A subspace estimator for fixed rank perturbations of large random matrices

Walid Hachem [a,*], Philippe Loubaton [b], Xavier Mestre [c], Jamal Najim [a], Pascal Vallet [b]

[a] CNRS; Télécom Paristech, 46, rue Barrault, 75013 Paris, France

[b] IGM LabInfo, UMR 8049, Institut Gaspard Monge, Université Paris Est Marne-la-Vallée, 5, Bd Descartes, Champs sur Marne, 77454 Marne La Vallée Cedex 2, France

[c] Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Parc Mediterrani de la Tecnologia - Building B4, Av. Carl Friedrich Gauss 7, 08860 - Castelldefels Barcelona, Spain

ARTICLE INFO

ABSTRACT

This paper deals with the problem of parameter estimation based on certain eigenspaces of the empirical covariance matrix of an observed multidimensional time series, in the case where the time series dimension and the observation window grow to infinity at the same pace. In the area of large random matrix theory, recent contributions studied the behavior of the extreme eigenvalues of a random matrix and their associated eigenspaces when this matrix is subject to a fixed-rank perturbation. The present work is concerned with the situation where the parameters to be estimated determine the eigenspace structure of a certain fixed-rank perturbation of the empirical covariance matrix. An estimation algorithm in the spirit of the well-known MUSIC algorithm for parameter estimation is developed. It relies on an approach recently developed by Benaych-Georges and Nadakuditi (2011) [8,9], relating the eigenspaces of extreme eigenvalues of the empirical covariance matrix with eigenspaces of the perturbation matrix. First and second order analyses of the new algorithm are performed.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Parameter estimation algorithms based on the estimation of an eigenspace of the autocorrelation matrix of an observed multivariate time series are very popular in the areas of statistics and signal processing. Applications of such algorithms include the estimation of the angles of arrival of plane waves impinging on an array of antennas, the estimation of the frequencies of superimposed sine waves, or the resolution of multiple paths of a radio signal. Denoting by N the signal dimension (*e.g.*, the number of antennas) and by n the length of the time observation window, the observed time series is represented by a $N \times n$ random matrix $\Sigma_n = X_n + P_n$ where X_n and P_n are respectively the so-called noise and signal matrices. In many applications, P_n is represented as

$$P_n = B(\varphi_1, \ldots, \varphi_r)S_n^*, \tag{1}$$

where $(\varphi_1, \ldots, \varphi_r)$ are the $r \leq \min(N, n)$ deterministic parameters to be estimated, B is a $N \times r$ matrix of the form $B(\varphi_1, \ldots, \varphi_r) = \begin{bmatrix} b(\varphi_1) & \cdots & b(\varphi_r) \end{bmatrix}$ where $b(\varphi)$ is a known $\mathbb{C}^N$-valued function of φ, and the S_n is an unknown $n \times r$ matrix with rank r representing the signals transmitted by the r emitting sources. As usual (and unless stated otherwise),

---

* Corresponding author.
E-mail addresses: walid.hachem@telecom-paristech.fr (W. Hachem), loubaton@univ-mlv.fr (P. Loubaton), xavier.mestre@cttc.cat (X. Mestre), jamal.najim@telecom-paristech.fr (J. Najim), vallet@univ-mlv.fr (P. Vallet).

A^* stands for the Hermitian adjoint of matrix A. It will be assumed in this work that this matrix is deterministic. Often, the noise matrix X_n is a complex random matrix such that the real and imaginary parts of its elements are $2Nn$ independent random variables with common probability law $\mathcal{N}(0, 1/(2n))$. In this case, we shall say that $\sqrt{n}X_n$ is a standard Gaussian matrix.

We shall consider here "direction of arrival" vector functions $b(\varphi)$ that are typically met in the field of antenna processing. These functions are written as

$$b(\varphi) = N^{-1/2} \left[\exp(-\imath D\ell\varphi) \right]_{\ell=0}^{N-1}$$

with domain $\varphi \in [0, \pi/D]$ where D is a positive real constant and $\imath^2 = -1$. Assuming that the angular parameters φ_k are all different, the well-known MUSIC (MUltiple SIgnal Classification, [27,11]) algorithm for estimating these parameters from Σ_n relies on the following simple idea. Assume that $\sqrt{n}X_n$ is standard Gaussian and let Π be the orthogonal projection matrix on the eigenspace of $\mathbb{E}\Sigma_n\Sigma_n^* = BS_n^*S_nB^* + I_N$ associated with the r largest eigenvalues, where I_N is the $N \times N$ identity matrix. Obviously, Π is the orthogonal projector on the column space of $B(\varphi_1, \ldots, \varphi_r)$. As a consequence, the angles φ_k coincide with the zeros of the function $b(\varphi)^*(I - \Pi)b(\varphi)$ on $[0, \pi/D]$. Since $\|b(\varphi)\| = 1$, they equivalently coincide with the maximum values (at one) of the so-called localization function $\chi(\varphi) = b(\varphi)^*\Pi b(\varphi)$.

In practice, Π is classically replaced with the orthogonal projection matrix $\widehat{\Pi}$ on the eigenspace associated with the r largest eigenvalues of $\Sigma_n\Sigma_n^*$. Assuming N is fixed and $n \to \infty$, and assuming furthermore that $S_n^*S_n$ converges to some matrix $O > 0$ in this asymptotic regime, the $\Sigma\Sigma^* \xrightarrow{\text{a.s.}} BOB^* + I_N$ by the Law of Large Numbers (a.s. stands for almost surely). Hence, the random variable $\chi_{\text{classical}}(\varphi) = b(\varphi)^*\widehat{\Pi}b(\varphi)$ a.s. converges to $\chi(\varphi)$, and it is standard to estimate the arrival angles as local maxima of $\chi_{\text{classical}}(\varphi)$.

However, in many practical situations, the signal dimension N and the window length n are of the same order of magnitude in which case the spectral norm of $\widehat{\Pi} - \Pi$ is not small, as we shall see below. In these situations, it is often more relevant to assume that both N and n converge to infinity at the same pace, while the number of parameters r is kept fixed. The subject of this paper is to develop a new estimator better suited to this asymptotic regime, and to study its first and second order behavior with the help of large random matrix theory.

In large random matrix theory, much has been said about the spectral behavior of $X_nX_n^*$ in this asymptotic regime, for a wide range of statistical models for X_n. In particular, it is frequent that the spectral measure of this matrix converges to a compactly supported limiting probability measure π, and that the extreme eigenvalues of $X_nX_n^*$ a.s. converge to the edges of this support. Considering that Σ_n is the sum of X_n and a fixed-rank perturbation, it is well-known that $\Sigma_n\Sigma_n^*$ also has the limiting spectral measure π [1, Lemma 2.2]. However, the largest eigenvalues of $\Sigma_n\Sigma_n^*$ have a special behavior. Under some conditions, these eigenvalues leave the support of π, and in this case, their related eigenspaces give valuable information on the eigenspaces of P_n. This paper shows how the angles φ_k can be estimated from these eigenspaces.

The problem of the behavior of the extreme eigenvalues of large random matrices subjected to additive or multiplicative low rank perturbations (often called "spiked models") has received a great deal of interest in the recent years. In this regard, the authors of [4,5,25] study the behavior of the extreme eigenvalues of a sample covariance matrix when the population covariance matrix has all but finitely many eigenvalues equal to one, a problem described in [20]. Ref. [13] is devoted to the extreme eigenvalues of a Wigner matrix that incurs a fixed-rank additive perturbation. Fluctuations of these eigenvalues are studied in [4,26,25,3,13,12,6].

Recently, Benaych-Georges and Nadakuditi have proposed in [8,9] a powerful technique for characterizing the behavior of extreme eigenvalues and their associated eigenspaces for three generic spiked models. The models $X_n + P_n$ and $(I_n + P_n)X_n$ when both X_n and P_n are Hermitian and P_n is low-rank, and the model that encompasses ours $(X_n + P_n)(X_n + P_n)^*$ where X_n and P_n are rectangular. One feature of this approach is that it uncovers simple relations between the extreme eigenvalues and their associated eigenspaces on the one hand, and certain quadratic forms involving resolvents related with the non-perturbed matrix X_n on the other hand. This makes the method particularly well-suited (but not limited to) the situation where X_n is unitarily or bi-unitarily invariant, a situation that we shall consider in this paper. Indeed, in this situation, these quadratic forms exhibit a particularly simple behavior in the considered large dimensional asymptotic regime.

In this paper, we make use of the approach of [8,9] to develop a new subspace estimator of the angles φ_k based on the eigenspaces of the isolated eigenvalues of $\Sigma_n\Sigma_n^*$. We perform the first and second order analyses of this estimator that we call the "Spike MUSIC" estimator. Our mathematical developments differ somehow from those of [8,9] and could have their own interest. They are based on two simple ingredients: the first is an analogue of the Poincaré–Nash inequality for the Haar distributed unitary matrices which has been recently discovered by Pastur and Vasilchuk [24], and the second is a contour integration method by means of which the first and second order analyses are done. The key step of the second order analysis of our estimator lies in the establishment of a Central Limit Theorem on the quadratic forms $b(\varphi_i)^*\widehat{\Pi}_i b(\varphi_i)$ where the $\widehat{\Pi}_i$ are the orthogonal projection matrices on certain eigenspaces of $\Sigma_n\Sigma_n^*$ associated with the isolated eigenvalues. The employed technique can easily be used to study the fluctuations of projections of other types of vectors on these eigenspaces.

We now state our general assumptions and introduce some notations.

Assumptions and notations

We now state the general assumptions of the paper. Consider the sequence of $N \times n$ matrices $\Sigma_n = X_n + P_n$ where the following assumption holds.

Assumption A1. The dimensions N, n satisfy: $N \leq n$, $n \to \infty$ and

$$\frac{N}{n} \to c \in (0, 1]$$

(notation for this asymptotic regime: $n \to \infty$).

The following assumption on X_n is widely used in the random matrix literature [18,23].

Assumption A2. Matrices X_n are random $N \times n$ bi-unitarily invariant matrices, *i.e.*, each X_n admits the singular value decomposition $X_n = L_n \Gamma_n R_n^*$ where L_n, the $N \times N$ matrix Γ_n and R_n are independent, L_n is Haar distributed on the group $\mathcal{U}(N)$ of unitary $N \times N$ matrices, and R_n is a $n \times N$ submatrix of a Haar distributed matrix on $\mathcal{U}(n)$.

We recall that the Stieltjes transform of a probability measure π on the real line is the complex function

$$m(z) = \int \frac{1}{t - z} \pi(dt),$$

analytic on $\mathbb{C}_+ = \{z : \Im(z) > 0\}$.

Assumption A3. Let $Q_n(z) = (X_n X_n^* - zI_N)^{-1}$ be the resolvent associated with $X_n X_n^*$ and let $\alpha_n(z) = N^{-1}\mathrm{tr}Q_n(z)$. For every $z \in \mathbb{C}_+$, $\alpha_n(z)$ a.s. converges to a deterministic function $m(z)$ which is the Stieltjes transform of a probability measure π supported by the compact interval $[\lambda_-, \lambda_+]$.

Assumption A4. The quantity $\|X_n X_n^*\|$ a.s. converges to λ_+ as $n \to \infty$, where $\|\cdot\|$ denotes the spectral norm.

Let $\widetilde{Q}_n(z) = (X_n^* X_n - zI_n)^{-1}$ and $\tilde{\alpha}_n(z) = n^{-1}\mathrm{tr}\widetilde{Q}_n(z)$. Equivalently to the convergence assumed by Assumption A3, one may assume that $\tilde{\alpha}_n(z)$ a.s. converges on $\mathbb{C}_+$ to a deterministic function $\tilde{m}(z)$ which is the Stieltjes transform of a probability measure $\tilde{\pi}$. In that case, $\tilde{m}(z) = cm(z) - (1 - c)/z$ and $\tilde{\pi} = c\pi + (1 - c)\delta_0$.

Remark 1. In the areas of signal processing and communication theory, the noise matrix X_n satisfying Assumptions A2–A4 is such that $\sqrt{n}X_n$ is standard Gaussian—see for instance [21,15].

We first make a general assumption on matrices P_n; it will be specified later, and adapted to the context of the MUSIC algorithm.

Assumption A5. Matrices P_n are deterministic with a fixed rank equal to r for all n large enough. Denoting by $P_n = U_n \Omega_n V_n^*$ a singular value decomposition of P_n, the matrix of singular values $\Omega_n = \mathrm{diag}(\omega_{1,n}, \dots, \omega_{r,n})$ with $\omega_{1,n} \geq \omega_{2,n} \geq \cdots \geq \omega_{r,n}$ converges to

$$O = \begin{bmatrix} \omega_1 I_{j_1} & & \\ & \ddots & \\ & & \omega_s I_{j_s} \end{bmatrix}, \tag{2}$$

where $\omega_1 > \cdots > \omega_s > 0$ and $j_1 + \cdots + j_s = r$.

Notations

As usual, if $z \in \mathbb{C}$, we shall denote by $\Re(z)$ and $\Im(z)$ its real and imaginary parts. We shall denote by $\xrightarrow{\text{a.s.}}$ (resp. $\xrightarrow{\mathcal{P}}$, $\xrightarrow{\mathcal{D}}$) the almost sure convergence (resp. convergence in probability, in distribution). We denote by $\delta_{i,j}$ the Kronecker delta ($= 1$ if $i = j$ and 0 otherwise).

The eigenvalues of $\Sigma_n \Sigma_n^*$ are $\hat{\lambda}_{1,n} \geq \hat{\lambda}_{2,n} \geq \cdots \geq \hat{\lambda}_{N,n}$. Associated eigenvectors will be denoted by $\hat{u}_{1,n}, \hat{u}_{2,n}, \dots, \hat{u}_{N,n}$. For $k \in \{1, \dots, r\}$, we shall denote by $i(k)$ the index $i \in \{1, \dots, s\}$ such that $j_1 + \cdots + j_{i-1} < k \leq j_1 + \cdots + j_i$. For $i = 1, \dots, s$, we shall denote by $\widehat{\Pi}_{i,n}$ the orthogonal projection matrix on the eigenspace of $\Sigma_n \Sigma_n^*$ associated with the eigenvalues $\hat{\lambda}_{k,n}$ such that $i(k) = i$, *i.e.*, $\widehat{\Pi}_{i,n} = \sum_{k:i(k)=i} \hat{u}_{k,n}\hat{u}_{k,n}^*$ when this eigenspace is defined. Columns of U_n (see A5) will be denoted by $u_{1,n}, \dots, u_{r,n}$. Given i, the orthogonal projection matrix on the eigenspace of $P_n P_n^*$ associated with the eigenvalues $\omega_{k,n}^2$ such that $i(k) = i$ will be $\Pi_{i,n} = \sum_{k:i(k)=i} u_{k,n}u_{k,n}^*$. Indices n and N will often be dropped for readability.

