ROBUST G-MUSIC

Romain Couillet[1], Abla Kammoun[2].

[1]Telecommunication Department, Supélec, Gif sur Yvette, France.
[2]King Abdullah University of Science and Technology, Saudi Arabia.

ABSTRACT

An improved MUSIC algorithm for direction-of-arrival estimation is introduced that accounts both for large array sizes N comparatively with the number of independent observations n and for the impulsiveness of the background environment (e.g., presence of outliers in the observations). This method derives from the spiked G-MUSIC algorithm proposed in [1] and from the recent works by one of the authors on the random matrix analysis of robust scatter matrix estimators [2]. The method is shown to be asymptotically consistent where classical approaches are not. This superiority is corroborated by simulations.

Index Terms— Random matrix theory, MUSIC, robust estimation, elliptical distribution.

I. INTRODUCTION

The recent advances in random matrix theory have allowed for several successive improvements of statistical inference methods when the number of observations n is not large compared to the population size N. In particular in [3], a generalization of the well-known MUSIC algorithm [4] for direction-of-arrival (DoA) estimation, called G-MUSIC, was proposed which brings consistence to MUSIC in the regime where $N, n \to \infty$ with $N/n \to c \in (0, \infty)$.[1] This work assumes that the signals are random and that each signal subspace has a rank growing with N. Further extensions went in the direction of deterministic but unknown signals instead [5], and then in the direction of small signal subspace dimensions compared to N (the so-called spiked model) [1]. Finally, the generalization to non-white Gaussian noise was established in [6].

As a common denominator, all the works, which rely on subspace characterization of the sample covariance matrix of the observed samples, demand that the eigenvalues of the noise subspace have a limiting compact support. This allows one to isolate and identify the "signal eigenvalues" as those eigenvalues found away from this support. However, the eigenvalues of the noise subspace may not meet this

[1]G-MUSIC mostly consists of a properly weighted version of the MUSIC algorithm which ensures consistency in the doubly-asymptotic regime.

condition when sample outliers occur, entailing false alarms, or when the background noise is very impulsive, leading then to a possibly unbounded eigenvalue support from which no signal eigenvalue can be isolated. In these scenarios, none of MUSIC or G-MUSIC (regular or spiked) are theoretically acceptable.

More recently, the second author proposed a new analytical framework of robust estimation of large dimensional scatter matrices in the series of works [7], [2], [8]. Robust estimation of scatter exactly aims at mitigating the effect of outliers and impulsiveness in the observed datasets. In particular, it is shown in [2] that the limiting eigenvalue distribution of impulsive samples has a bounded support with no eigenvalue escaping. The present article hinges on this observation by proposing a new MUSIC improvement, called robust G-MUSIC, which works under the spiked signal space assumption and accounts both for the large dimension of the antenna array and for the impulsive nature of the noise. The key idea is to trade the sample covariance matrix for a robust scatter estimate as studied in [2]. It appears that robust G-MUSIC manages to recover both the signal powers and their DoA in impulsive scenarios where MUSIC and G-MUSIC cannot. In particular, and maybe quite surprisingly at first sight, the low rank population signal subspace (if powerful enough) engenders isolated – hence detectable – eigenvalues in the robust scatter estimate found beyond a certain detectability threshold, while low rank sample outliers remain either contained in the noise subspace or found as isolated eigenvalues below the detectability threshold. Simulations are performed that corroborate the theoretical analysis.

The remainder of the article introduces the array processing system model in Section II and our main results in Section III.

II. MODEL

We consider the model

$$y_i = \sum_{l=1}^{L} \sqrt{p_l} a_l s_{li} + \sqrt{\tau_i} w_i, \ i \in \{1, \dots, n\} \quad (1)$$

with $a_l = a(\theta_l) \in \mathbb{C}^N$ a unit norm directional vector parametrized by $\theta_l \in [0, 2\pi)$, $p_l \in \mathbb{R}^+$ a power parameter, $s_{l,j} \in \mathbb{C}$ random i.i.d. variables with zero mean and unit

variance, τ_i random i.i.d. variables with distribution ν such that $\int t\nu(dt) = 1$ and $\int t^{1+\varepsilon}\nu(dt) < \infty$ for some $\varepsilon > 0$, and $w_j \in \mathbb{C}^N$ a unitarily invariant norm-$\sqrt{N}$ vector. As for L, it is assumed finite.

