+\kappa(\mathbb{E}(\hat{y}_{N,t}))} + \mathcal{O}(M^{-2})\ .
\end{aligned} \tag{65}$$

Now recall the definition of $y_{N,t} = (1+\kappa(y_{N,t}))^{-1}$. One can prove easily that $y \mapsto \kappa(y)$ is a contraction, i.e. that there exists $k_N < 1$ such that

$$|\kappa(y_1) - \kappa(y_2)| \le k_N|y_1 - y_2|, \quad \forall y_1, y_2 > 0\ ,$$

and that $\limsup_{N,n} k_N < 1$. Using the mere definition of $y_{N,t}$ and (65), we obtain:

$$\begin{aligned}
\mathbb{E}(\hat{y}_{N,t}) - y_{N,t} &= \frac{1}{1+\kappa(\mathbb{E}(\hat{y}_{N,t}))} - \frac{1}{1+\kappa(y_{N,t})} + \mathcal{O}(M^{-2})\ ,\\
&= \frac{\kappa(y_{N,t}) - \kappa(\mathbb{E}(y_{N,t}))}{(1+\kappa(\mathbb{E}(y_{N,t})))(1+\kappa(y_{N,t}))} + \mathcal{O}(M^{-2})\ .
\end{aligned}$$

Hence,

$$|\mathbb{E}(\hat{y}_{N,t}) - y_{N,t}| \le k_N\,|\mathbb{E}(\hat{y}_{N,t}) - y_{N,t}| + \mathcal{O}(M^{-2})\ ,$$

thus proving that $|\mathbb{E}(\hat{y}_{N,t}) - y_{N,t}| = \mathcal{O}(M^{-2})$, which concludes the proof.

REFERENCES

[1] X. Mestre, "On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5353–5368, Nov. 2008.

[2] P. Vallet, P. Loubaton, and X. Mestre, "Improved subspace estimation for multivariate observations of high dimension: the deterministic signals case," *IEEE Trans. Inf. Theory*, 2010, submitted for publication. [Online]. Available: http://arxiv.org/abs/1002.3234

[3] X. Mestre and M. Lagunas, "Modified Subspace Algorithms for DoA Estimation With Large Arrays," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 598–614, Feb. 2008.

[4] Ø. Ryan and M. Debbah, "Channel Capacity Estimation using Free Probability Theory," *IEEE Trans. Signal Process.*, vol. 56, no. 11, Nov. 2008.

[5] P. Vallet and P. Loubaton, "A G-estimator of the MIMO channel ergodic capacity," in *Proc. IEEE International Symposium on Information Theory (ISIT'09)*, 2009, pp. 774–778.

[6] V. L. Girko, "Ten years of general statistical analysis," unpublished. [Online]. Available: http://www.general-statistical-analysis.girko.freewebspace.com/chapter14.pdf

[7] ——, *An Introduction to Statistical Analysis of Random Arrays*. VSP, 1998.

[8] V. A. Marčenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math USSR-Sbornik*, vol. 1, no. 4, pp. 457–483, Apr. 1967.

[9] 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.

[10] L. A. Pastur, "A simple approach to global regime of random matrix theory," in *Mathematical results in statistical mechanics*. World Scientific Publishing, 1999, pp. 429–454.

[11] W. Hachem, O. Khorunzhy, P. Loubaton, J. Najim, and L. A. Pastur, "A new approach for capacity analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, 2008.

[12] D. A. Harville, *Matrix algebra from a statistician's perspective*. New York: Springer-Verlag, 1997.

[13] A. W. Van der Vaart, *Asymptotic statistics*. New York: Cambridge University Press, 2000.

[14] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[15] J. W. Silverstein and S. Choi, "Analysis of the limiting spectral distribution of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 295–309, 1995.

[16] R. J. Murihead, *Aspects of Multivariate Statistical Theory*. New York: Wiley Series in Probability and Mathematical Statistics, 1982.

[17] 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.