ON THE CONSISTENCY OF THE G-MUSIC DOA ESTIMATOR

P.Vallet [2], W.Hachem [1], P.Loubaton [2], X.Mestre [3], J.Najim [1]

[1] Telecom Paristech (CNRS UMR 5141), 46 rue Barrault 75634 Paris (France)
[2] IGM (CNRS UMR 8049), 5 Bd. Descartes 77454 Marne-la-Vallée (France)
[3] CTTC, Av. Carl Friedrich Gauss 08860 Castelldefels, Barcelona (Spain)
{hachem, najim}@telecom-paristech.fr, {loubaton,vallet}@univ-mlv.fr, xavier.mestre@cttc.cat

ABSTRACT

Recently, a new subspace DoA estimation method (called "G-MUSIC") has been proposed, in the context where the number of available snapshots N is of the same order of magnitude than the number of sensors M. In this context, the traditional subspace methods tend to fail because the empirical covariance matrix of the observations is a poor estimate of the true covariance matrix. G-MUSIC is based on a new consistent estimator of the localization function in the regime where M and N tend to $+\infty$ at the same rate. However, the consistency of the DoA estimator was not adressed. The purpose of this paper is to prove the consistency of the angles estimator in the previous asymptotic regime.

***Index Terms*—** Random matrices, MUSIC, DoA estimation.

1. INTRODUCTION

Subspace DoA (Directions of Arrival) estimation methods using antenna arrays (such as MUSIC) has been extensively studied in the past because they offer a good trade-off between performance and complexity. Their statistical performance has mainly been characterized when the number of snapshots N converges to $+\infty$ while the number of antennas M remains fixed. In practice, the corresponding conclusions are valid in finite sample size if N is much larger than M. However, this assumption is often not realistic if the number of antennas is large because the number of available snapshots may be limited. In order to study the statistical performance of the subspace DoA estimates in this context, Mestre et al. [1] proposed to consider the asymptotic regime in which M and N converge to $+\infty$ at the same rate, i.e. $M, N \to +\infty$, $\frac{M}{N}$ converges towards a positive constant. Using large random matrix theory (RMT) results, [1] proved that the traditional DoA subspace estimators are asymptotically biased, and proposed, assuming the source signals are i.i.d, a consistent estimator of the localization function which outperforms the standard ones for realistic values of M and N. Later, [2] proposed a similar estimator assuming the source signals are unknown deterministic quantities. Neither one of these papers imposes restrictive conditions on the number of signals K, which may scale-up with M, N. In [2], the consistency of the localization function estimator was proved, but the consistency of the DoA estimator was not adressed. Therefore, the purpose of this paper is to continue the study of this DoA estimator, and prove its consistency, still in the regime where $M, N \to +\infty$ at the same rate. In this context, the key point is to prove that the localization function estimator derived in [2] is uniformly consistent.

The paper is organized as follows. In section 2, we introduce the model and notations used in the paper. In section 3, we review some basic results concerning the G-MUSIC method introduced in [2]. In section 4, we prove the consistency of the DoA estimator.

2. MODEL AND PROBLEM STATEMENT

We consider K narrow band deterministic source signals impinging on an antennas array of M elements with $K < M$. At time n, the received snapshot (of size M) writes

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n, \tag{1}$$

where $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ is the $M \times K$ matrix containing the K linearly independent steering vectors, $\mathbf{v}_n$ is a Gaussian vector satisfying $\mathbb{E}[\mathbf{v}_n\mathbf{v}_n^H] = \sigma^2\mathbf{I}_M$, and $\mathbf{s}_n$ is a vector containing the K source signals received at time n. Assuming we collect N snapshots, (1) is equivalent to