Paper organization

The paper is organized as follows. Section 2 is devoted to the mathematical preliminaries. The general approach is described in Section 3. The Spike MUSIC algorithm is presented in Section 4 along with a first order study of this algorithm. Fluctuations of the estimates of the φ_k are studied in Section 5 under the form of a Central Limit Theorem.

2. Preliminary mathematical results

We shall need the two following results. The first one is well-known [24]. The second result, due to Pastur and Vasilchuk, is the unitary analogue of the well-known Poincaré–Nash inequality.

Lemma 1. *Let* $W = [w_{ij}]$ *be a random matrix Haar distributed on* $\mathcal{U}(n)$. *Then*

$$\mathbb{E}\left[w_{ij}w_{i'j'}^*\right] = \frac{1}{n}\delta_{i,i'}\delta_{j,j'}.$$

Lemma 2 (*[24,23]*). *Let* $\Phi : \mathcal{U}(n) \to \mathbb{C}$ *be a function that admits a* C^1 *continuation to an open neighborhood of* $\mathcal{U}(n)$ *in the whole algebra of* $n \times n$ *complex matrices. Then*

$$\operatorname{var}\Phi(W_n) = \mathbb{E}|\Phi(W_n)|^2 - |\mathbb{E}\Phi(W_n)|^2 \leq \frac{1}{n}\sum_{j,k=1}^n \mathbb{E}\left|\Phi'(W_n) \cdot \left(\mathbf{e}_j\mathbf{e}_k^\mathsf{T}W_n\right)\right|^2$$

where $\mathbb{E}$ *is the expectation with respect to the Haar measure on* $\mathcal{U}(n)$, *where* Φ' *is the differential of* Φ *as a function on* $\mathbb{R}^{2n^2}$ *acting on the matrix* $\mathbf{e}_j\mathbf{e}_k^\mathsf{T}W_n$ *seen as an element of* $\mathbb{R}^{2n^2}$, *and where* $\mathbf{e}_j = [0\cdots0\,1\,0\cdots0]^*$ *is the jth canonical vector of* $\mathbb{C}^n$.

Given a small $\varepsilon_1 > 0$, let O_n be the probability event

$$O_n = \left\{\|X_nX_n^*\| \leq \lambda_+ + \varepsilon_1\right\}. \tag{3}$$

By Assumption A4, $\mathbb{1}_{O_n} \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$.

Lemma 3. *Let Assumption* A2 *hold true and let* u, v *be two unit norm deterministic* $N \times 1$ *vectors such that* $u^*v = 0$. *Then for any* z *with* $\Re(z) > \lambda_+ + \varepsilon_1$,

$$\mathbb{E}\left|\mathbb{1}_{O_n} \times u^*\left(Q(z) - \alpha(z)I\right)u\right|^p \leq \frac{K_p}{N^{p/2}d(z, \lambda_+ + \varepsilon_1)^p},$$

$$\mathbb{E}\left|\mathbb{1}_{O_n} \times u^*Q(z)v\right|^p \leq \frac{K_p}{N^{p/2}d(z, \lambda_+ + \varepsilon_1)^p},$$

where the constant K_p *only depends on* p, *and where* $d(z, z')$ *is the Euclidean distance between* z *and* z' *in* $\mathbb{C}$.

Proof. Recall that $X = L\Gamma R^*$ by Assumption A2; let $D = (\Gamma^2 - zI)^{-1}$; write:

$$\begin{bmatrix} u^* \\ v^* \end{bmatrix}(Q - \alpha I)\begin{bmatrix} u & v \end{bmatrix} = \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix}\left(D - \frac{\operatorname{tr}D}{N}I\right)\begin{bmatrix} w_1 & w_2 \end{bmatrix}.$$

Thanks to A2, w_1 and w_2 are the first two columns of a $N \times N$ unitary Haar distributed matrix $W = [w_{ij}]$ independent of D. Let $M = \mathbb{1}_{O_n} \times \left(D - N^{-1}(\operatorname{tr}D)I\right)$ and $\Phi_i(W) = w_1^*Mw_i$ for $i = 1, 2$. Then $\mathbb{E}\Phi_1(W) = \mathbb{E}\Phi_2(W) = 0$ by Lemma 1. Applying Lemma 2 to Φ_i after noticing that $\Phi_i'(W) \cdot A = \mathbf{e}_1^\mathsf{T}A^*Mw_i + w_1^*MA\mathbf{e}_i$ for any $N \times N$ matrix A, we obtain:

$$\mathbb{E}|\Phi_i|^2 = \operatorname{var}(\Phi_i) \leq \frac{1}{N}\sum_{j,k=1}^N \mathbb{E}\left|w_{k1}^*[MW]_{ji} + [W^*M]_{1j}w_{ki}\right|^2,$$

$$\leq \frac{2}{N}\mathbb{E}\left(\|Mw_i\|^2 + \|Mw_1\|^2\right),$$

$$\leq \frac{8}{Nd(z, \lambda_+ + \varepsilon_1)^2}.$$

We now proceed by induction; assume that the result is true until $p \geq 1$. Applying Lemma 2 to $\Phi_i^{(p+1)/2}$, we obtain:

$$\operatorname{var}\left(\Phi_i^{\frac{p+1}{2}}\right) \leq \frac{1}{N}\sum_{j,k=1}^N \mathbb{E}\left|\frac{p+1}{2}\Phi_i^{\frac{p-1}{2}}\Phi_i'(W) \cdot \left(\mathbf{e}_j\mathbf{e}_k^\mathsf{T}W\right)\right|^2,$$

$$\leq \frac{(p+1)^2}{2N}\mathbb{E}\left(|\Phi_i|^{p-1}\left(\|Mw_i\|^2 + \|Mw_1\|^2\right)\right),$$

$$\leq \frac{2(p+1)^2K_{p-1}}{d(z, \lambda_+ + \varepsilon_1)^{p+1}N^{(p+1)/2}}.$$

Using again the induction hypothesis, we get:

$$
\mathbb{E}\,|\Phi_i|^{p+1} = \mathrm{var}\!\left(\Phi_i^{\frac{p+1}{2}}\right) + \left|\mathbb{E}\Phi_i^{\frac{p+1}{2}}\right|^2
$$

$$
\leq \frac{2(p+1)^2 K_{p-1} + K_{(p+1)/2}^2}{d(z, \lambda_+ + \varepsilon_1)^{p+1} N^{(p+1)/2}} = \frac{K_{p+1}}{d(z, \lambda_+ + \varepsilon_1)^{p+1} N^{(p+1)/2}},
$$

which concludes the proof. □

Lemma 4. *Let Assumption A2 hold true; let u, v be two unit norm deterministic vectors with respective dimensions $N \times 1$ and $n \times 1$. Then for any z such as $\Re(z) > \lambda_+ + \varepsilon_1$,*

$$
\mathbb{E}\left|\mathbb{1}_{O_n} \times u^* X \widetilde{Q}(z)v\right|^p \leq \frac{K_p}{n^{p/2} d(z, \lambda_+ + \varepsilon_1)^p}.
$$

Proof. Let $C = \Gamma(\Gamma^2 - zI)^{-1}$. By Assumption A2, $u^* X \widetilde{Q}(z)v = w^* C \tilde{w} = \Phi(w)$ where w is a vector uniformly distributed on the unit sphere of $\mathbb{C}^N$, $\tilde{w}$ is a vector uniformly distributed on the unit sphere of $\mathbb{C}^n$ and truncated to its first N elements, and w, $\tilde{w}$ and C are independent. The lemma is proved as above by applying Lemma 2 to Φ and by taking the expectation with respect to the law of w. □

Lemma 5. *Let Assumptions A1–A4 hold true. Let $\mathcal{C}$ be a closed path of $\mathbb{C}$ such that $\min_{z \in \mathcal{C}} \Re(z) > \lambda_+$. Fix the integer $r \leq N$ and let U_n and V_n be two deterministic isometry matrices with dimensions $N \times r$ and $n \times r$ respectively. Then*

$$
\sup_{z \in \mathcal{C}} \|U_n^* (Q_n(z) - m(z)I_N) U_n\| \xrightarrow[n \to \infty]{\text{a.s.}} 0,
$$

$$
\sup_{z \in \mathcal{C}} \|V_n^* (\widetilde{Q}_n(z) - \tilde{m}(z)I_n) V_n\| \xrightarrow[n \to \infty]{\text{a.s.}} 0,
$$

$$
\sup_{z \in \mathcal{C}} \|U_n^* X_n \widetilde{Q}_n(z) V_n\| \xrightarrow[n \to \infty]{\text{a.s.}} 0.
$$

Proof. Recall the definition (3) of the set O_n and assume that ε_1 is chosen such that $\min_{z \in \mathcal{C}} \Re(z) > \lambda_+ + \varepsilon_1$; let

$$
h_n(z) = \mathbb{1}_{O_n} \times U_n^* (Q_n(z) - \alpha_n(z)I_N) U_n.
$$

For any $\ell, s \leq r$, $[h_n]_{\ell,s}$ is a holomorphic function on $\mathbb{C} - [0, \lambda_+ + \varepsilon_1]$. Consider a denumerable sequence of points (z_k) in $\mathbb{C} - [0, \lambda_+ + \varepsilon_1]$ with an accumulation point in that set. By Lemma 3 with $p = 3$, the Markov inequality and Borel–Cantelli's lemma, there exists a probability one set on which $[h_n(z_k)]_{\ell,s} \to 0$ for every k. Moreover, the $|[h_n(z_k)]_{\ell,s}|$ are uniformly bounded on any compact set of $\mathbb{C} - [0, \lambda_+ + \varepsilon_1]$. By the normal family theorem, every n-sequence of $[h_n]_{\ell,s}$ contains a further subsequence which converges uniformly on the compact set $\mathcal{C} \subset \mathbb{C} - [0, \lambda_+ + \varepsilon_1]$ to a holomorphic function that we denote h^*. Since $h^*(z_k) = 0$ for all k, $h^*(z) = 0$ on $\mathcal{C}$, then $|[h_n(z)]_{\ell,s}|$ converges uniformly to zero on $\mathcal{C}$ with probability one, and thanks to Assumption A4, $\|U^* (Q(z) - \alpha(z)I) U\| \to 0$ uniformly on $\mathcal{C}$ with probability one. The same argument, used in conjunction with Assumption A3, shows that with probability one, $\alpha(z) - m(z) \to 0$ uniformly on $\mathcal{C}$, and the first assertion is proven. The second and third assertions are proven similarly, the third being obtained with the help of Lemma 4. □

3. Fixed rank perturbations: first order behavior

We first recall a result on matrix analysis that can be found in [19, Theorem 7.3.7].

Lemma 6. *Given a $N \times n$ matrix A with $N \leq n$, let $\mathbf{A}$ be the matrix:*

$$
\mathbf{A} = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}.
$$

Then $\sigma_1, \ldots, \sigma_N$ are the singular values of A if and only if $\sigma_1, \ldots, \sigma_N, -\sigma_1, \ldots, -\sigma_N$ in addition to $n - N$ zeros are the eigenvalues of $\mathbf{A}$. Furthermore, a pair (u, v) of unit norm vectors is a pair of (left, right) singular vectors of A associated with the singular value σ if and only if $\begin{bmatrix} u/\sqrt{2} \\ v/\sqrt{2} \end{bmatrix}$ is a unit norm eigenvector of $\mathbf{A}$ associated with the eigenvalue σ.

Along the ideas in [8,9], we now characterize the behavior of the largest eigenvalues of $\Sigma \Sigma^*$, and then focus on their eigenspaces.