Our objective is to estimate the powers $p_1, \ldots, p_L$ along with the angles $\theta_1, \ldots, \theta_L$ from the independent samples $y_1, \ldots, y_n$ in the regime $N, n \to \infty$ with $N/n \to c \in (0, 1)$.

Denoting $\tilde{w}_i = [s_{1i}, \ldots, s_{L1}, w_i^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{N+L}$ (made of independent entries of zero mean and unit variance) and $A_i = [\sqrt{p_1}a_1, \ldots, \sqrt{p_L}a_L, \sqrt{\tau_i}I_N]$, we have $y_i = A_i\tilde{w}_i$. We also denote

$$A_iA_i^* = \sum_{l=1}^{L} p_l a_l a_l^* + \tau_i I_N \triangleq B + \tau_i I_N.$$

We now introduce Maronna's robust estimator of scatter $\hat{C}_N$ as in [2]. The matrix $\hat{C}_N$ is defined, when it exists, as the unique solution to the fixed-point matrix-valued equation in Z:

$$Z = \frac{1}{n}\sum_{i=1}^{n} u\left(\frac{1}{N}y_i^* Z^{-1} y_i\right) y_i y_i^* \qquad (2)$$

where $u : \mathbb{R}^+ \to \mathbb{R}^+$ is a nonnegative non-increasing function with $u(0)$ finite and such that $\phi(x) = xu(x)$ is increasing and bounded with $\phi_\infty = \lim_{x\to\infty} \phi(x) \in (1, 1/c)$.

We recall from [2] that, if $\hat{C}_N$ exists, then it can be written

$$\hat{C}_N = \frac{1}{n}\sum_{i=1}^{n} v\left(\frac{1}{n}y_i^* \hat{C}_{(i)}^{-1} y_i\right) y_i y_i^*$$

where $v : x \mapsto u \circ g^{-1}$, $g : x \mapsto x/(1 - c\phi(x))$, and $\hat{C}_{(i)} = \hat{C}_N - \frac{1}{n}u\left(\frac{1}{n}y_i^* Z^{-1} y_i\right) y_i y_i^*$. We shall also denote $\psi(x) = xv(x)$. Recall that v is non-increasing and ψ is increasing with limit $\psi_\infty = \phi_\infty/(1 - c\phi_\infty)$.

III. MAIN RESULTS

III-A. Theoretical results

Our main results are twofold. Proposition 1 first generalizes [2, Theorem 2] on the asymptotic behavior of $\hat{C}_N$ in the large N, n regime. Then Theorem 1 and Theorem 2 provide statistical inference methods that will be used to estimate the powers $p_1, \ldots, p_L$ and angles $\theta_1, \ldots, \theta_L$.

Proposition 1 (Asymptotic Behavior): Let $\hat{C}_N$ be defined as the unique solution of (2) when it exists or arbitrarily otherwise. Then, as $N, n \to \infty$ with $N/n \to c \in (0, 1)$,

$$\|\hat{C}_N - \hat{S}_N\| \xrightarrow{\text{a.s.}} 0 \qquad (3)$$

where

$$\hat{S}_N \triangleq \frac{1}{n}\sum_{i=1}^{n} v(\tau_i\gamma) A_i\tilde{w}_i\tilde{w}_i^* A_i^*$$

with γ the unique positive solution to

$$1 = \int \frac{\psi(t\gamma)}{1 + c\psi(t\gamma)}\nu(dt). \qquad (4)$$

Sketch of Proof: We only provide an intuitive sketch of the proof. The complete proof is found in an extended version of this article. The asymptotic uniqueness of $\hat{C}_N$ is an extension of [2, Theorem 1] which follows from the same ideas and is therefore not presented here. The interesting aspect is the convergence (3) which relates to [2, Theorem 2]. Note in particular that, for $B = 0$, the result matches [2, Theorem 2]. Following the intuitive approach delineated in [2, Section 2.2], denoting $\hat{C}_{(i)} = \hat{C}_N - \frac{1}{n}u(\frac{1}{N}y_i^*\hat{C}_N^{-1}y_i)y_iy_i^*$, we expect that, for large N, n

$$\frac{1}{N}y_i^*\hat{C}_{(i)}^{-1}y_i \simeq \frac{1}{N}\operatorname{tr} A_j A_j^* \hat{C}_N^{-1}$$
$$= \frac{1}{N}\operatorname{tr} B\hat{C}_N^{-1} + \tau_j\frac{1}{N}\operatorname{tr}\hat{C}_N^{-1}.$$