$$\mathbf{\Sigma}_N = \mathbf{B}_N + \mathbf{W}_N,$$

with $\mathbf{\Sigma}_N = \frac{1}{\sqrt{N}}[\mathbf{y}_1, \ldots, \mathbf{y}_N]$, $\mathbf{B}_N = \frac{1}{\sqrt{N}}[\mathbf{A}\mathbf{s}_1, \ldots, \mathbf{A}\mathbf{s}_N]$, and $\mathbf{W}_N$ built as $\mathbf{\Sigma}_N$ and $\mathbf{B}_N$. In this paper, we will assume that $M < N$ and that matrix $\mathbf{S}_N = [\mathbf{s}_1, \ldots, \mathbf{s}_N]$ has full rank K. Therefore $\mathbf{B}_N\mathbf{B}_N^H$ has rank K, and we assume that its positive eigenvalues have multiplicity one and are denoted by $0 = \lambda_{1,N} = \ldots = \lambda_{M-K,N} < \lambda_{M-K+1,N} < \ldots < \lambda_{M,N}$. We also denote by $\mathbf{u}_{1,N}, \ldots, \mathbf{u}_{M,N}$ the corresponding unit norm eigenvectors.

The MUSIC algorithm is based on the fact that the K angles $\theta_1, \ldots, \theta_K$ are solutions to the equation $\eta_N(\theta) = 0$, with

$\eta_N(\theta)$ the localization function defined by

$$\eta_N(\theta) = \sum_{k=1}^{M-K} \mathbf{a}(\theta)^H \mathbf{u}_{k,N} \mathbf{u}_{k,N}^H \mathbf{a}(\theta),$$

where $\sum_{k=1}^{M-K} \mathbf{u}_{k,N}\mathbf{u}_{k,N}^H$ is the projection matrix onto the "noise subspace", defined as the orthogonal complement of the column space of matrix $\mathbf{A}\mathbf{A}^H$. To estimate these angles, we need to estimate the quantity $\eta_N(\theta)$, and the traditional approach consists in replacing the "noise" eigenvectors $\mathbf{u}_{1,N},\ldots,\mathbf{u}_{M-K,N}$ of $\mathbf{B}_N\mathbf{B}_N^H$ by those of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^H$, the empirical covariance matrix of the observations.

We denote by $\hat{\lambda}_{1,N} \leq \ldots \leq \hat{\lambda}_{M,N}$ the eigenvalues of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^H$ and by $\hat{\mathbf{u}}_{1,M}, \ldots, \hat{\mathbf{u}}_{M,N}$ the corresponding unit norm eigenvectors. The traditional estimator for the localization function thus writes

$$\hat{\eta}_{trad,N}(\theta) = \sum_{k=1}^{M-K} \mathbf{a}(\theta)^H \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^H \mathbf{a}(\theta). \qquad (2)$$

Thanks to law of large number, the previous estimator is consistent in the case where $N \to \infty$ while M is kept constant, i.e $c_N = M/N \to 0$. However, when $M, N \to +\infty$ in such a way that $c_N \to c > 0$, the previous estimator does not converge anymore to η_N.

For the remainder of the paper, all convergences will be considered under the regime $c_N \to c \in]0,1[$, as $M, N \to \infty$, and will be directly referred by "$N \to +\infty$".

3. REVIEW OF THE G-MUSIC METHOD

We begin by stating some classical results in RMT. Define $\mathbf{Q}_N(z) = \left(\mathbf{\Sigma}_N\mathbf{\Sigma}_N^H - z\mathbf{I}_M\right)^{-1}$ the resolvent of matrix $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^H$, and its normalized trace by $\hat{m}_N(z) = \frac{1}{M}\operatorname{Tr}\mathbf{Q}_N(z)$. Define the probability measure $\hat{\mu}_N = \frac{1}{M}\sum_{k=1}^M \delta_{\hat{\lambda}_{k,N}}$, with δ_a the Dirac measure at point a. $\hat{\mu}_N$ is usually called the "empirical spectral measure" of matrix $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^H$. Then $\hat{m}_N(z)$ is the Stieltjes transform of $\hat{\mu}_N$, i.e $\hat{m}_N(z) = \int_{\mathbb{R}} \frac{\mathrm{d}\hat{\mu}_N(\lambda)}{\lambda - z}$.