Asymptotic behavior of the largest eigenvalues of $\Sigma \Sigma^*$

We start with an informal description of the approach. By Lemma 6, λ is an eigenvalue of $\Sigma \Sigma^*$ if and only if $\det(\boldsymbol{\Sigma} - \sqrt{\lambda}I) = 0$ where $\boldsymbol{\Sigma} = \begin{bmatrix} 0 & \Sigma \\ \Sigma^* & 0 \end{bmatrix}$. Writing:

$$\boldsymbol{\Sigma} = \begin{bmatrix} 0 & X \\ X^* & 0 \end{bmatrix} + \begin{bmatrix} U & 0 \\ 0 & V\Omega \end{bmatrix} \begin{bmatrix} 0 & I_r \\ I_r & 0 \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & \Omega V^* \end{bmatrix} \triangleq \mathbf{X} + BJB^*, \tag{4}$$

and assuming that $x > 0$ is not a singular value of X, we have:

$$\det(\boldsymbol{\Sigma} - xI) = \det(\mathbf{X} - xI + BJB^*) = \det(J)\det(\mathbf{X} - xI)\det(J + B^*(\mathbf{X} - xI)^{-1}B),$$

after noticing that $J = J^{-1}$. Using the formula for the inversion of a partitioned matrix (see [19])

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{12}^*)^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{12}^*A_{11}^{-1}A_{12})^{-1} \\ -(A_{22} - A_{12}^*A_{11}^{-1}A_{12})^{-1}A_{12}^*A_{11}^{-1} & (A_{22} - A_{12}^*A_{11}^{-1}A_{12})^{-1} \end{bmatrix},$$

we obtain:

$$\mathbf{Q}(x) = (\mathbf{X} - xI)^{-1} = \begin{bmatrix} -xI & X \\ X^* & -xI \end{bmatrix}^{-1} = \begin{bmatrix} xQ(x^2) & X\widetilde{Q}(x^2) \\ \widetilde{Q}(x^2)X^* & x\widetilde{Q}(x^2) \end{bmatrix}. \tag{5}$$

Therefore,

$$\det(\boldsymbol{\Sigma} - xI) = \det(J)\det(\mathbf{X} - xI)\det\widehat{H}(x),$$

where

$$\widehat{H}_n(x) = \begin{bmatrix} xU^*Q(x^2)U & I_r + U^*X\widetilde{Q}(x^2)V\Omega \\ I_r + \Omega V^*\widetilde{Q}(x^2)X^*U & x\Omega V^*\widetilde{Q}(x^2)V\Omega \end{bmatrix}$$

whence for n large enough, the isolated eigenvalues of $\Sigma \Sigma^*$ above λ_+ will coincide with the zeros of $\det\widehat{H}(\sqrt{x})$ that lie above λ_+. Under Assumptions A1–A5, Lemma 5 shows that $\widehat{H}(x)$ a.s. converges to

$$H(x) = \begin{bmatrix} xm(x^2)I_r & I_r \\ I_r & x\tilde{m}(x^2)O^2 \end{bmatrix}.$$

Consider the equation

$$\det H(\sqrt{x}) = \det\left(xm(x)\tilde{m}(x)O^2 - I_r\right) = 0,$$

and notice that the function

$$g(x) = xm(x)\tilde{m}(x) = x\left(\int \frac{1}{t-x}\pi(dt)\right)\left(c\int \frac{1}{t-x}\pi(dt) - \frac{1-c}{x}\right) \tag{6}$$

decreases from $g(\lambda_+^+) = \lim_{x\downarrow\lambda_+} g(x)$ to zero on (λ_+, ∞). Let $\omega_1^2 > \cdots > \omega_q^2$ be those among the diagonal elements of O^2 that satisfy $\omega_i^2 > 1/g(\lambda_+^+)$. Equation $g(x) = \omega_i^{-2}$ will have a unique solution $x = \rho_i > \lambda_+$ for any $i = 1, \ldots, q$, while it will have no solution larger than λ_+ for $i > q$. It is then expected that any eigenvalue $\hat{\lambda}_{k,n}$ of $\Sigma_n\Sigma_n^*$ for which $i(k) \leq q$ (remember the definition of $i(k)$ provided in the paragraph "Assumptions and Notations" in Section 1), will converge to ρ_i, while $\hat{\lambda}_{j_1+\cdots+j_q+1,n} \to \lambda_+$ almost surely.

These facts are formalized in the following theorem, shown in [7,9].

Theorem 1. *Let Assumptions A1–A5 hold true; let q be the maximum index such that $\omega_q^2 > 1/g(\lambda_+^+)$. Let ρ_i be the unique real number $>\lambda_+$ satisfying $\omega_i^2 g(\rho_i) = 1$ for $i = 1, \ldots, q$. Then*

$$\hat{\lambda}_{j_1+\cdots+j_{i-1}+\ell,n} \xrightarrow[n\to\infty]{\text{a.s.}} \rho_i$$

for $i = 1, \ldots, q$ and $\ell = 1, \ldots, j_i$ while

$$\hat{\lambda}_{j_1+\cdots+j_q+1,n} \xrightarrow[n\to\infty]{\text{a.s.}} \lambda_+.$$

In the case where $\sqrt{n}X$ is a standard Gaussian matrix, π is the Marčenko–Pastur distribution with support $\text{supp}(\pi) = [\lambda_-, \lambda_+] = [(1 - \sqrt{c})^2, (1 + \sqrt{c})^2]$, and

$$m(x) = \frac{1}{2cx}\left(1 - c - x + \sqrt{(1 - c - x)^2 - 4cx}\right) \tag{7}$$

for $x \in (\lambda_+, \infty)$. After a few derivations, we obtain the following.

Corollary 1. *Assume that $\sqrt{n}X$ is standard Gaussian. Let q be the maximum index such that $\omega_q^2 > \sqrt{c}$. Then*

$$\hat{\lambda}_{j_1+\cdots+j_{i-1}+\ell,n} \xrightarrow[n\to\infty]{\text{a.s.}} \frac{(\omega_i^2+1)(\omega_i^2+c)}{\omega_i^2} \quad \text{for } i = 1, \ldots, q,$$

and $\hat{\lambda}_{j_1+\cdots+j_q+1,n} \xrightarrow[n\to\infty]{\text{a.s.}} (1+\sqrt{c})^2.$

We now turn our attention to the eigenspaces of the isolated eigenvalues.

Asymptotic behavior of certain bilinear forms

Recall the definition of s as provided in Assumption A5. Given $i \le s$, assume that $\omega_i^2 > 1/g(\lambda_+^+)$. Given two $N \times 1$ deterministic sequences of vectors $b_{1,n}$ and $b_{2,n}$ with bounded norms, we shall find here a simple asymptotic relation between $b_{1,n}^* \widehat{\Pi}_{i,n} b_{2,n}$ and $b_{1,n}^* \Pi_{i,n} b_{2,n}$, that will be at the basis of the Spike MUSIC algorithm. A close problem has been considered in [9]. We consider here a different technique, based on a contour integration and on the use of Lemmas 3 and 4. This method lends itself easily to the first and second order analyses of the Spike MUSIC algorithm that we shall develop in the following sections.

Writing $\mathbf{b}_i = \begin{bmatrix} b_i \\ 0 \end{bmatrix}$ with $i = 1, 2$, we have by virtue of Lemma 6:

$$b_1^* \widehat{\Pi}_i b_2 = \frac{-1}{\imath\pi} \oint_{\mathcal{C}_{i,n}} \mathbf{b}_1^* (\Sigma - zI)^{-1} \mathbf{b}_2 \, dz,$$

where $\mathcal{C}_{i,n}$ is a positively oriented circle that encloses the only singular values $\sqrt{\hat{\lambda}_{k,n}}$ of Σ_n for which $i(k) = i$. Recalling (4) and using Woodbury's identity [19, Section 0.7.4] together with the fact that $J = J^{-1}$, we obtain:

$$b_1^* \widehat{\Pi}_i b_2 = \frac{-1}{\imath\pi} \oint_{\mathcal{C}_i} \mathbf{b}_1^* \mathbf{Q}(z)\mathbf{b}_2 \, dz + \frac{1}{\imath\pi} \oint_{\mathcal{C}_i} \mathbf{b}_1^* \mathbf{Q}(z)B \left(J + B^*\mathbf{Q}(z)B\right)^{-1} B^*\mathbf{Q}(z)\mathbf{b}_2 \, dz.$$

Using (5), we obtain after a straightforward calculation:

$$b_{1,n}^* \widehat{\Pi}_{i,n} b_{2,n} = \frac{-1}{\imath\pi} \oint_{\mathcal{C}_{i,n}} \mathbf{b}_{1,n}^* \mathbf{Q}_n(z)\mathbf{b}_{2,n} \, dz + \frac{1}{\imath\pi} \oint_{\mathcal{C}_{i,n}} \hat{a}_{1,n}^*(z)\widehat{H}_n(z)^{-1}\hat{a}_{2,n}(z) \, dz \tag{8}$$

where[1]

$$\hat{a}_{\ell,n}(z) = \begin{bmatrix} zU_n^*Q_n(z^2) \\ \Omega_n V_n^*\widetilde{Q}_n(z^2)X_n^* \end{bmatrix} b_{\ell,n},$$

$$\hat{a}_{\ell,n}^*(z) = b_{\ell,n}^* \begin{bmatrix} zQ_n(z^2)U_n & X_n\widetilde{Q}_n(z^2)V_n\Omega_n \end{bmatrix}. \tag{9}$$

Intuitively, the first integral is zero for n large enough and the second is close to

$$T_{i,n} = \frac{1}{\imath\pi} \oint_{\gamma_i} a_{1,n}^*(z)H(z)^{-1}a_{2,n}(z) \, dz,$$

where γ_i is a small enough positively oriented circle which does not meet the image of $\text{supp}(\pi)$ by $x \mapsto \sqrt{x}$ nor any of the $\sqrt{\rho_\ell}$ and such that only $\sqrt{\rho_i} \in \text{Int}(\gamma_i)$, the interior of the disk defined by γ_i (see Fig. 1), $a_{\ell,n}^*(z) = b_{\ell,n}^* \begin{bmatrix} zm(z^2)U_n & 0 \end{bmatrix}$, and

$$a_{\ell,n}(z) = \begin{bmatrix} zm(z^2)U_n^* \\ 0 \end{bmatrix} b_{\ell,n}.$$

The approximation $b_1^* \widehat{\Pi}_i b_2 \simeq T_i$ will be justified rigorously below. For the moment, let us develop the expression of T_i. Defining the $r \times r$ matrices:

$$\mathcal{I}_i = \begin{bmatrix} 0 & & \\ & I_{j_i} & \\ & & 0 \end{bmatrix},$$

where the integers j_i are defined in Assumption A5, we have

$$H(z)^{-1} = \sum_{i=1}^{s} \frac{1}{z^2 m(z^2)\tilde{m}(z^2)\omega_i^2 - 1} \begin{bmatrix} z\tilde{m}(z^2)\omega_i^2 & -1 \\ -1 & zm(z^2) \end{bmatrix} \otimes \mathcal{I}_i, \tag{10}$$

---

[1] Notice that $\hat{a}_{\ell,n}^*(z)$ as defined is *not* the Hermitian adjoint of $\hat{a}_{\ell,n}(z)$. Despite this ambiguity, we introduce this notation which remains natural and widespread in Signal Processing.
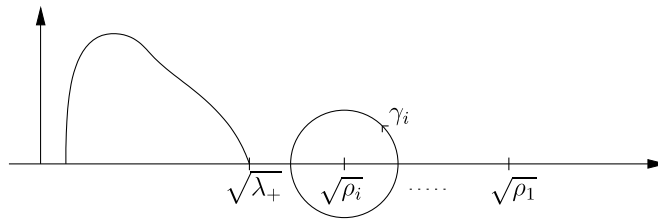
Fig. 1. The contour γ_i w.r.t. the support of the limit singular value distribution of X_n and the other $\sqrt{\rho_\ell}$.

which leads to

$$T_i = \frac{1}{\imath\pi} \sum_{\ell=1}^{s} b_1^* \Pi_\ell b_2 \oint_{\gamma_i} \frac{z^3 m(z^2)^2 \tilde{m}(z^2) \omega_\ell^2}{z^2 m(z^2) \tilde{m}(z^2) \omega_\ell^2 - 1} \, dz$$

$$= \frac{1}{2\imath\pi} \sum_{\ell=1}^{s} b_1^* \Pi_\ell b_2 \oint_{\gamma_i'} \frac{wm(w)^2 \tilde{m}(w) \omega_\ell^2}{wm(w) \tilde{m}(w) \omega_\ell^2 - 1} \, dw$$

by making the change of variable $w = z^2$. Observe that the path γ_i' now encloses ρ_i only. Recall that $wm(w)\tilde{m}(w)\omega_\ell^2 - 1 = 0$ if and only if $w = \rho_\ell$ for every ℓ such that $\omega_\ell^2 > 1/g(\lambda_+^+)$, and since $g(w) = wm(w)\tilde{m}(w)$ is decreasing on (λ_+, ∞), these zeros are simple. As a result, the integrals above are equal to zero for $\ell \neq i$, and the integrand has a simple pole at $w = \rho_i$ for $\ell = i$. By the Residue Theorem, we have:

$$T_i = \frac{1}{\imath\pi} \oint_{\gamma_i} a_1^*(z) H(z)^{-1} a_2(z) \, dz = \frac{\rho_i m(\rho_i)^2 \tilde{m}(\rho_i)}{(\rho_i m(\rho_i) \tilde{m}(\rho_i))'} b_1^* \Pi_i b_2 \tag{11}$$

where the denominator at the right hand side is the derivative of the function $\lambda \mapsto \lambda m(\lambda)\tilde{m}(\lambda)$ at $\lambda = \rho_i$. We now make this argument more rigorous.

Theorem 2. *Let Assumptions A1–A5 hold true. For a given $i \leq s$, assume that $\omega_i^2 > 1/g(\lambda_+^+)$. Let $(b_{1,n})$ and $(b_{2,n})$ be two sequences of deterministic vectors with bounded norms. Then*

$$b_{1,n}^* \widehat{\Pi}_{i,n} b_{2,n} - \frac{\rho_i m(\rho_i)^2 \tilde{m}(\rho_i)}{(\rho_i m(\rho_i) \tilde{m}(\rho_i))'} b_{1,n}^* \Pi_{i,n} b_{2,n} \xrightarrow[n\to\infty]{\text{a.s.}} 0.$$

Proof. Write

$$\widehat{T_i} = \frac{1}{\imath\pi} \oint_{\gamma_i} \hat{a}_1^*(z) \widehat{H}(z)^{-1} \hat{a}_2(z) \, dz.$$

Then, with probability one, $b_1^* \widehat{\Pi}_i b_2 = \widehat{T_i}$ for n large enough. Indeed, on the set O_n (as defined in (3)), the singular values of Σ greater than $\sqrt{\lambda_+} + \varepsilon_1$ coincide with the poles of $\widehat{H}(z)$ which are greater than $\sqrt{\lambda_+} + \varepsilon_1$ by the argument preceding Theorem 1. On this set, the first integral on the right hand side (r.h.s.) of (8) is zero, and by Theorem 1, the second integral can be replaced with $\int_{\gamma_i}$ with probability one for n large enough. By Lemma 5, the differences $\widehat{H}(z) - H(z)$, $\hat{a}_1(z) - a_1(z)$, and $\hat{a}_2(z) - a_2(z)$ a.s. converge to zero, uniformly on γ_i. Hence $\widehat{T_i} - T_i \xrightarrow{\text{a.s.}} 0$. $\quad\square$

4. The spike MUSIC estimation algorithm

Algorithm description

We now consider the application context described in the introduction, and assume that $P_n = B_n(\varphi_1, \ldots, \varphi_r) S_n^*$ where $B_n(\varphi_1, \ldots, \varphi_r) = [b_n(\varphi_1) \cdots b_n(\varphi_k)]$, and $b_n(\varphi) = N^{-1/2} [\exp(-\imath D\ell\varphi)]_{\ell=0}^{N-1}$ with domain $\varphi \in [0, \pi/D]$. When the φ_k are different, one can check that $B_n^* B_n \to I_r$ as $n \to \infty$. In most practical cases of interest, $S_n^* S_n \to O^2$ where O is given by Eq. (2). In these conditions, due to $B_n^* B_n \to I_r$, the diagonal elements of O are the limits of the singular values of P_n and Assumption A5 holds true.