Denoting $\bar{\alpha}_N = \frac{1}{N}\operatorname{tr} B\hat{C}_N^{-1}$ and $\bar{\gamma}_N = \frac{1}{N}\operatorname{tr}\hat{C}_N^{-1}$, we expect from random matrix intuition that $\bar{\alpha}_N - \alpha_N \xrightarrow{\text{a.s.}} 0$ and $\bar{\gamma}_N - \gamma_N \xrightarrow{\text{a.s.}} 0$ as $N, n \to \infty$ for some deterministic equivalents $\alpha_N, \gamma_N > 0$. If so, approximating $\frac{1}{N}y_i^*\hat{C}_{(i)}^{-1}y_i$ by $\alpha_N + \tau_j\gamma_N$ in the expression of $\hat{C}_{(i)}$ and applying classical random matrix results (for instance from [9, Theorem 1]), we must have

$$\gamma_N = \frac{1}{N}\operatorname{tr}\left(\frac{1}{n}\sum_{j=1}^{n}\frac{v(\alpha_N + \tau_j\gamma_N)A_jA_j^*}{1 + e_j}\right)^{-1}$$

where the e_j are uniquely defined as the solution of

$$e_j = v(\alpha_N + \tau_j\gamma_N)$$
$$\times \frac{1}{n}\operatorname{tr} A_jA_j^*\left(\frac{1}{n}\sum_{j=1}^{n}\frac{v(\alpha_N + \tau_j\gamma)A_jA_j^*}{1 + e_j}\right)^{-1}.$$

The latter further simplifies as

$$e_j = cv(\alpha_N + \tau_j\gamma_N)(\alpha_N + \tau_j\gamma_N) = c\psi(\alpha_N + \tau_j\gamma_N).$$

This then gives the system

$$\gamma_N = \frac{1}{N}\operatorname{tr}\left(\frac{1}{n}\sum_{j=1}^{n}\frac{v(\alpha_N + \tau_j\gamma_N)}{1 + c\psi(\alpha_N + \tau_j\gamma_N)}(B + \tau_j I_N)\right)^{-1}$$

$$\alpha_N = \frac{1}{N}\operatorname{tr} B\left(\frac{1}{n}\sum_{j=1}^{n}\frac{v(\alpha_N + \tau_j\gamma_N)}{1 + c\psi(\alpha_N + \tau_j\gamma_N)}(B + \tau_j I_N)\right)^{-1}.$$

It now follows from the small-rank property and norm boundedness of the vectors composing B that $\alpha_N \to 0$. Also, the small rank of B does not alter asymptotically the trace term in γ_N so that B can be discarded in this expression in the limit. Hence $\gamma_N \to \gamma$ which, outside a zero measure set, satisfies (4) from the strong law of large numbers. ∎

Proposition 1 states that, for many problems dealing with $\hat{C}_N$, which is a complicated matrix defined as the solution of a fixed-point equation, one can instead use $\hat{S}_N$, a much simpler object of the bi-correlated sample covariance matrix

$$w_j = \int \frac{v(t\gamma)\nu(dt)}{\left(1 + \delta(\hat{\lambda}_j)tv(t\gamma)\right)^2} \left[\int \frac{v(t\gamma)\nu(dt)}{1 + \delta(\hat{\lambda}_j)tv(t\gamma)} \left(1 - \frac{1}{c}\int \frac{\delta(\hat{\lambda}_j)^2 t^2 v(t\gamma)^2 \nu(dt)}{\left(1 + \delta(\hat{\lambda}_j)tv(t\gamma)\right)^2}\right)\right]^{-1} \tag{5}$$