Theorem 1 ([3]). *Let $z \in \mathbb{C}\backslash\mathbb{R}^+$. Then as $N \to \infty$,*

$$\hat{m}_N(z) - m_N(z) \to 0 \quad a.s,$$

where $m_N(z)$ is the Stieltjes transform of a deterministic probability measure μ_N, i.e $m_N(z) = \int_{\mathbb{R}} \frac{\mathrm{d}\mu_N(\lambda)}{\lambda - z}$, and satisfies the equation

$$\begin{aligned} m_N(z) &= \frac{1}{M}\operatorname{Tr}\mathbf{T}_N(z) \\ &= (1+\sigma^2 c_N m_N(z))\frac{1}{M}\operatorname{Tr}\left(\mathbf{B}_N\mathbf{B}_N^H - w_N(z)\right)^{-1}, \end{aligned}$$

with $w_N(z)$ defined by

$$w_N(z) = z(1+\sigma^2 c_N m_N(z))^2 - \sigma^2 c_N(1 - \sigma^2 c_N m_N(z)).$$

The previous theorem implies that $\hat{\mu}_N - \mu_N \to 0$ weakly, w.p.1. as $N \to \infty$. Thus μ_N represents a deterministic approximation of measure $\hat{\mu}_N$. From [2], we also have the almost-sure convergence $\mathbf{a}(\theta)^H\left(\mathbf{Q}_N(z) - \mathbf{T}_N(z)\right)\mathbf{a}(\theta) \to 0$ and $\hat{w}'_N(z) - w'_N(z) \to 0$ as $N \to \infty$, where $\hat{w}'_N(z)$ is the derivative of $\hat{w}_N(z)$ given by

$$\begin{aligned} \hat{w}_N(z) &= z\left(1+\sigma^2 c_N \hat{m}_N(z)\right)^2 \\ &\quad - \sigma^2(1-c_N)\left(1+\sigma^2 c_N \hat{m}_N(z)\right). \qquad (3) \end{aligned}$$

The support of μ_N provides some insights on the localization of the eigenvalues $\{\hat{\lambda}_{k,N} : k = 1,\ldots,M\}$. It is proved in [2] that this support is the union of a finite number (say Q) of disjoint compact intervals, denoted $[x_{q,N}^-, x_{q,N}^+]$. Before giving results concerning the G-MUSIC method, two assumptions are required, basically expressing the fact that the "signal subspace" must be well separated from the "noise subspace". Firstly, it is assumed that $\exists t_1^-, t_1^+, t_2^-, t_2^+ \in \mathbb{R}_*^+$, independent of N, such that for all large N,

$$t_1^- < x_{1,N}^- < x_{1,N}^+ < t_1^+ < t_2^- < x_{2,N}^- < \ldots < x_{Q,N}^- < t_2^+.$$

Secondly, no eigenvalue of $\mathbf{B}_N\mathbf{B}_N^H$, except 0, must belong to the interval $]w_N(t_1^-), w_N(t_1^+)[$. These assumptions, known as "eigenvalue separation condition", imply that a.s., for all $\epsilon > 0$, $\hat{\lambda}_{1,N},\ldots,\hat{\lambda}_{M-K,N} \in [t_1^- - \epsilon, t_1^+ + \epsilon]$ and $\hat{\lambda}_{M-K+1,N},\ldots,\hat{\lambda}_{M,N} \in [t_2^- - \epsilon, t_2^+ + \epsilon]$, for all large N. Choose $\epsilon > 0$ such that $t_1^- - 6\epsilon > 0$ and $t_1^+ + 6\epsilon < t_2^-$, and define $\mathcal{R}_y$ ($y > 0$) the rectangle

$$\mathcal{R}_y = \{u + \mathbf{i}v : u \in [t_1^- - 3\epsilon, t_1^+ + 3\epsilon], v \in [-y, y]\}.$$