In the area of signal processing, the positive real numbers ω_i^2 are called the Signal to Noise Ratios (SNRs) associated with the r sources. Assumption A5 becomes the following.

Assumption A6. Matrices P_n of dimension $N \times n$ are deterministic and are written as

$$P_n = B_n(\varphi_1, \ldots, \varphi_r) S_n^*$$

where r is a fixed integer, $B_n(\varphi_1, \ldots, \varphi_r) = \begin{bmatrix} b_n(\varphi_1) & \cdots & b_n(\varphi_r) \end{bmatrix}$ is a $N \times r$ matrix, $b_n(\varphi) = N^{-1/2} \left[\exp(-\imath D\ell\varphi) \right]_{\ell=0}^{N-1}$ on $\varphi \in [0, \pi/D]$, and the φ_k are all different. Matrix S_n of dimensions $n \times r$ satisfies:

$$\sqrt{n}(S_n^* S_n - O^2) = \mathcal{O}(1)$$

as $n \to \infty$, where O is defined in Assumption A5, and $\mathcal{O}$ is the classical Landau notation.

The assumption over the speed of convergence of S^*S will be needed only for the purpose of the second order analysis. It is satisfied by most practical systems met in the field of signal processing. We moreover observe that it is possible to relax the assumption that O is diagonal at the expense of a more complicated second order analysis.

In order for the algorithm to be able to estimate the r angles, it is necessary that the perturbation P gives rise to r isolated eigenvalues, a fact that is stated in the following assumption.

Assumption A7. Recall the definition (6) of function g, let λ_+ as defined in A3 and let $g(\lambda_+^+) = \lim_{x \downarrow \lambda_+} g(x)$. Let the ω_i's as defined in A5, then:

$$\omega_r^2 > \frac{1}{g(\lambda_+^+)}.$$

The Spike MUSIC algorithm goes like this. The localization function $\chi(\varphi)$ defined in the introduction is also written as $\chi(\varphi) = \sum_{i=1}^s b(\varphi)^* \Pi_i b(\varphi)$. Given φ, the results of the previous section (Theorems 1 and 2 with $b_1 = b_2 = b(\varphi)$) show us that:

$$\hat{\chi}_n(\varphi) = \sum_{k=1}^r |b_n(\varphi)^* \hat{u}_{k,n}|^2 \zeta(\hat{\lambda}_{k,n}), \tag{12}$$

where

$$\zeta(\lambda) = \frac{(\lambda m(\lambda) \tilde{m}(\lambda))'}{\lambda m(\lambda)^2 \tilde{m}(\lambda)} \tag{13}$$

is a consistent estimator of $\chi_n(\varphi)$ in the asymptotic regime described by A1. By searching for the maxima of $\hat{\chi}(\varphi)$, we infer that we obtain consistent estimates of the angles or arrival. Observe that this algorithm requires the knowledge of the Stieltjes Transform of the limit spectral measure of XX^* (available if the statistical description of the noise is known) and the number r of emitting sources. Notice that when this number is unknown, it can be estimated along the ideas described in *e.g.* [10,22].

We now perform the first order analysis of this algorithm.

First order analysis of the spike MUSIC algorithm

We now formalize the argument of the previous paragraph and we push it further to show the consistency "up to the order n" of the Spike MUSIC estimator. We shall need this speed to perform the second order analysis (Lemma 9).

Theorem 3. *Let Assumptions A1–A6 hold true. Then for all $k = 1, \ldots, r$, there exists a local maximum $\hat{\varphi}_{k,n}$ of $\hat{\chi}_n(\varphi)$ such that*

$$n(\hat{\varphi}_{k,n} - \varphi_k) \xrightarrow[n\to\infty]{\text{a.s.}} 0.$$

The proof of this theorem is performed in two steps. With an approach similar to the one used in Section 3, we first prove that $\hat{\chi}(\varphi) - \chi(\varphi) \xrightarrow{\text{a.s.}} 0$, and the convergence is uniform on $\varphi \in [0, \pi/D]$ (Proposition 1). Next, following the technique of [16,17], we prove that this uniform a.s. convergence leads to Theorem 3.

In the sequel, we write:

$$\hat{a}(z, \varphi) = \begin{bmatrix} zU^*Q(z^2) \\ \Omega V^* \widetilde{Q}(z^2) X^* \end{bmatrix} b(\varphi) \quad \text{and} \quad a(z, \varphi) = \begin{bmatrix} zm(z^2)U^* \\ 0 \end{bmatrix} b(\varphi), \tag{14}$$

$$\hat{a}^*(z, \varphi) = b^*(\varphi)[zQ(z^2)U \quad X\widetilde{Q}(z^2)V\Omega],$$

$$a^*(z, \varphi) = b(\varphi)[zm(z^2)U \quad 0].$$

Beware that $\hat{a}^*$ and a^* are not the Hermitian adjoints of $\hat{a}$ and a (see the footnote associated to Eq. (9)).

Proposition 1. *In the setting of Theorem 3,*

$$\max_{\varphi \in [0, \pi/D]} \left| \hat{\chi}_n(\varphi) - \chi_n(\varphi) \right| \xrightarrow[n \to \infty]{\text{a.s.}} 0.$$

Proof. Write

$$\hat{\chi}(\varphi) - \chi(\varphi) = \sum_{k=1}^{r} (\zeta(\hat{\lambda}_k) - \zeta(\rho_{i(k)})) |b(\varphi)^* \hat{u}_k|^2 + \sum_{i=1}^{s} \left(\zeta(\rho_i) b(\varphi)^* \widehat{\Pi}_i b(\varphi) - b(\varphi)^* \Pi_i b(\varphi) \right).$$

By Theorem 1 and the continuity of ζ on $(\lambda_+, +\infty)$, the first term at the r.h.s. goes to zero a.s. and uniformly in φ. Consider the second term. Let γ_i be a small enough positively oriented circle which does not meet $\text{supp}(\pi) \cup \{\sqrt{\rho_1}, \ldots, \sqrt{\rho_s}\}$ and such that only $\sqrt{\rho_i} \in \text{Int}(\gamma_i)$. Since $\hat{\lambda}_k \xrightarrow{\text{a.s.}} \rho_{i(k)}$,

$$\max_i \max_\varphi \left| b(\varphi)^* \widehat{\Pi}_i b(\varphi) - \widehat{T}_i(\varphi) \right| = 0$$

a.s. for n large enough, where

$$\widehat{T}_i(\varphi) = \frac{1}{\imath \pi} \oint_{\gamma_i} \hat{a}^*(z, \varphi) \widehat{H}(z)^{-1} \hat{a}(z, \varphi) \, dz.$$

Recalling Eq. (11), it will therefore be enough to prove that

$$\max_{1 \le i \le s} \max_{\varphi \in [0, \pi/D]} |Z_i(\varphi)| \xrightarrow[n \to \infty]{\text{a.s.}} 0,$$

where

$$Z_i(\varphi) = \frac{1}{\imath \pi} \oint_{\gamma_i} \left(\hat{a}^*(z, \varphi) \widehat{H}(z)^{-1} \hat{a}(z, \varphi) - a^*(z, \varphi) H(z)^{-1} a(z, \varphi) \right) \, dz.$$

We have

$$\max_\varphi |Z_i(\varphi)| \le 2R \int_0^1 \max_\varphi e(\sqrt{\rho_i} + R e^{2\imath \pi \theta}, \varphi) \, d\theta$$

where R is the radius of γ_i and where

$$\begin{aligned} e(z, \varphi) &= \left| \hat{a}^*(z, \varphi) \widehat{H}(z)^{-1} \hat{a}(z, \varphi) - a^*(z, \varphi) H(z)^{-1} a(z, \varphi) \right| \\ &\le \left| (\hat{a}^* - a^*) H^{-1} \hat{a} \right| + \left| a H^{-1} (\hat{a} - a) \right| + \left| \hat{a}^* (\widehat{H}^{-1} - H^{-1}) \hat{a} \right|. \end{aligned}$$

Since $\|H^{-1}\|$, $\max_\varphi \|a\|$ and $\max_\varphi \|\hat{a}\|$ are bounded on γ_i, $e(z, \varphi)$ satisfies on this path

$$e(z, \varphi) \le K \left(\|\hat{a}(z, \varphi) - a(z, \varphi)\| + \|\widehat{H}(z)^{-1} - H(z)^{-1}\| \right).$$

By Lemma 5 and the fact that $\|H^{-1}\|$ is bounded on γ_i, the term $\|\widehat{H}^{-1} - H^{-1}\| = \|\widehat{H}^{-1}(H - \widehat{H}) H^{-1}\|$ converges to zero uniformly on γ_i with probability one. To obtain the result, we prove that $\|\hat{a} - a\| \xrightarrow{\text{a.s.}} 0$ and that this convergence is uniform on $(z, \varphi) \in \gamma_i \times [0, \pi/D]$. Let us focus on the first term $z u_1^*(Q(z^2) - m(z^2)I) b(\varphi)$ of $\hat{a} - a$, where we recall that u_1 is the first column of U. Since $\|b(\varphi)\| = \|u_1\| = 1$,

$$|z u_1^*(Q(z^2) - m(z^2)I) b(\varphi)| \le |z u_1^*(Q(z^2) - \alpha(z^2)I) b(\varphi)| + |z(\alpha(z^2) - m(z^2))|.$$

With probability one, the second term converges to zero on γ_i, and the convergence is uniform (along the principle of the proof of Lemma 5). Since

$$\sup_n \max_\varphi \|n^{-1} b'(\varphi)\| = \sup_n \max_\varphi \left\| n^{-1} N^{-1/2} \left[\ell D \exp(-\imath \ell D \varphi) \right]_{\ell=0}^{N-1} \right\| < \infty,$$

the term

$$\xi(z, \varphi) = \mathbb{1}_{0_n} \times z u_1^*(Q(z^2) - \alpha(z^2)I) b(\varphi)$$

satisfies

$$|\xi(z_1, \varphi_1) - \xi(z_2, \varphi_2)| \le K(n|\varphi_1 - \varphi_2| + |z_1 - z_2|)$$

for every (z_1, φ_1), (z_2, φ_2) in $\gamma_i \times [0, \pi/D]$. Therefore, it will be enough to prove that

$$\max_{(z, \varphi) \in A_n \times B_n} \xi(z, \varphi) \xrightarrow[n \to \infty]{\text{a.s.}} 0$$

where A_n contains n regularly spaced points in γ_i and B_n contains n^2 regularly spaced points in $[0, \pi/D]$. This can be obtained from Lemma 3 with $p = 9$, the Markov inequality and Borel–Cantelli's lemma. The other terms of $\hat{a} - a$ can be handled similarly. □

We now prove Theorem 3 by following the ideas of [16,17]. To that end, we need the following lemma, proven in [14].

Lemma 7. *Let (c_N) be a sequence of real numbers belonging to a compact of $[-1/2, 1/2]$ and converging to c. Let*

$$q_N(c_N) = \frac{1}{N} \sum_{k=0}^{N-1} \exp(-2\imath\pi k c_N).$$

Then the following hold true:

$$q_N(c_N) \xrightarrow[N\to\infty]{} 0 \quad \text{if } c \neq 0,$$

$$q_N(c_N) \xrightarrow[N\to\infty]{} 0 \quad \text{if } c = 0 \text{ and } N|c_N - c| \to \infty,$$

$$q_N(c_N) \xrightarrow[N\to\infty]{} \exp(-\imath\pi d)\mathrm{sinc}(d) \quad \text{if } c = 0 \text{ and } N|c_N - c| \to d,$$

where sinc *stands as usual for sine cardinal.*

Proof of Theorem 3. We start by observing that $\chi(\varphi) = d(\varphi)^*(B^*B)^{-1}d(\varphi)$ where B is the matrix defined in A6 and where $d(\varphi) = \left[b(\varphi_k)^*b(\varphi)\right]_{k=1}^r$. By Lemma 7, $B^*B \to I_r$; hence $\chi(\varphi) - \|d(\varphi)\|^2 \to 0$.

In the remainder of the proof, we shall stay in the probability one set where the uniform convergence in the statement of Proposition 1 holds true. Taking $k = 1$ without loss of generality, we shall show that any sequence $\hat{\varphi}_{1,n}$ for which $\hat{\chi}(\hat{\varphi}_{1,n})$ attains its maximum in the closure of a small neighborhood of φ_1 satisfies $N(\hat{\varphi}_{1,n} - \varphi_1) \to 0$. Given a sequence of such $\hat{\varphi}_{1,n}$, assume we can extract a subsequence $\hat{\varphi}_{1,n^*}$ such that $N|\hat{\varphi}_{1,n^*} - \varphi_1| \to \infty$. In this case, Lemma 7 and the observations made above on the structure of $\chi(\varphi)$ show that $\chi(\hat{\varphi}_{1,n^*}) \to 0$. Since $\max_\varphi |\hat{\chi}(\varphi) - \chi(\varphi)| \to 0$, $\hat{\chi}(\hat{\varphi}_{1,n^*}) \to 0$. But $\hat{\chi}(\varphi_1) \to \chi(\varphi_1) = 1$, which contradicts the fact that $\hat{\varphi}_{1,n^*}$ maximizes $\hat{\chi}$. Hence the sequence $N(\hat{\varphi}_{1,n^*} - \varphi_1)$ belongs to a compact. Assume that $N(\hat{\varphi}_{1,n^*} - \varphi_1) \not\to 0$. If we take a further subsequence of the latter that converges to a constant $d \neq 0$, then by Lemma 7, $\hat{\chi}$ converges to $\mathrm{sinc}(d)^2 < 1$ along this subsequence, which also raises a contradiction. This proves the theorem. □

5. Second order analysis of the spike MUSIC estimator

In order to perform the second order analysis, we also assume the following.

Assumption A8. Let λ_-, λ_+, α and m be as in A3. Then for any $z \in \mathbb{C} - [\lambda_-, \lambda_+]$, $\sqrt{n}\,(\alpha(z) - m(z))$ converges in probability to zero.