$$\hat{w}_j = \frac{1}{n}\sum_{i=1}^{n} \frac{v(\hat{\tau}_i)}{\left(1 + \hat{\delta}(\hat{\lambda}_j)\hat{\tau}_i v(\hat{\tau}_i\hat{\gamma}_N)\right)^2} \left[\frac{1}{n}\sum_{i=1}^{n} \frac{v(\hat{\tau}_i)}{1 + \hat{\delta}(\hat{\lambda}_j)\hat{\tau}_i v(\hat{\tau}_i\hat{\gamma}_N)} \left(1 - \frac{1}{N}\sum_{i=1}^{n} \frac{\hat{\delta}(\hat{\lambda}_j)^2 \hat{\tau}_i^2 v(\hat{\tau}_i)^2}{\left(1 + \hat{\delta}(\hat{\lambda}_j)\hat{\tau}_i v(\hat{\tau}_i\hat{\gamma}_N)\right)^2}\right)\right]^{-1}. \tag{6}$$

type, extensively analyzed in [10]. Indeed Proposition 1 ensures as a fundamental corollary that, denoting $\lambda_1(X) \leq \ldots \leq \lambda_N(X)$ the ordered eigenvalues of a Hermitian matrix X,

$$\max_{1 \leq i \leq N} |\lambda_i(\hat{C}_N) - \lambda_i(\hat{S}_N)| \xrightarrow{\text{a.s.}} 0$$

so that the largest eigenvalues of $\hat{C}_N$ coincide asymptotically with those of $\hat{S}_N$ and, as far as (first order) G-estimation is concerned, we can equivalently work with $\hat{C}_N$ or $\hat{S}_N$.

A thorough analysis of the expression of $\hat{S}_N$ now reveals that it has a limiting eigenvalue distribution with compact support and may exhibit isolated eigenvalues outside its limiting support. These isolated eigenvalues may be of two types. Some may first arise as sample outliers from large values of τ_i which, as established by [2, Theorem 2], cannot exceed a limiting value. In particular, since $\hat{S}_N \preceq \frac{\phi_\infty}{\gamma(1-c\phi_\infty)}\frac{1}{n}\sum_{i=1}^{n}w_iw_i^*$ when $B = 0$, from standard random matrix considerations, we find that, for arbitrary $\varepsilon > 0$, these eigenvalues cannot exceed $S^+ + \varepsilon$ infinitely often, where we defined

$$S^+ \triangleq \frac{\phi_\infty}{\gamma(1 - c\phi_\infty)}(1 + \sqrt{c})^2.$$

On the opposite, by the presence of the small rank $B \neq 0$ in each term $A_iA_i^*$, a deeper analysis (similar to [11]) reveals that some other eigenvalues may be found *beyond* $S^+ + \varepsilon$ infinitely often. The value S^+, although not optimal, will from now on constitute our signal detectability threshold for our improved signal classification algorithm based on $\hat{C}_N$.

We are now in position to introduce our first main result.

Theorem 1 (Estimation with known ν): Denote $q_1 \geq \ldots \geq q_L > 0$ the positive eigenvalues of B (they may depend of N), $u_1, \ldots, u_L$ their respectively associated eigenvectors in B, and $\hat{u}_1, \ldots, \hat{u}_N$ the eigenvectors of $\hat{C}_N$ respectively associated with the eigenvalues $\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_N$. For $x > 0$ sufficiently large, call $\delta(x)$ the unique negative solution of

$$\delta(x) = c\left(-x + \int \frac{tv(t\gamma)}{1 + \delta(x)tv(t\gamma)}\nu(dt)\right)^{-1}.$$

Then, recalling the definition of γ as the solution of (4), we have the following two results:

1. Eigenvalue estimation. For each k with $\liminf_N q_k$ sufficiently large,

$$-c\left(\delta(\hat{\lambda}_k)\int \frac{v(t\gamma)}{1 + \delta(\hat{\lambda}_k)tv(t\gamma)}\nu(dt)\right)^{-1} - q_k \xrightarrow{\text{a.s.}} 0.$$

2. Localization function estimation. Assume that $\liminf_N q_L$ is large enough. Denote $\eta_N, \hat{\eta}_N : [0, 2\pi) \to \mathbb{R}$ the functions given by

$$\eta_N(\theta) = 1 - \sum_{k=1}^{L} a(\theta)^* u_k u_k^* a(\theta)$$

$$\hat{\eta}_N(\theta) = 1 - \sum_{k=1}^{L} w_k a(\theta)^* \hat{u}_k \hat{u}_k^* a(\theta)$$

where $w_j > 0$ is given in (5). Then, for each $\theta \in [0, 2\pi)$,

$$\hat{\eta}_N(\theta) - \eta_N(\theta) \xrightarrow{\text{a.s.}} 0$$

as $N, n \to \infty$ and $N/n \to c$.