Denote $\partial\mathcal{R}_y^-$ its boundary, clockwise oriented. From [2], under these assumptions, for all large N, we can express η_N as

$$\eta_N(\theta) = \frac{1}{2\pi\mathbf{i}}\oint_{\partial\mathcal{R}_y^-} \mathbf{a}(\theta)^H\mathbf{T}_N(z)\mathbf{a}(\theta)\frac{w'_N(z)}{1+\sigma^2 c_N m_N(z)}\mathrm{d}z.$$

In this context, the G-MUSIC method, proposed in [2], is based on the following estimator of the localization function,

$$\hat{\eta}_N(\theta) = \frac{1}{2\pi\mathbf{i}}\oint_{\partial\mathcal{R}_y^-} \mathbf{a}(\theta)^H\mathbf{Q}_N(z)\mathbf{a}(\theta)\frac{\hat{w}'_N(z)}{1+\sigma^2 c_N \hat{m}_N(z)}\mathrm{d}z. \qquad (4)$$

In practice, the contour $\partial\mathcal{R}_y$ is not known, but if the number of sources K is available, the integral in (4) can be solved using residue theorem, to obtain an explicit formula depending only on the observations matrix $\mathbf{\Sigma}_N$. For this, we have to study the location of the poles of the integrand in (4), i.e $\{\hat{\lambda}_{k,N} : k = 1,\ldots,M\}$ and the zeros of $1+\sigma^2 c_N \hat{m}_N(z)$. Following the idea of [3], we define the matrix $\hat{\mathbf{\Omega}}_N = \hat{\mathbf{\Lambda}}_N + \frac{\sigma^2 c_N}{M}\mathbf{1}\mathbf{1}^T$, with $\mathbf{1} = [1,\ldots,1]^T$ and $\hat{\omega}_{1,N} \leq \ldots \leq \hat{\omega}_{M,N}$ its eigenvalues. $\hat{\mathbf{\Lambda}}_N$ is the matrix of eigenvalues of $\mathbf{\Sigma}_N\mathbf{\Sigma}_N^H$. It can be shown that the zeros of $1+\sigma^2 c_N \hat{m}_N(z)$ (the denominator in (4)) are eigenvalues of $\hat{\mathbf{\Omega}}_N$. It is proved in [2] that the "eigenvalue separation condition" implies that a.s., for all $\epsilon > 0$, $\hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M-K,N} \in [t_1^- - \epsilon, t_1^+ + \epsilon]$ and $\hat{\omega}_{M-K+1,N}, \ldots, \hat{\omega}_{M,N} \in [t_2^- - \epsilon, t_2^+ + \epsilon]$, for all large N. Thus, the poles $\{\hat{\lambda}_{k,N}, \hat{\omega}_{k,N} : k = 1, \ldots, M-K\}$ are enclosed by $\partial\mathcal{R}_y$, w.p.1. for all large N, which ensures that the integral (4) can be solved.

For a fixed angle θ, the estimator (4) is consistent when $M, N \to \infty$ while $c_N \to c \in]0,1[$. Note that at this step, K may be constant or scale-up with N, but in this paper, we will assume K constant, to lighten the proof in section 4. This assumption can be weakened and we can also prove the consistency if K increases with N, at an appropriate rate.