Remark 2. If $\sqrt{n}X$ is standard Gaussian and if $c_n = N/n$ satisfies $\sqrt{n}(c_n - c) \to 0$, then Assumption A8 is satisfied. Indeed, call $m_n(z)$ the Stieltjes Transform of the Marčenko–Pastur distribution, *i.e.*, the analytic continuation of (7), when c is replaced with c_n, and let π_n be the associated probability measure. For $z \in \mathbb{C} - [\lambda_-, \lambda_+]$, function $f(x) = (x - z)^{-1}$ is analytic outside the support of π_n for n large, and [2, Theorem 1.1] can be applied to show that $\sqrt{n}(\alpha_n(z) - m_n(z)) \xrightarrow{\mathcal{P}} 0$. When $\sqrt{n}(c_n - c) \to 0$, it is furthermore clear that $\sqrt{n}(m_n(z) - m(z)) \to 0$.

The main result of this section is the following.

Theorem 4. *Let Assumptions A1–A8 hold true. Then the estimates $\hat{\varphi}_{k,n}$ satisfy*

$$n^{3/2} \begin{bmatrix} \hat{\varphi}_{1,n} - \varphi_1 \\ \vdots \\ \hat{\varphi}_{r,n} - \varphi_r \end{bmatrix} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \begin{bmatrix} \sigma_1^2 I_{j_1} & & \\ & \ddots & \\ & & \sigma_s^2 I_{j_s} \end{bmatrix}\right) \tag{15}$$

where

$$\sigma_i^2 = \frac{6}{c^2 D^2}\left(\frac{m'(\rho_i) - m(\rho_i)^2}{c m(\rho_i)^2} + \omega_i^2\left(m(\rho_i) + \rho_i m'(\rho_i)\right)\right), \quad 1 \leq i \leq s.$$

When $\sqrt{n}X$ is standard Gaussian, plugging the r.h.s. of (7) into this expression leads after some derivations to the following corollary.
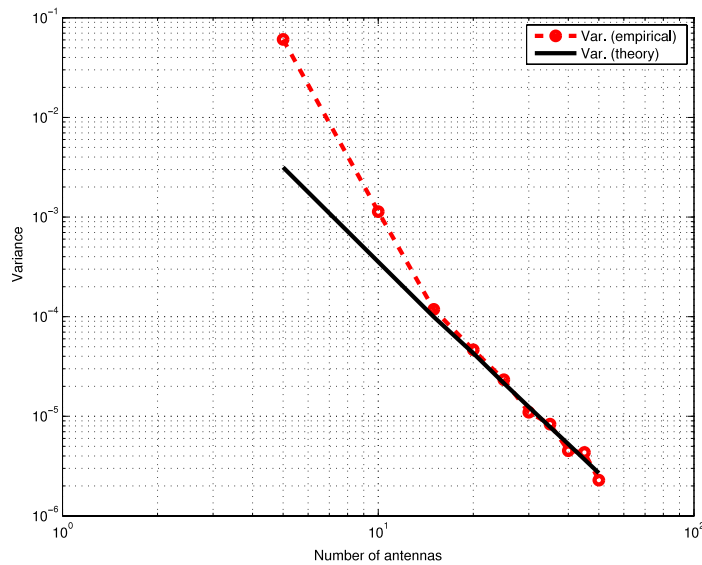
Fig. 2. Spike MUSIC algorithm, variance *vs. N*.

Corollary 2. *If* $\sqrt{n}X$ *is standard Gaussian and if* $\sqrt{n}(c_n - c) \to 0$, *the convergence* (15) *holds true with*

$$\sigma_i^2 = \frac{6}{c^2 D^2} \frac{\omega_i^2 + 1}{\omega_i^4 - c}.$$

This corollary calls for some comments.

Remark 3 (*Efficiency at High SNR*). Recalling that $\omega_i^2 > \sqrt{c}$ is the condition for the existence of a corresponding isolated eigenvalue (Corollary 1), we observe that the estimator variance for φ_k goes to infinity as the corresponding ω_i^2 decreases to $\sqrt{c}$. At the other extreme, this variance behaves like $6c^{-2}D^{-2}\omega_i^{-2}$ as $\omega_i^2 \to \infty$. It is useful to notice that this asymptotic variance coincides with the Cramér–Rao bound for estimating φ_k [28]. In other words, the Spike MUSIC estimator is efficient at high SNR when the noise matrix is standard Gaussian.

A numerical illustration

In order to illustrate the convergence and the fluctuations of the Spike MUSIC algorithm, we simulate a radio signal transmission satisfying Assumptions A1–A8. We consider $r = 2$ emitting sources located at the angles 0.5 and 1 rad., and a number of receiving antennas ranging from $N = 5$ to $N = 50$. The observation window length is set to $n = 2N$ (hence $c = 0.5$). The noise matrix X_n is such that $\sqrt{n}X_n$ is standard Gaussian. The source powers are assumed equal, so that the matrix O given by Eq. (2) is written $O = \omega I_2$, and the Signal to Noise Ratio for any source is SNR $= 10\log_{10}\omega^2$ decibels. In Fig. 2, the SNR is set to 10 dB, and the empirical variance of $\hat{\varphi}_{1,n} - \varphi_1$ (red curve) is computed over 2000 runs. The variance provided by Corollary 2 is also plotted versus N. We observe a good fit between the variance predicted by Corollary 2 and the empirical variance after $N = 15$ antennas. In Fig. 3, the variance is plotted as a function of the SNR, the number of antennas being fixed to $N = 20$. The empirical variance is computed over 5000 runs. The Cramér–Rao Bound is also plotted. The empirical variance fits the theoretical one from SNR $\approx$ 6 dB upwards.

Proof of Theorem 4. We start with some additional notations and definitions. Matrix $B = [b(\varphi_1), \ldots, b(\varphi_r)]$ will be often written as $B = [b_1, \ldots, b_r]$ or in block form as $B = [B_1, \ldots, B_s]$ where B_i has j_i columns. We shall also write $B' = [b'(\varphi_1), \ldots, b'(\varphi_r)]$ and $B'' = [b''(\varphi_1), \ldots, b''(\varphi_r)]$ where $b'(\varphi)$ and $b''(\varphi)$ are respectively the first and second derivatives of $b(\varphi)$. We shall also use the short hand notations $B' = [b'_1, \ldots, b'_r]$ and $B'' = [b''_1, \ldots, b''_r]$. Matrix $B^\perp = [b_1^\perp, \ldots, b_r^\perp]$ will be defined by the equation

$$\frac{1}{n}B' = -\frac{\imath c D}{2}B + \frac{cD}{2\sqrt{3}}B^\perp. \tag{16}$$

Finally, if x_n, y_n are random sequences, we denote by $x_n \asymp y_n$ the convergence $x_n - y_n \overset{\mathcal{P}}{\to} 0$.
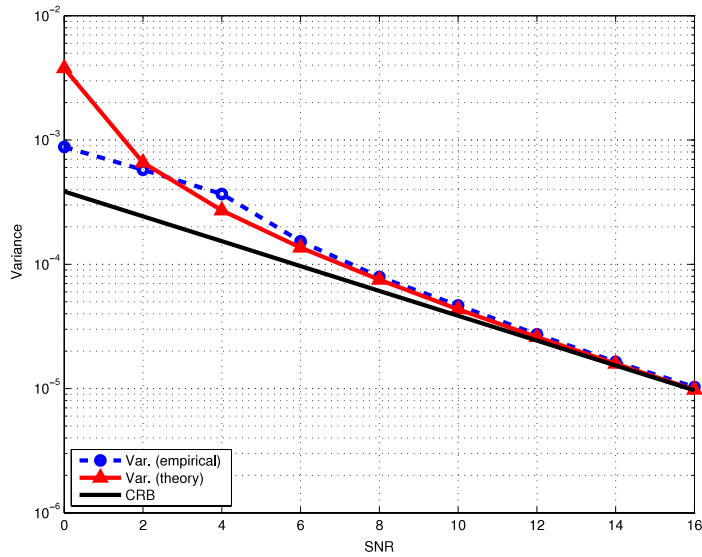
Fig. 3. Spike MUSIC algorithm, variance *vs.* the SNR.

We now state some preliminary results. In the following, we say that the complex random vector η is governed by the law $\mathcal{CN}(0, R)$ where R is a nonnegative Hermitian matrix if the real vector $\begin{bmatrix} \Re(\eta) \\ \Im(\eta) \end{bmatrix}$ has the law $\mathcal{N}\left(0, \frac{1}{2}\begin{bmatrix} \Re(R) & -\Im(R) \\ \Im(R) & \Re(R) \end{bmatrix}\right)$. The following proposition, whose proof is postponed to Appendix A, is crucial. □

Proposition 2. *Let Assumptions* A1–A4 *hold true. Let* $t \leq N$ *be a fixed integer, let* $W = [w_1, \ldots, w_t]$ *and* $\widetilde{W} = [\tilde{w}_1, \ldots, \tilde{w}_t]$ *be deterministic isometry matrices with dimensions* $N \times t$ *and* $n \times t$ *respectively. Let* ρ *be a real number such that* $\rho > \lambda_+$. *Then*

$$\xi_n = \sqrt{n}\left(W^*\left(Q(\rho) - \alpha(\rho)I_N\right)W, \ \widetilde{W}^*\left(\widetilde{Q}(\rho) - \tilde{\alpha}(\rho)I_n\right)\widetilde{W}, \ W^*X\widetilde{Q}(\rho)\widetilde{W}\right)$$

is tight.

Assume that t *is even. Given real numbers* $\rho_1, \ldots, \rho_{t/2}$ *all strictly greater than* λ_+, *the* $t \times 1$ *random vector*

$$\eta_n = \left[\sqrt{N}\left(w_k^*Q(\rho_k)w_{t/2+k}\right)_{1\leq k\leq t/2}, \ \sqrt{n}\left(w_k^*X\widetilde{Q}(\rho_k)\tilde{w}_k\right)_{1\leq k\leq t/2}\right]^T$$

converges in distribution towards $\mathcal{CN}(0, R)$ *with*

$$R = \begin{bmatrix} \operatorname{diag}\left(m'(\rho_k) - m(\rho_k)^2\right)_{k=1}^{t/2} & 0 \\ 0 & \operatorname{diag}\left(m(\rho_k) + \rho_k m'(\rho_k)\right)_{k=1}^{t/2} \end{bmatrix}.$$

Writing $Q - mI = (Q - \alpha I) + (\alpha - m)I$, and similarly for $\widetilde{Q}$, we obtain the following.

Corollary 3. *Assume in addition that Assumption* A8 *is satisfied. Then*

$$\xi_n = \sqrt{n}\left(W^*\left(Q(\rho) - m(\rho)I_N\right)W, \ \widetilde{W}^*\left(\widetilde{Q}(\rho) - \tilde{m}(\rho)I_n\right)\widetilde{W}, \ W^*X\widetilde{Q}(\rho)\widetilde{W}\right)$$

is tight.

Intuitively, tightness of ξ_n leads to the tightness of the $\sqrt{n}(\hat{\lambda}_{k,n} - \rho_{i(k)})$. This is formalized by the following proposition, proven in Appendix B.

Proposition 3. *Assume the setting of Theorem* 4. *Then the sequences* $\sqrt{n}(\hat{\lambda}_{k,n} - \rho_{i(k)})$ *are tight for* $1 \leq k \leq r$.

The following lemma is proven in Appendix C.

Lemma 8. *Let Assumptions* A5 *and* A6 *hold true. Then the following convergences hold true:*

$$B^*B \xrightarrow[n \to \infty]{} I_r,$$

$$\frac{1}{n^2}B^*B'' \xrightarrow[n \to \infty]{} -\left(\frac{c^2 D^2}{3}\right) I_r,$$

$$(B^\perp)^*B^\perp \xrightarrow[n \to \infty]{} I_r,$$

$$(B^\perp)^*B \xrightarrow[n \to \infty]{} 0,$$

$$\|\Pi_i - \Pi_{B_i}\| \xrightarrow[n \to \infty]{} 0 \quad \text{for all } i = 1, \ldots, s$$

where Π_{B_i} is the orthogonal projection matrix on the column space of B_i.

We now enter the proof of Theorem 4.

Recall the definitions (12) and (13) of $\hat{\chi}$ and ζ. In most of the proof, we shall focus on $\sqrt{n}(\hat{\varphi}_{1,n} - \varphi_1)$. Recalling that $\hat{\chi}'(\hat{\varphi}_1) = 0$ and performing a Taylor–Lagrange expansion of $\hat{\chi}'$ around φ_1, we obtain

$$0 = \hat{\chi}'(\hat{\varphi}_1) = \hat{\chi}'(\varphi_1) + (\hat{\varphi}_1 - \varphi_1)\hat{\chi}''(\varphi_1) + \frac{(\hat{\varphi}_1 - \varphi_1)^2}{2}\hat{\chi}^{(3)}(\bar{\varphi}_1),$$

where $\hat{\chi}^{(3)}$ is the third derivative of $\hat{\chi}$ and where $\bar{\varphi}_1 \in [\varphi_1 \wedge \hat{\varphi}_1, \varphi_1 \vee \hat{\varphi}_1]$. Hence

$$n^{3/2}(\hat{\varphi}_1 - \varphi_1) = -\frac{n^{-1/2}\hat{\chi}'(\varphi_1)}{n^{-2}\hat{\chi}''(\varphi_1) + 0.5 n^{-2}(\hat{\varphi}_1 - \varphi_1)\hat{\chi}^{(3)}(\bar{\varphi}_1)}.$$

We start by characterizing the asymptotic behavior of the denominator of this equation.