4. CONSISTENCY OF THE DOA

In this section, we derive the consistency of the DoA estimator, in the case where the steering vectors follow the model $\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}[1, e^{\mathbf{i}\theta}, \ldots, e^{\mathbf{i}(M-1)\theta}]$, which corresponds to a uniform linear array of antennas, whose elements are located at half the wavelength [1] . The proof of the consistency splits in two steps. We first derive a uniform consistency for the localization function estimator, and we transfer the consistency to the angle estimator by a standard argument. The uniform consistency requires the use of a regularization trick to force the poles of the integrand in (4) to be confined in a certain set. Define the set $\mathcal{T}_\epsilon = [t_1^- - \epsilon, t_1^+ + \epsilon] \cup [t_2^- - \epsilon, t_2^+ + \epsilon]$ and the event $\mathcal{A} = \{\hat{\lambda}_{1,N}, \ldots, \hat{\lambda}_{M,N}, \hat{\omega}_{1,N}, \ldots, \hat{\omega}_{M,N} \in \mathcal{T}_\epsilon\}$. It turns out that $\mathbb{P}(\mathcal{A}^c) = \mathcal{O}\left(\frac{1}{N^l}\right)$ for all $l \in \mathbb{N}$ (the proof is omitted due to lack of space). This result allows to control the moments of the integrand in (4). Indeed, let $\phi \in \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R}^+)$ (the space of non-negative test functions on $\mathbb{R}$), such that

$$\phi(t) = \begin{cases} 1 & \text{for} \quad t \in \mathcal{T}_\epsilon \\ 0 & \text{for} \quad t \in \mathbb{R}\backslash\{\mathcal{T}_{2\epsilon}\} \end{cases} .$$

Define [2] $\hat{\chi}_N = \det \phi(\mathbf{\Sigma}_N \mathbf{\Sigma}_N^*) \det \phi(\hat{\mathbf{\Omega}}_N)$. Note that $\hat{\chi}_N = 1$ on the event $\mathcal{A}$. When z approaches the real axis on the contour $\partial\mathcal{R}_y$, the integrand in (4) can be unbounded on the event $\mathcal{A}^c$, due to the poles at $\{\hat{\lambda}_{k,N}, \hat{\omega}_{k,N} : k = 1, \ldots, M\}$. However, when it is multiplied ("regularized") by $\hat{\chi}_N$, the integrand is bounded on $\partial\mathcal{R}_y$, and thus its moments are all finite.

Proposition 1. *For all $l \in \mathbb{N}$,*

$$\mathbb{E}\left|\mathbf{a}(\theta)^H \left(\mathbf{Q}_N(z) \frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} \hat{\chi}_N - \mathbf{T}_N(z) \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)}\right) \mathbf{a}(\theta)\right|^{2l} \leq \frac{1}{N^l} \mathrm{P}_1(|z|)\, \mathrm{P}_2\left(\frac{1}{\mathrm{dist}(z, \mathrm{supp}(\phi))}\right).$$

where $\mathrm{P}_1, \mathrm{P}_2$ are two polynomials with positive coefficients independent of N, θ.

Proof. The proof is omitted here due to lack of space. It widely relies on the Gaussian tools used in [2], and on the fact that the regularization does not introduce any change in the rythm of convergence of the moments of $\hat{w}'_N(z) - w'_N(z)$ and $\mathbf{a}(\theta)\left(\mathbf{Q}_N(z) - \mathbf{T}_N(z)\right)\mathbf{a}(\theta)$ (since $\mathbb{P}(\mathcal{A}^c) = \mathcal{O}(\frac{1}{N^l})$). □

As a consequence, we can upper-bound the moments of the localization function estimator regularized by the quantity $\hat{\chi}_N$. More precisely, using Jensen's inequality and Fubini's theorem, for each $l \in \mathbb{N}$, we obtain

$$\begin{aligned}
&\mathbb{E}\left|\hat{\eta}_N(\theta)\hat{\chi}_N - \eta_N(\theta)\right|^{2l} \\
&\leq \frac{1}{(2\pi)^{2l}} \oint_{\partial\mathcal{R}_y^-} \mathbb{E}\left|\mathbf{a}(\theta)^H \left(\mathbf{Q}_N(z) \frac{\hat{w}'_N(z)}{1 + \sigma^2 c_N \hat{m}_N(z)} \hat{\chi}_N - \mathbf{T}_N(z) \frac{w'_N(z)}{1 + \sigma^2 c_N m_N(z)}\right) \mathbf{a}(\theta)\right|^{2l} |\mathrm{d}z|^{2l} \\
&= \mathcal{O}\left(\frac{1}{N^l}\right).
\end{aligned}$$

With these results, we can now prove the uniform consistency of the estimator of the localization function.