Lemma 9. *Assume that the setting of Theorem 4 holds true. Then,*

$$\frac{\hat{\chi}''(\varphi_1)}{n^2} + (\hat{\varphi}_1 - \varphi_1)\frac{\hat{\chi}^{(3)}(\bar{\varphi}_1)}{2n^2} \xrightarrow[n \to \infty]{\text{a.s.}} -\frac{c^2 D^2}{6}.$$

Proof. We have

$$\frac{\hat{\chi}''(\varphi_1)}{n^2} = \frac{2}{n^2}\sum_{k=1}^{r}\zeta(\hat{\lambda}_k)|(b_1')^*\widehat{u}_k|^2 + \frac{2}{n^2}\sum_{k=1}^{r}\Re\left(\zeta(\hat{\lambda}_k)b_1^*\widehat{u}_k\widehat{u}_k^*b_1''\right),$$

$$\frac{\chi''(\varphi_1)}{n^2} = \frac{2}{n^2}(b_1')^*UU^*b_1' + \frac{2}{n^2}\Re\left(b_1^*UU^*b_1''\right). \tag{17}$$

Theorem 1 along with the continuity of ζ on (λ_+, ∞), and Theorem 2 show that

$$\frac{1}{n^2}\hat{\chi}''(\varphi_1) - \frac{1}{n^2}\chi''(\varphi_1) \xrightarrow[n \to \infty]{\text{a.s.}} 0.$$

Writing

$$\frac{1}{n^2}\chi''(\varphi_1) = \frac{2}{n^2}\sum_{i=1}^{s}\left((b_1')^*\Pi_i b_1' + \Re(b_1^*\Pi_i b_1'')\right),$$

we have

$$\frac{1}{n^2}(b_1')^*\Pi_i b_1' = \left(-\frac{\imath c D}{2}b_1 + \frac{cD}{2\sqrt{3}}b_1^\perp\right)^*\Pi_i\left(-\frac{\imath c D}{2}b_1 + \frac{cD}{2\sqrt{3}}b_1^\perp\right),$$

$$\xrightarrow[n \to \infty]{} \frac{c^2 D^2}{4}\delta_{i,0}$$

by the first, fourth and fifth assertions of Lemma 8. By the same lemma,

$$\frac{1}{n^2}b_1^*\Pi_i b_1'' - \frac{\delta_{i,0}}{n^2}b_1^*b_1'' \to 0 \quad \text{and} \quad \frac{1}{n^2}b_1^*b_1'' \to -\frac{c^2 D^2}{3}.$$

Hence $n^{-2}\hat{\chi}''(\varphi_1) \to -c^2 D^2/6$.

Furthermore, it is easily seen that $n^{-3}\hat{\chi}^{(3)}(\bar{\varphi}_1)$ is bounded. Since $n(\hat{\varphi}_1-\varphi_1) \xrightarrow{\text{a.s.}} 0$ by Theorem 3, $n^{-2}(\hat{\varphi}_1-\varphi_1)\hat{\chi}^{(3)}(\bar{\varphi}_1) \xrightarrow{\text{a.s.}} 0$, which establishes the result. □

We now turn to the numerator $n^{-1/2}\hat{\chi}'(\varphi_1) = 2n^{-1/2}\sum_{k=1}^{r}\zeta(\hat{\lambda}_k)\Re\left(b_1^*\hat{u}_k\hat{u}_k^*b_1'\right)$, and start with the following lemma.

Lemma 10. *Assume that the setting of Theorem 4 holds true. Then*

$$\frac{1}{\sqrt{n}}\hat{\chi}'(\varphi_1) - 2\Re(\xi) \xrightarrow{\mathcal{P}} 0,$$

where

$$\xi = \sum_{i=1}^{s}\frac{\zeta(\rho_i)}{\imath\pi\sqrt{n}}\oint_{\gamma_i}\left(\hat{a}^*(z,\varphi_1)\widehat{H}(z)^{-1}\hat{a}_\varphi'(z,\varphi_1) - a^*(z,\varphi_1)H(z)^{-1}a_\varphi'(z,\varphi_1)\right)dz, \qquad (18)$$

and where the deterministic circle γ_i encloses $\rho_i^{1/2}$ only and:

$$\hat{a}_\varphi'(z,\varphi) = \frac{\partial\hat{a}(z,\varphi)}{\partial\varphi} = \begin{bmatrix} zU^*Q(z^2) \\ \Omega V^*\widetilde{Q}(z^2)X^* \end{bmatrix}b'(\varphi),$$

$$a_\varphi'(z,\varphi) = \frac{\partial a(z,\varphi)}{\partial\varphi} = \begin{bmatrix} zm(z^2)U^* \\ 0 \end{bmatrix}b'(\varphi).$$

Proof. Recall the definition of $\hat{\chi}$ as given in (12). A direct computation yields:

$$\hat{\chi}'(\varphi) = 2\sum_{k=1}^{r}\zeta(\hat{\lambda}_{k,n})\Re\left(b_1^*(\varphi)\hat{u}_k\hat{u}_k^*b_1'(\varphi)\right),$$

$$= 2\sum_{i=1}^{s}\sum_{k:i(k)=i}\zeta(\hat{\lambda}_{k,n})\Re\left(b_1^*(\varphi)\hat{u}_k\hat{u}_k^*b_1'(\varphi)\right).$$

Recall that r and s are fixed and independent from n by A5. We start by showing that

$$\frac{1}{\sqrt{n}}\hat{\chi}'(\varphi_1) - \frac{2}{\sqrt{n}}\sum_{i=1}^{s}\zeta(\rho_i)\Re\left(b_1^*\widehat{\Pi}_ib_1'\right) \xrightarrow[n\to\infty]{\mathcal{P}} 0. \qquad (19)$$

Since $\sqrt{n}(\zeta(\hat{\lambda}_{k,n}) - \zeta(\rho_{i(k)}))$ is tight as a corollary of Proposition 3, it will be enough to prove that $n^{-1}\Re\left(b_1^*\hat{u}_k\hat{u}_k^*b_1'\right) \to 0$ in probability for every k. By the definition (16) of $B^\perp$, we have

$$\frac{1}{n}\Re\left(b_1^*\hat{u}_k\hat{u}_k^*b_1'\right) = \frac{cD}{2\sqrt{3}}\Re\left(b_1^*\hat{u}_k\hat{u}_k^*b_1^\perp\right).$$

By the Cauchy–Schwarz inequality,

$$\left|b_1^*\hat{u}_k\hat{u}_k^*b_1^\perp\right|^2 \leq b_1^*\widehat{\Pi}_{i(k)}b_1\,(b_1^\perp)^*\widehat{\Pi}_{i(k)}b_1^\perp.$$

By Theorem 2,

$$b_1^*\widehat{\Pi}_{i(k)}b_1(b_1^\perp)^*\widehat{\Pi}_{i(k)}b_1^\perp - \zeta(\rho_{i(k)})^{-2}\,b_1^*\Pi_{i(k)}b_1(b_1^\perp)^*\Pi_{i(k)}b_1^\perp \xrightarrow{\text{a.s.}} 0,$$

and by Lemma 8, $b_1^*\Pi_{i(k)}b_1(b_1^\perp)^*\Pi_{i(k)}b_1^\perp \to 0$ (consider alternatively the cases $i(k)=1$ and $i(k)>1$) which proves (19).
    Now, applying (8) and (14), and taking up an argument used in the proof of Theorem 2, we have

$$2\sum_{i=1}^{s}\frac{\zeta(\rho_i)}{\sqrt{n}}\Re\left(b_1^*\widehat{\Pi}_ib_1'\right) = 2\sum_{i=1}^{s}\Re\left(\frac{-\zeta(\rho_i)}{\imath\pi\sqrt{n}}\oint_{\mathcal{C}_i}\begin{bmatrix} b_1^* & 0 \end{bmatrix}\mathbf{Q}(z)\begin{bmatrix} b_1' \\ 0 \end{bmatrix}dz\right)$$

$$+ 2\sum_{i=1}^{s}\Re\left(\frac{\zeta(\rho_i)}{\imath\pi\sqrt{n}}\oint_{\mathcal{C}_i}\hat{a}^*(z,\varphi_1)\widehat{H}(z)^{-1}\hat{a}_\varphi'(z,\varphi_1)dz\right)$$

$$= 2\sum_{i=1}^{s}\Re\left(\frac{\zeta(\rho_i)}{\imath\pi\sqrt{n}}\oint_{\gamma_i}\hat{a}^*(z,\varphi_1)\widehat{H}(z)^{-1}\hat{a}_\varphi'(z,\varphi_1)dz\right)$$

with probability one for n large. On the other hand, recalling (11), we have

$$0 = \chi'(\varphi_1) = 2 \sum_{i=1}^{s} \Re \left(\frac{\zeta(\rho_i)}{\iota\pi} \oint_{\gamma_i} a^*(z, \varphi_1) H(z)^{-1} a'_\varphi(z, \varphi_1) \, dz \right),$$

which proves the result. □

Write $\widehat{H}(z) = H(z) + E(z)$ and $\hat{a}(z, \varphi) = a(z, \varphi) + e(z, \varphi)$. To be more specific,

$$E(z) = \begin{bmatrix} zU^*(Q(z^2) - m(z^2)I_N)U & U^*X\widetilde{Q}(z^2)V\Omega \\ \Omega V^*\widetilde{Q}(z^2)X^*U & z\Omega V^*(\widetilde{Q}(z^2) - \tilde{m}(z^2)I_n)V\Omega \end{bmatrix} \tag{20}$$

and

$$e(z, \varphi) = \begin{bmatrix} zU^* \left(Q(z^2) - m(z^2)I \right) \\ \Omega V^*\widetilde{Q}(z^2)X^* \end{bmatrix} b(\varphi).$$

Write $e'_\varphi(z, \varphi) = \partial e(z, \varphi)/\partial \varphi$. For a given $z \in \gamma_i$, $\widehat{H}^{-1} = H^{-1} - H^{-1}EH^{-1} + \mathcal{O}(\|E\|^2)$. This suggests the following development

$$
\begin{aligned}
\xi &= \sum_{i=1}^{s} \left(\frac{\zeta(\rho_i)}{\iota\pi\sqrt{n}} \oint_{\gamma_i} a^*(z, \varphi_1) H(z)^{-1} e'_\varphi(z, \varphi_1) \, dz + \frac{\zeta(\rho_i)}{\iota\pi\sqrt{n}} \oint_{\gamma_i} e^*(z, \varphi_1) H(z)^{-1} a'_\varphi(z, \varphi_1) \, dz \right. \\
&\quad \left. - \frac{\zeta(\rho_i)}{\iota\pi\sqrt{n}} \oint_{\gamma_i} a^*(z, \varphi_1) H(z)^{-1} E(z) H(z)^{-1} a'_\varphi(z, \varphi_1) \, dz + q_i \right) \\
&= \sum_{i=1}^{s} (X_{1,i} + X_{2,i} + X_{3,i} + q_i)
\end{aligned}
$$

where the terms q_i are "higher order terms" that appear when we expand the r.h.s. of (18). We first handle the terms $X_{k,i}$'s, then q_i.

The terms $X_{1,i}$

Writing $U_n = [U_{1,n} \cdots U_{s,n}]$ and $V_n = [V_{1,n} \cdots V_{s,n}]$ where both $U_{i,n}$ and $V_{i,n}$ have j_i columns, and recalling (10), we have

$$
\begin{aligned}
X_{1,i} &= \frac{\zeta(\rho_i)}{\iota\pi\sqrt{n}} \sum_{\ell=1}^{s} \oint_{\gamma_i} [zm(z^2)b_1^*U_\ell \quad 0] \times \frac{\begin{bmatrix} z\tilde{m}(z^2)\omega_\ell^2 & -1 \\ -1 & zm(z^2) \end{bmatrix} \otimes I_{j_\ell}}{z^2 m(z^2)\tilde{m}(z^2)\omega_\ell^2 - 1} \begin{bmatrix} zU_\ell^* \left(Q(z^2) - m(z^2)I \right) b_1' \\ \omega_\ell V_\ell^*\widetilde{Q}(z^2)X^*b_1' \end{bmatrix} dz \\
&= \frac{\zeta(\rho_i)}{\iota\pi\sqrt{n}} \sum_{\ell=1}^{s} \oint_{\gamma_i} \frac{z^3\omega_\ell^2 m(z^2)\tilde{m}(z^2)b_1^*\Pi_\ell \left(Q(z^2) - m(z^2)I \right) b_1'}{z^2 m(z^2)\tilde{m}(z^2)\omega_\ell^2 - 1} dz \\
&\quad - \frac{\zeta(\rho_i)}{\iota\pi\sqrt{n}} \sum_{\ell=1}^{s} \oint_{\gamma_i} \frac{\omega_\ell zm(z^2)b_1^*U_\ell V_\ell^*\widetilde{Q}(z^2)X^*b_1'}{z^2 m(z^2)\tilde{m}(z^2)\omega_\ell^2 - 1} dz \\
&= \frac{\zeta(\rho_i)}{2\iota\pi\sqrt{n}} \sum_{\ell=1}^{s} \oint_{\gamma_i'} \frac{w\omega_\ell^2 m(w)\tilde{m}(w)b_1^*\Pi_\ell \left(Q(w) - m(w)I \right) b_1'}{wm(w)\tilde{m}(w)\omega_\ell^2 - 1} dw \\
&\quad - \frac{\zeta(\rho_i)}{2\iota\pi\sqrt{n}} \sum_{\ell=1}^{s} \oint_{\gamma_i'} \frac{\omega_\ell m(w)b_1^*U_\ell V_\ell^*\widetilde{Q}(w)X^*b_1'}{wm(w)\tilde{m}(w)\omega_\ell^2 - 1} dw
\end{aligned}
$$

where γ_i' encloses ρ_i only. These integrals are zero for $\ell \neq i$. For large n and with probability one, none of the numerators has a pole within γ_i'; hence by the Residue Theorem

$$X_{1,i} = \frac{b_1^*\Pi_i \left(Q(\rho_i) - m(\rho_i)I \right) b_1'}{\sqrt{n}m(\rho_i)} - \frac{\omega_i b_1^*U_i V_i^*\widetilde{Q}(\rho_i)X^*b_1'}{\sqrt{n}}$$

a.s. for n large enough.