Proposition 2. *With probability one,*

$$\sup_{\theta \in [-\pi,\pi]} |\hat{\eta}_N(\theta) - \eta_N(\theta)| \xrightarrow[N \to \infty]{} 0. \tag{5}$$

Proof. Consider a set ϑ_N of N^2 points denoted $\Theta_1, \ldots, \Theta_{N^2}$, evenly spaced in the interval $[-\pi, \pi]$. Fix $\theta \in [-\pi, \pi]$, and let (θ_N) a sequence such that $\theta_N \in \vartheta_N$ and $|\theta - \theta_N| = \mathcal{O}\left(\frac{1}{N^2}\right)$. We consider the decomposition,

$$\begin{aligned}
\hat{\eta}_N(\theta) - \eta_N(\theta) = \hat{\eta}_N(\theta) - \hat{\eta}_N(\theta_N) + \hat{\eta}_N(\theta_N) - \eta_N(\theta_N) \\
+ \eta_N(\theta_N) - \eta_N(\theta).
\end{aligned} \tag{6}$$

The third term of the RHS of (6) is deterministic and easy to compute. Since $\sup_{\theta \in [-\pi,\pi]} \|\mathbf{a}(\theta) - \mathbf{a}(\theta_N)\| = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$, we get $\sup_{\theta \in [-\pi,\pi]} |\eta_N(\theta_N) - \eta_N(\theta)| = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$. For the second term, we write

$$\mathbb{P}\left(\sup_{\nu \in \vartheta_N} |\hat{\eta}_N(\nu) - \eta_N(\nu)| \geq \epsilon\right) \leq \sum_{k=1}^{N^2} \mathbb{P}\left(|\hat{\eta}_N(\Theta_k) - \eta_N(\Theta_k)| \geq \epsilon\right).$$

Choose $l \geq 4$. Then $\mathbb{P}(\mathcal{A}^c) = \mathcal{O}\left(\frac{1}{N^l}\right)$ and we get

$$\sum_{k=1}^{N^2} \mathbb{P}\left(\{|\hat{\eta}_N(\Theta_k) - \eta_N(\Theta_k)| \geq \epsilon\} \cap \mathcal{A}^c\right) = \mathcal{O}\left(\frac{1}{N^{l-2}}\right).$$

[1] In fact, the precise model for the components of $\mathbf{a}(\theta)$ is $e^{\mathbf{i}k\pi\sin(\theta)}$.

[2] By applying the function ϕ to a Hermitian matrix, we implicitly represent the action of ϕ on the corresponding eigenvalues.

Using the fact that $\hat{\eta}_N = \hat{\eta}_N \hat{\chi}_N$ on $\mathcal{A}$,

$$\begin{aligned}&\mathbb{P}\left(\{|\hat{\eta}_N(\Theta_k) - \eta_N(\Theta_k)| \geq \epsilon\} \cap \mathcal{A}\right) \\ &\quad \leq \mathbb{P}\left(|\hat{\eta}_N(\Theta_k)\hat{\chi}_N - \eta_N(\Theta_k)| \geq \epsilon\right) \\ &\quad \leq \frac{1}{\epsilon^{2l}} \mathbb{E}\left|\hat{\eta}_N(\Theta_k)\hat{\chi}_N - \eta_N(\Theta_k)\right|^{2l}.\end{aligned}$$

This last term is $\mathcal{O}\left(\frac{1}{N^l}\right)$ from the above bounds, and thus,

$$\mathbb{P}\left(\sup_{\nu \in \vartheta_N} |\hat{\eta}_N(\nu) - \eta_N(\nu)| \geq \epsilon\right) = \mathcal{O}\left(\frac{1}{N^{l-2}}\right).$$