Due to the bounded character of $\|n^{-1}b'\|$ and to Corollary 3, $X_{1,i}$ is tight for every i. By Lemma 8,

$$X_{1,i} \asymp \delta_{i-1,0} \left(\frac{b_1^* \left(Q(\rho_1) - m(\rho_1)I \right) b_1'}{\sqrt{n}m(\rho_1)} - \frac{\omega_1 b_1^*U_1 V_1^*\widetilde{Q}(\rho_1)X^*b_1'}{\sqrt{n}} \right).$$

The terms $X_{2,i}$

We have here

$$
X_{2,i} = \frac{\zeta(\rho_i)}{\imath\pi\sqrt{n}} \sum_{\ell=1}^{s} \oint_{\gamma_i} \left[zb_1^* \left(Q(z^2) - m(z^2)I \right) U_\ell \quad \omega_\ell b_1^* X \widetilde{Q}(z^2) V_\ell \right]
$$
$$
\times \frac{\begin{bmatrix} z\tilde{m}(z^2)\omega_\ell^2 & -1 \\ -1 & zm(z^2) \end{bmatrix} \otimes I_{j_\ell}}{z^2 m(z^2)\tilde{m}(z^2)\omega_\ell^2 - 1} \times \begin{bmatrix} zm(z^2) U_\ell^* b_1' \\ 0 \end{bmatrix} dz
$$
$$
= \frac{\zeta(\rho_i)}{2\imath\pi\sqrt{n}} \sum_{\ell=1}^{s} \oint_{\gamma_i'} \frac{wm(w)\tilde{m}(w)\omega_\ell^2 b_1^* \left(Q(w) - m(w)I \right) \Pi_\ell b_1'}{wm(w)\tilde{m}(w)\omega_\ell^2 - 1} dw
$$
$$
- \frac{\zeta(\rho_i)}{2\imath\pi\sqrt{n}} \sum_{\ell=1}^{s} \oint_{\gamma_i'} \frac{\omega_\ell m(w) b^* X \widetilde{Q}(w) V_\ell U_\ell^* b_1'}{wm(w)\tilde{m}(w)\omega_\ell^2 - 1} dw
$$
$$
= \frac{b_1^* \left(Q(\rho_i) - m(\rho_i)I \right) \Pi_i b_1'}{\sqrt{n} m(\rho_i)} - \frac{\omega_i b_1^* X \widetilde{Q}(\rho_i) V_i U_i^* b_1'}{\sqrt{n}} \quad \text{w.p. 1 for large } n
$$
$$
\asymp \delta_{i-1,0} \left(\frac{b_1^* \left(Q(\rho_i) - m(\rho_i)I \right) \Pi_1 b_1'}{\sqrt{n} m(\rho_i)} - \frac{\omega_i b_1^* X \widetilde{Q}(\rho_i) V_1 U_1^* b_1'}{\sqrt{n}} \right)
$$

by Corollary 3 and Lemma 8.

The terms $X_{3,i}$

From (10) and (20), we have

$$
X_{3,i} = -\frac{\zeta(\rho_i)}{\imath\pi\sqrt{n}} \oint_{\gamma_i} \sum_{p,\ell=1}^{s} \left[zm(z^2) b_1^* U \quad 0 \right]
$$
$$
\times \frac{\left(\begin{bmatrix} z\tilde{m}(z^2)\omega_p^2 & -1 \\ -1 & zm(z^2) \end{bmatrix} \otimes \mathit{l}_p \right) E(z) \left(\begin{bmatrix} z\tilde{m}(z^2)\omega_\ell^2 & -1 \\ -1 & zm(z^2) \end{bmatrix} \otimes \mathit{l}_\ell \right)}{(z^2 m(z^2)\tilde{m}(z^2)\omega_p^2 - 1)(z^2 m(z^2)\tilde{m}(z^2)\omega_\ell^2 - 1)} \begin{bmatrix} zm(z^2) U^* b_1' \\ 0 \end{bmatrix} dz
$$
$$
= -\frac{\zeta(\rho_i)}{\imath\pi\sqrt{n}} \oint_{\gamma_i} \sum_{p,\ell=1}^{s} \left[\omega_p^2 z^2 m(z^2)\tilde{m}(z^2) \quad -zm(z^2) \right]
$$
$$
\times \frac{\begin{bmatrix} zb_1^* \Pi_p (Q(z^2) - m(z^2)I) \Pi_\ell b_1' & \omega_\ell b_1^* \Pi_p X \widetilde{Q}(z^2) V_\ell U_\ell^* b_1' \\ \omega_p b_1^* U_p V_p^* \widetilde{Q}(z^2) X^* \Pi_\ell b_1' & z\omega_p\omega_\ell b_1^* U_p V_p^* (\widetilde{Q}(z^2) - \tilde{m}(z^2)) V_\ell U_\ell^* b_1' \end{bmatrix}}{(z^2 m(z^2)\tilde{m}(z^2)\omega_p^2 - 1)(z^2 m(z^2)\tilde{m}(z^2)\omega_\ell^2 - 1)} \begin{bmatrix} \omega_\ell^2 z^2 m(z^2)\tilde{m}(z^2) \\ -zm(z^2) \end{bmatrix} dz
$$
$$
= -\frac{\zeta(\rho_i)}{2\imath\pi} \oint_{\gamma_i'} \sum_{p,\ell=1}^{s} \frac{G_{p,\ell}(w)}{(wm(w)\tilde{m}(w)\omega_p^2 - 1)(wm(w)\tilde{m}(w)\omega_\ell^2 - 1)} dw
$$

where

$$
G_{p\ell}(w) = n^{-1/2} \left(\omega_p^2\omega_\ell^2 w^2 m(w)^2 \tilde{m}(w)^2 b_1^* \Pi_p (Q(w) - m(w)I) \Pi_\ell b_1' \right.
$$
$$
- \omega_p\omega_\ell^2 wm(w)^2 \tilde{m}(w) b_1^* U_p V_p^* \widetilde{Q}(w) X^* \Pi_\ell b_1' - \omega_p^2\omega_\ell wm(w)^2 \tilde{m}(w) b_1^* \Pi_p X \widetilde{Q}(w) V_\ell U_\ell^* b_1'
$$
$$
\left. + \omega_p\omega_\ell wm(w)^2 b_1^* U_p V_p^* (\widetilde{Q}(w) - \tilde{m}(w)I) V_\ell U_\ell^* b_1' \right).
$$

For large n and with probability one, the $G_{p\ell}(w)$ are holomorphic functions in a domain enclosing γ_i', and $G_{p\ell}(w)$ does not cancel any of the terms of the denominator. The integrals of all terms in the sum such that $p \neq i$ and $\ell \neq i$ are zero. Each of the integrands of the terms $p = i, \ell \neq i$ or $p \neq i, \ell = i$ has a pole with degree one, and the corresponding integrals are of the form $K_{i\ell}G_{i\ell}(\rho_i)$ or $K_{pi}G_{pi}(\rho_i)$ where the $K_{i\ell}$ and K_{pi} are real constants. By inspecting the expression of $G_{p\ell}$ and by using Corollary 3 and Lemma 8, it can be seen that these terms converge to zero in probability. It remains to study the term $p = \ell = i$, which has a degree 2 pole. Recalling that the residue of a meromorphic function $f(z)$ that has a pole with degree 2

at z_0 is $\lim_{z \to z_0} d\left((z - z_0)^2 f(z)\right)/dz$ and letting $g_\ell(z) = zm(z)\tilde{m}(z)\omega_\ell^2 - 1$, the integral of this term is

$$\zeta(\rho_i)\left(\frac{G_{ii}(\rho_i)g_i''(\rho_i)}{g_i'(\rho_i)^3} - \frac{G_{ii}'(\rho_i)}{g_i'(\rho_i)^2}\right).$$

Thanks to Corollary 3 and Lemma 8, $\Re(G_{ii}(\rho_i)) \xrightarrow{\mathcal{P}} 0$. The same can be said about $G_{ii}'(\rho_i)$ after a simple modification of Proposition 2 and Corollary 3. In conclusion,

$$\forall i = 1, \ldots, s, \quad \Re(X_{3,i}) \xrightarrow{\mathcal{P}} 0.$$

The terms q_i

These are the higher order terms that appear when we expand the right hand side of (18). We shall work here on one of these terms, namely

$$\varepsilon = \frac{\zeta(\rho_i)}{\iota\pi\sqrt{n}} \oint_{\gamma_i} a^*(z, \varphi_1)\left(\widehat{H}(z)^{-1} - H(z)^{-1} + H(z)^{-1}E(z)H(z)^{-1}\right) a_\varphi'(z, \varphi_1)\, dz$$

and show that $\varepsilon \xrightarrow{\mathcal{P}} 0$. The other higher order terms can be handled similarly. Writing $z = \sqrt{\rho_i} + R\exp(2\iota\pi\theta)$ on the circle γ_i, we have

$$|\varepsilon| \leq K\sqrt{n} \int_0^1 \|\widehat{H}(z)^{-1} - H(z)^{-1} + H(z)^{-1}E(z)H(z)^{-1}\|\, d\theta$$

where K is a constant whose value can change from line to line, but which remains independent from n. Let ϕ be a function from $[0, 1]$ to a normed vector space. If ϕ is twice differentiable on $(0, 1)$, then it is known that $\|\phi(1) - \phi(0) - \phi'(0)\| \leq \sup_{t \in (0,1)} 0.5\|\phi''(t)\|$.

Setting $\phi(t) = (H + tE)^{-1}$ and recalling that $\hat{H} = H + E$, we have $\phi(1) = \hat{H}$, $\phi(0) = H$ and $\phi''(t) = (H + tE)^{-1}E(H + tE)^{-1}E(H + tE)^{-1}$; hence

$$\|\widehat{H}(z)^{-1} - H(z)^{-1} + H(z)^{-1}E(z)H(z)^{-1}\| \leq K\|E(z)\|^2$$

for $z \in \gamma_i$. Write $Q - mI = (Q - \alpha I) + (\alpha - m)I$ and $\widetilde{Q} - \tilde{m}I = (\widetilde{Q} - \tilde{\alpha}I) + (\tilde{\alpha} - \tilde{m})I$, and decompose E as defined in (20) as $E = E_1 + E_2$ where

$$E_1(z) = \begin{bmatrix} zU^*(Q(z^2) - \alpha(z^2)I_N)U & U^*X\widetilde{Q}(z^2)V\Omega \\ \Omega V^*\widetilde{Q}(z^2)X^*U & z\Omega V^*(\widetilde{Q}(z^2) - \tilde{\alpha}(z^2)I_n)V\Omega \end{bmatrix},$$

$$E_2(z) = \begin{bmatrix} zU^*((\alpha(z^2) - m(z^2))I_N)U & 0 \\ 0 & z\Omega V^*(\tilde{\alpha}(z^2) - \tilde{m}(z^2))I_n V\Omega \end{bmatrix}.$$

Consider any element of E_1, for instance $zu_1^*(Q(z^2) - \alpha(z^2)I)u_1$. By Lemma 3,

$$\sqrt{n}\mathbb{E}\left(\int_0^1 \mathbb{1}_{\mathcal{O}_n} |u_1^*(Q - \alpha)u_1|^2\, d\theta\right) = \sqrt{n}\int_0^1 \mathbb{E}\mathbb{1}_{\mathcal{O}_n} |u_1^*(Q - \alpha)u_1|^2\, d\theta \leq \frac{K}{\sqrt{n}}$$

which shows that $\sqrt{n}\int_0^1 \|E_1\|^2 d\theta \xrightarrow{\mathcal{P}} 0$.

We now prove that $\sqrt{n}\int_0^1 \|E_2\|^2 d\theta \xrightarrow{\mathcal{P}} 0$. In the space of probability measures on $\mathbb{R}$ endowed with the weak convergence metric, in order to prove that a sequence converges weakly to μ, it is enough to prove that from any sequence, we can extract a subsequence along which the weak convergence to μ holds true. We shall show along this principle that $\sqrt{n}\int_0^1 \|E_2\|^2 d\theta \xrightarrow{\mathcal{P}} 0$. Consider the term $\sqrt{n}(\alpha - m)$. Let (z_k) be a denumerable sequence of points in $\mathbb{C} - [0, \lambda_+]$ with an accumulation point in that set. By A8, from every sequence, there is subsequence n_ℓ such that $\sqrt{n_\ell}(\alpha_{n_\ell}(z_1) - m(z_1)) \to 0$ almost surely (recall that the convergence in probability implies the a.s. convergence along a subsequence). By Cantor's diagonal argument, we can extract a subsequence (call it again n_ℓ) such that $\sqrt{n_\ell}(\alpha_{n_\ell}(z_k) - m(z_k)) \to 0$ almost surely for every k. By the normal family theorem, there is a subsequence along which the function $\sqrt{n_\ell}(\alpha_{n_\ell} - m) \to 0$ uniformly on γ_i a.s. Repeating the argument for $\sqrt{n}(\tilde{\alpha} - \tilde{m})$, there is a subsequence n_ℓ along which $\sqrt{n_\ell}\int_0^1 \|E_2\|^2 d\theta \xrightarrow{a.s.} 0$, hence weakly. Necessarily, $\sqrt{n_\ell}\int_0^1 \|E_2\|^2 d\theta$ converges weakly to zero. Now since the weak convergence to a constant is equivalent to the convergence in probability to the same constant, we obtain the desired result. We have finally shown that:

$$\forall i = 1, \ldots, s, \quad q_i \xrightarrow{\mathcal{P}} 0.$$

Final derivations

Write $\hat{\chi}' = \left[\hat{\chi}'(\varphi_1), \ldots, \hat{\chi}'(\varphi_r)\right]$. Generalizing the previous argument to all the φ_k and gathering the results, we obtain

$$
n^{-1/2}\hat{\chi}' \asymp 2\Re \left[\frac{b_k^* \left(Q(\rho_{i(k)}) - m(\rho_{i(k)})I\right) b_k'}{\sqrt{n}m(\rho_{i(k)})} + \frac{b_k^* \left(Q(\rho_{i(k)}) - m(\rho_{i(k)})I\right) \Pi_{i(k)}b_k'}{\sqrt{n}m(\rho_{i(k)})} \right.
$$
$$
\left. - \frac{\omega_{i(k)}b_k^*U_{i(k)}V_{i(k)}^*\widetilde{Q}(\rho_{i(k)})X^*b_k'}{\sqrt{n}} - \frac{\omega_{i(k)}b_k^*X\widetilde{Q}(\rho_{i(k)})V_{i(k)}U_{i(k)}^*b_k'}{\sqrt{n}} \right]_{k=1}^{r}
$$
$$
\asymp \frac{cD}{\sqrt{3}} \sqrt{n}\Re \left[\frac{b_k^*Q(\rho_{i(k)})b_k^\perp}{m(\rho_{i(k)})} - \omega_{i(k)}b_k^*U_{i(k)}V_{i(k)}^*\widetilde{Q}(\rho_{i(k)})X^*b_k^\perp \right]_{k=1}^{r} .
$$

By Lemma 8, matrix $A = \left[V_{i(k)}U_{i(k)}^*b_k\right]_{k=1}^{r}$ satisfies $A^*A \to I_r$. Recall from the same lemma that $B^*B \to I_r$, $(B^\perp)^*B^\perp \to I_r$ and $(B^\perp)^*B \to 0$. Hence, Proposition 2 can be applied to the r.h.s. of this expression, and $n^{-1/2}\hat{\chi}'$ converges in law to

$$
\mathcal{N}\left(0, \frac{c^2D^2}{6}\mathrm{diag}\left(\frac{m'(\rho_{i(k)}) - m(\rho_{i(k)})^2}{cm(\rho_{i(k)})^2} + \omega_{i(k)}^2 \left(m(\rho_{i(k)}) + \rho_{i(k)}m'(\rho_{i(k)})\right) \right)_{k=1}^{r} \right).
$$

It remains to recall Lemmas 9 and 10 to terminate the proof of Theorem 4.