Therefore, since $l \geq 4$, we deduce with Borel-Cantelli lemma that $\sup_{\nu \in \vartheta_N} |\hat{\eta}_N(\nu) - \eta_N(\nu)| \to 0$ a.s., as $N \to \infty$. For the first term in (6), using as previously the event $\mathcal{A}$,

$$\begin{aligned}&\mathbb{P}\left(\sup_{\theta \in [-\pi,\pi]} |\hat{\eta}_N(\theta) - \hat{\eta}_N(\theta_N)| \geq \epsilon\right) \\ &\leq \mathbb{P}\left(\sup_{\theta \in [-\pi,\pi]} |\hat{\eta}_N(\theta) - \hat{\eta}_N(\theta_N)|\, \hat{\chi}_N \geq \epsilon\right) + \mathbb{P}\left(\mathcal{A}^c\right), \\ &\leq \frac{1}{\epsilon^l} \mathbb{E}\left[\sup_{\theta \in [-\pi,\pi]} |\hat{\eta}_N(\theta) - \hat{\eta}_N(\theta_N)|^l\, \hat{\chi}_N\right] + \mathcal{O}\left(\frac{1}{N^l}\right).\end{aligned}$$

By Jensen's inequality,

$$\begin{aligned}&|\hat{\eta}_N(\theta) - \hat{\eta}_N(\theta_N)|^l\, \hat{\chi}_N \\ &\quad \leq C \oint_{\partial \mathcal{R}_y^-} \|\mathbf{a}(\theta) - \mathbf{a}(\theta_N)\|^l \, \|\mathbf{Q}_N(z)\|^l \\ &\qquad \times \left|\frac{\hat{w}_N'(z)}{1 + \sigma^2 c_N \hat{m}_N(z)}\right|^l \hat{\chi}_N \, |\mathrm{d}z|^l, \\ &\quad \leq \frac{1}{N^{l/2}} \oint_{\partial \mathcal{R}_y^-} \mathrm{P}_1(|z|)\, \mathrm{P}_2\left(\frac{1}{\mathrm{dist}(z, \mathrm{supp}(\phi))}\right) |\mathrm{d}z|^l, \\ &\quad = \mathcal{O}\left(\frac{1}{N^{l/2}}\right),\end{aligned}$$

with $C > 0$ a constant and $\mathrm{P}_1, \mathrm{P}_2$ two polynomials with positive coefficients, all non-random and independent of N, θ. This eventually shows, using again Borel-Cantelli lemma, that $\sup_{\theta \in [-\pi,\pi]} |\hat{\eta}_N(\theta) - \hat{\eta}_N(\theta_N)| \to 0$ w.p.1. □

We now adress the consistency of the DoA estimate. For this, we follow the approach of [4].

Lemma 1. *Let (α_M) a real-valued sequence of a compact subset of $(-0.5, 0.5]$, and converging to α as $M \to \infty$. Define $q_M(\alpha_M) = \frac{1}{M}\sum_{k=1}^M e^{-\mathbf{i}2\pi k \alpha_M}$. If $\alpha \neq 0$ or if $\alpha = 0$ and $M|\alpha_M| \to \infty$, then $q_M(\alpha_M) \to 0$. If $\alpha = 0$ and $M\alpha_M \xrightarrow[M \to \infty]{} \beta \in \mathbb{R}$, then $q_M(\alpha_M) \to e^{\mathbf{i}\beta}\mathrm{sinc}(\beta)$.*

Choose K disjoint intervals $\mathcal{I}_1, \ldots, \mathcal{I}_K$, such that $\theta_k \in \mathcal{I}_k$. Denote by $\hat{\theta}_{k,N}$ the G-MUSIC estimate of θ_k, defined as

$$\hat{\theta}_{k,N} = \arg\min_{\theta \in \mathcal{I}_k} |\hat{\eta}_N(\theta)|.$$

Theorem 2. *For $k = 1, \ldots, K$, with probability one,*

$$N(\hat{\theta}_{k,N} - \theta_k) \xrightarrow[N \to \infty]{} 0.$$

Proof. Since $\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ is the projection matrix onto the "signal subspace", the true localization function can be written as $\eta_N(\theta) = 1 - \mathbf{a}(\theta)^H \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{a}(\theta)$. We consider the estimation of angle θ_k. By definition, $|\hat{\eta}_N(\hat{\theta}_{k,N})| \leq |\hat{\eta}_N(\theta_k)|$. From (5) and the equality $\eta_N(\theta_k) = 0$, we have $|\hat{\eta}_N(\hat{\theta}_{k,N})| \to 0$ w.p.1., as $N \to \infty$. Consequently,

$$\begin{aligned}|\eta_N(\hat{\theta}_{k,N})| &\leq |\eta_N(\hat{\theta}_{k,N}) - \hat{\eta}_N(\hat{\theta}_{k,N})| + |\hat{\eta}_N(\hat{\theta}_{k,N})| \\ &\leq \sup_{\theta \in [-\pi,\pi]} |\eta_N(\theta) - \hat{\eta}_N(\theta)| + |\hat{\eta}_N(\hat{\theta}_{k,N})| \\ &\xrightarrow[N \to \infty]{a.s.} 0. \end{aligned} \tag{7}$$

Moreover, remark that from lemma 1, the $K \times K$ matrix $(\mathbf{A}^H\mathbf{A})^{-1}$ converges to $\mathbf{I}_K$ as $N \to \infty$. Since $(\hat{\theta}_{k,N})$ is bounded, we can extract a converging subsequence $(\hat{\theta}_{k,\varphi(N)})$. Let $\alpha_N = \hat{\theta}_{k,\varphi(N)} - \theta_k$. From lemma 1, if $\alpha_N \to \alpha \neq 0$ or $\alpha_N \to 0$ and $N|\alpha_N| \to \infty$ as $N \to \infty$, then

$$\mathbf{a}(\hat{\theta}_{k,\varphi(N)})^H \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{a}(\hat{\theta}_{k,\varphi(N)}) \xrightarrow[N \to \infty]{a.s.} 0,$$

and thus $\eta_N(\hat{\theta}_{k,\varphi(N)}) \to 1$, a contradiction with (7). Therefore we can find a further subsequence, still denoted with index $\varphi(N)$, such that $\hat{\theta}_{k,\varphi(N)} - \theta_k \to 0$ and $N|\hat{\theta}_{k,\varphi(N)} - \theta_k| \leq C < \infty$, with C a positive constant. Since the subsequence $(N|\hat{\theta}_{k,\varphi(N)} - \theta_k|)$ is bounded, we can extract a further subsequence, again denoted with index $\varphi(N)$, converging to $\beta \in [-\pi, \pi]$. From lemma 1, if $\beta \neq 0$, we get

$$\eta_{\varphi(N)}(\hat{\theta}_{k,\varphi(N)}) \xrightarrow[N \to \infty]{a.s.} 1 - \mathrm{sinc}(\beta)^2 > 0,$$

which is again in contradiction with (7). Therefore, $\beta = 0$ and all converging subsequences extracted from $(N|\hat{\theta}_{k,\varphi(N)} - \theta_k|)$ converge to 0, which of course implies that the whole sequence $(N|\hat{\theta}_{k,\varphi(N)} - \theta_k|)$ converges to 0. Applying iteratively the same argument to all the previous subsequences, we finally end up with $N(\hat{\theta}_{k,N} - \theta_k) \to 0$ w.p.1., as $N \to \infty$. □

5. REFERENCES

[1] X. Mestre and M.A. Lagunas, "Modified subspace algorithms for DoA estimation with large arrays," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 598, 2008.

[2] P. Vallet, P. Loubaton, and X. Mestre, "Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case," *submitted*, 2010, arXiv: 1002.3234.

[3] Z. Bai and J.W. Silverstein, *Spectral analysis of large dimensional random matrices*, Springer Verlag, 2010.

[4] P. Ciblat, P. Loubaton, E. Serpedin, and G.B. Giannakis, "Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1922–1934, 2002.