Appendix A. Proof of Proposition 2

The tightness of ξ_n follows from Lemmas 3 and 4 with $p = 2$ and from the application of Chebyshev's inequality.

Let $Z = [z_{i,k}]_{i,k=1}^{N,t}$ and $\widetilde{Z} = [\tilde{z}_{i,k}]_{i,k=1}^{n,t}$ be $N \times t$ and $n \times t$ standard Gaussian random matrices chosen such that $Z, \widetilde{Z}$ and the $N \times N$ matrix Γ of singular values of X are independent. For $k = 1, \ldots, t/2$, let $D_k = \mathrm{diag}(d_{i,k})_{i=1}^{N} = (\Gamma^2 - \rho_k)^{-1}$ and $C_k = \mathrm{diag}(c_{i,k})_{i=1}^{N} = \Gamma(\Gamma^2 - \rho_k)^{-1}$. Then

$$
\eta_n \overset{\mathscr{D}}{=} \left[\sqrt{N}\left(\left[(Z^*Z)^{-1/2}Z^*\left(D_k - \frac{\mathrm{tr}D_k}{N}\right)Z(Z^*Z)^{-1/2}\right]_{k,k+t/2} \right)_{k=1,\ldots,t/2} \right.,
$$
$$
\left. \times \sqrt{n}\left(\left[(Z^*Z)^{-1/2}Z^*C_k\widetilde{Z}[1;N](\widetilde{Z}^*\widetilde{Z})^{-1/2}\right]_{k,k} \right)_{k=1,\ldots,t/2} \right]^{T}
$$

where $\widetilde{Z}[1;N]$ is $\widetilde{Z}$ truncated to its first N rows. By the Law of Large Numbers, $N^{-1}Z^*Z \to I_t$ and $n^{-1}\widetilde{Z}^*\widetilde{Z} \to I_t$ almost surely. Hence, if we show that the multidimensional random variables $A_{k,n} = N^{-1/2}Z^*(D_k - N^{-1}\mathrm{tr}D_k)Z$ and $B_{k,n} = N^{-1/2}Z^*C_k\widetilde{Z}[1;N]$ are tight for $k = 1, \ldots, t/2$, and

$$
\bar{\eta}_n = \frac{1}{\sqrt{N}}\left[\left(\left[Z^*\left(D_k - \frac{\mathrm{tr}D_k}{N}\right)Z\right]_{k,k+t/2} \right)_{k=1,\ldots,t/2} , \left(\left[Z^*C_k\widetilde{Z}[1;N]\right]_{k,k} \right)_{k=1,\ldots,t/2} \right]^{T}
$$

converges in law towards $\mathcal{CN}(0, R)$, the second result of Proposition 2 is proven. From A3 and A4,

$$
\frac{1}{N}\sum_{i=1}^{N}\left(d_{i,k} - \frac{\mathrm{tr}D_k}{N}\right)^2 = \frac{1}{N}\mathrm{tr}Q(\rho_k)^2 - \left(\frac{1}{N}\mathrm{tr}Q(\rho_k)\right)^2 \xrightarrow[n\to\infty]{\text{a.s.}} m'(\rho_k) - m(\rho_k)^2, \quad \text{and}
$$

$$
\frac{1}{N}\sum_{i=1}^{N}c_{i,k}^2 = \frac{1}{N}\mathrm{tr}Q(\rho_k) + \frac{\rho_k}{N}\mathrm{tr}Q(\rho_k)^2 \xrightarrow[n\to\infty]{\text{a.s.}} m(\rho_k) + \rho_km'(\rho_k)
$$

for all $k = 1, \ldots, t/2$. Recalling that Z and $\widetilde{Z}$ are standard Gaussian, it results that $\limsup_n \mathbb{E}\left[\|A_{k,n}\|^2 \| \Gamma_n\right]$ and $\limsup_n \mathbb{E}\left[\|B_{k,n}\|^2 \| \Gamma_n\right]$ are bounded w.p. 1 by a constant. Tightness of the $A_{k,n}$ and $B_{k,n}$ follows. Now we have

$$
\bar{\eta}_n = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[\left((d_{i,k} - N^{-1}\mathrm{tr}D_k)z_{i,k}^*z_{i,k+t/2}\right)_{k=1,\ldots,t/2} , \left(c_{i,k}z_{i,k}^*\tilde{z}_{i,k}\right)_{k=1,\ldots,t/2} \right]^{T}
$$
$$
= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{u}_{i,n}.
$$

Observe that the covariance matrix of $\bar{\eta}_n$ conditional to Γ_n converges almost surely to R. Moreover, thanks to A4, it is easy to see that the Lyapunov condition

$$\frac{1}{N^{1+a}} \sum_{i=1}^{n} \mathbb{E}\left[\|\mathbf{u}_{i,n}\|^{2(1+a)} \mid \Gamma_n\right] \xrightarrow[n\to\infty]{\text{a.s.}} 0$$

is satisfied for any $a > 0$; hence $\bar{\eta}_n \xrightarrow{\mathcal{L}} \mathcal{CN}(0, R)$ which completes the proof of Proposition 2.

Appendix B. Sketch of the proof of Proposition 3

For $k = 1, \ldots, r$, let $\bar{\rho}_{k,n}$ be the solutions of the equation $\omega_{k,n}^2 g(\rho) = 1$, where we recall that the $\omega_{k,n}^2$ are the diagonal elements of matrix Ω_n. Then, by a simple extension to the case $r \geq 1$ of the proof of [9, Theorem 2.15], one can show that the sequences $\sqrt{n}(\hat{\lambda}_{k,n} - \bar{\rho}_{k,n})$ are tight. To obtain the result, we show that $\sqrt{n}(\bar{\rho}_{k,n} - \rho_{i(k)}) = \mathcal{O}(1)$. Since g is decreasing, this amounts to showing that $\sqrt{n}(\omega_{k,n}^2 - \omega_{i(k)}^2) = \mathcal{O}(1)$. Since the non-zero eigenvalues of PP^* coincide with those of B^*BS^*S, it will be enough to prove that $\sqrt{n}(B^*BS^*S - O) = \mathcal{O}(1)$. It is clear that $B^*B = I_r + n^{-1}A$ where $\sup_n \|A\| < \infty$; hence $\sqrt{n}(B^*BO - O) \to 0$. By the last item in Assumption A6, $\sqrt{n}B^*B(S^*S - O) = \mathcal{O}(1)$, and the proposition is shown.

Appendix C. Proof of Lemma 8

Observing that

$$b'(\varphi) = \frac{-\imath D}{\sqrt{N}} \left[\ell \exp(-\imath D\ell\varphi)\right]_{\ell=0}^{N-1} \quad \text{and} \quad b''(\varphi) = \frac{-D^2}{\sqrt{N}} \left[\ell^2 \exp(-\imath D\ell\varphi)\right]_{\ell=0}^{N-1},$$

and using the fact that $N^{-(K+1)} \sum_{\ell=0}^{N-1} \ell^K \exp(\imath\alpha\ell) \to \delta_{\alpha,0}/(K+1)$ for $\alpha \in [-\pi, \pi]$, we have $B^*B \to I_r$, $n^{-1}B^*B' \to -(\imath cD/2)I_r$, $n^{-2}(B')^*B' \to (c^2D^2/3)I_r$, and $n^{-2}B^*B'' \to -(c^2D^2/3)I_r$.

Writing $B^\perp = 2\sqrt{3}(ncD)^{-1}B' + \imath\sqrt{3}B$ and replacing in the above convergences, the stated properties of $B^\perp$ become straightforward.

We now show the last convergence. Assume without loss of generality that $i = 1$ and recall that $S^*S \to O^2$. Consider the isometry matrices $W = B(B^*B)^{-1/2}$ and $Z = S(S^*S)^{-1/2}$, and let $A = (B^*B)^{1/2}(S^*S)^{1/2}$, resulting in $P = WAZ^*$. Notice that the singular values of A coincide with those of P apart from the zeros. Let π_1 be the orthogonal projection matrix on the eigenspace of AA^* associated with the eigenvalues $\omega_{1,n}^2, \ldots, \omega_{j_1,n}^2$. With these notations, $\Pi_1 = W\pi_1W^*$ and $\Pi_{B_1} = B_1(B_1^*B_1)^{-1}B_1^*$. We have $A \to O$; hence $\pi_1 \to \begin{bmatrix} I_{j_1} & 0 \\ 0 & 0 \end{bmatrix}$. Since $B^*B \to I$, for any vector x such that $\|x\| = 1$, we have $x^*\Pi_1 x - x^*B_1B_1^*x \to 0$, and $x^*\Pi_{B_1}x - x^*B_1B_1^*x \to 0$. Therefore, $x^*(\Pi_1 - \Pi_{B_1})x \to 0$, which proves the last result.

References

[1] Z.D. Bai, Methodologies in spectral analysis of large-dimensional random matrices, a review, Statist. Sinica 9 (3) (1999) 611–677. With comments by G.J. Rodgers and Jack W. Silverstein; and a rejoinder by the author.

[2] Z.D. Bai, J.W. Silverstein, CLT for linear spectral statistics of large-dimensional sample covariance matrices, Ann. Probab. 32 (1A) (2004) 553–605.

[3] Z. Bai, J.-f. Yao, Central limit theorems for eigenvalues in a spiked population model, Ann. Inst. H. Poincaré Probab. Statist. 44 (3) (2008) 447–474.

[4] J. Baik, G. Ben Arous, S. Péché, Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices, Ann. Probab. 33 (5) (2005) 1643–1697.

[5] J. Baik, J.W. Silverstein, Eigenvalues of large sample covariance matrices of spiked population models, J. Multivariate Anal. 97 (6) (2006) 1382–1408.

[6] F. Benaych-Georges, A. Guionnet, M. Maïda, Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices, January 2010. Arxiv Preprint arXiv:1009.0145.

[7] F. Benaych-Georges, R.R. Nadakuditi, The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices (v1), October 2009. ArXiv e-prints.

[8] F. Benaych-Georges, R.R. Nadakuditi, The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices, Adv. Math. (2011).

[9] F. Benaych-Georges, R.R. Nadakuditi, The singular values and vectors of low rank perturbations of large rectangular random matrices, March 2011. ArXiv e-prints.

[10] P. Bianchi, M. Debbah, M. Maida, J. Najim, Performance of statistical tests for single-source detection using random matrix theory, IEEE Trans. Inform. Theory 57 (4) (2011) 2400–2419.

[11] G. Bienvenu, L. Kopp, Adaptivity to background noise spatial coherence for high resolution passive methods, in: IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP'80, vol. 5, April 1980, pp. 307–310.

[12] M. Capitaine, C. Donati-Martin, D. Féral, Central limit theorems for eigenvalues of deformations of Wigner matrices, March 2009. Arxiv Preprint arXiv:0903.4740.

[13] M. Capitaine, C. Donati-Martin, D. Féral, The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations, Ann. Probab. 37 (1) (2009) 1–47.

[14] P. Ciblat, P. Loubaton, E. Serpedin, G.B. Giannakis, Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators, IEEE Trans. Inform. Theory 48 (7) (2002) 1922–1934.

[15] S. Geman, A limit theorem for the norm of random matrices, Ann. Probab. 8 (2) (1980) 252–261.

[16] E.J. Hannan, Non-linear time series regression, J. Appl. Probab. 8 (1971) 767–780.

[17] E.J. Hannan, The estimation of frequency, J. Appl. Probab. 10 (1973) 510–519.

[18] Fumio Hiai, Dénes Petz, The Semicircle Law, Free Random Variables and Entropy, in: Mathematical Surveys and Monographs, vol. 77, American Mathematical Society, Providence, RI, 2000.

[19] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 2007.
[20] I.M. Johnstone, On the distribution of the largest eigenvalue in principal components analysis, Ann. Statist. 29 (2) (2001) 295–327.
[21] V.A. Marčenko, L.A. Pastur, Distribution of eigenvalues in certain sets of random matrices, Mat. Sb. (NS) 72 (114) (1967) 507–536.
[22] Boaz Nadler, On the distribution of the ratio of the largest eigenvalue to the trace of a Wishart matrix, J. Multivariate Anal. 102 (2) (2011) 363–371.
[23] Leonid Pastur, Mariya Shcherbina, Eigenvalue Distribution of Large Random Matrices, in: Mathematical Surveys and Monographs, vol. 171, American Mathematical Society, Providence, RI, 2011.
[24] L. Pastur, V. Vasilchuk, On the law of addition of random matrices: covariance and the central limit theorem for traces of resolvent, in: Probability and Mathematical Physics, in: CRM Proc. Lecture Notes, vol. 42, Amer. Math. Soc., Providence, RI, 2007, pp. 399–416.
[25] D. Paul, Asymptotics of sample eigenstructure for a large dimensional spiked covariance model, Statist. Sinica 17 (4) (2007) 1617–1642.
[26] S. Péché, The largest eigenvalue of small rank perturbations of Hermitian random matrices, Probab. Theory Related Fields 134 (1) (2006) 127–173.
[27] R. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antennas and Propagation 34 (3) (1986) 276–280.
[28] P. Stoica, A. Nehorai, MUSIC, maximum likelihood, and Cramer–Rao bound, IEEE Trans. Acoust. Speech Signal Process. 37 (5) (1989) 720–741.