Proof: The proof starts by transferring the difficult analysis of the bahavior of $\hat{C}_N$ into the simpler analysis of $\hat{S}_N$, thanks to Theorem 1. This being done, as $\hat{S}_N$ is a random matrix of a rather standard form, the proof then follows similar steps as in the proof of the first order results in [12] (see also [11]), although for a slightly different spiked model. For lack of space, we do not further detail these here. The complete derivation can be found in the extended version of this article [13]. ∎

Theorem 1 states that, based on the largest eigenvalues of $\hat{C}_N$, one can consistently estimate the values of $q_1, \ldots, q_L$ and the localization functions $\eta_N(\theta)$ for each θ.

Note that the theorem mentions that q_k should be large enough. This sufficiently large value ensures that the number of eigenvalues found above the detectability threshold S^+ must equal at least k (see [13] for more precise statements). Without this assumption, the above estimates may be invalid as some of the largest eigenvalues may correspond to sample noise outliers or eigenvalues from the noise bulk. This condition is classical and relates to the often named *separability condition* by which population eigenvalues q_k engender sample eigenvalues found away from a detectability threshold in the sprectrum of $\hat{C}_N$ (or equivalently of $\hat{S}_N$).

Note importantly that, if the sample covariance matrix were used in place of $\hat{C}_N$ (therefore leading to a G-MUSIC

algorithm), the impulsiveness of the noise may not allow for the eigen-spectrum of noise subspace to be of compact support, making it impossible to separate signal eigenvalues.

Theorem 1 assumes ν to be known. Using instead the fact that $\gamma - \hat{\gamma}_N \xrightarrow{\text{a.s.}} 0$ where

$$\hat{\gamma}_N = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{N} y_i^* \hat{C}_{(i)}^{-1} y_i$$

and that τ_i can be well approximated by

$$\hat{\tau}_i \triangleq \frac{1}{\hat{\gamma}_N} \frac{1}{N} y_i^* \hat{C}_{(i)}^{-1} y_i$$

(where we recall that $\hat{C}_{(i)} = \hat{C}_N - \frac{1}{n} u(\frac{1}{N} y_i^* \hat{C}_N^{-1} y_i) y_i y_i^*$), we deduce the following alternative empirical estimator.

Theorem 2 (Empirical estimation): Define $\hat{\delta}(x)$ as the unique solution to

$$\hat{\delta}(x) = \frac{N}{n} \left(-x + \frac{1}{n} \sum_{i=1}^{n} \frac{\hat{\tau}_i v(\hat{\tau}_i \hat{\gamma}_N)}{1 + \hat{\delta}(x) \hat{\tau}_i v(\hat{\tau}_i \hat{\gamma}_N)} \right)^{-1}.$$

Then we have the following results

1. Eigenvalue estimation. For each k with $\liminf_N q_k$ large enough,

$$-\left(\hat{\delta}(\hat{\lambda}_k) \frac{1}{N} \sum_{i=1}^{n} \frac{v(\hat{\tau}_i \hat{\gamma}_N)}{1 + \hat{\delta}(\hat{\lambda}_k) \hat{\tau}_i v(\hat{\tau}_i \hat{\gamma}_N)} \right)^{-1} - q_k \xrightarrow{\text{a.s.}} 0.$$

2. Localization function estimation. Assume that $\liminf_N q_1$ is large enough. Denote $\hat{\eta}_N^{\text{emp}} : [0, 2\pi) \to \mathbb{R}$ the function given by

$$\hat{\eta}_N^{\text{emp}}(\theta) = 1 - \sum_{k=1}^{L} \hat{w}_k a(\theta)^* \hat{u}_k \hat{u}_k^* a(\theta)$$

where $\hat{w}_j$ is defined in (6). Then, for each $\theta \in [0, 2\pi)$,

$$\hat{\eta}_N^{\text{emp}}(\theta) - \eta_N(\theta) \xrightarrow{\text{a.s.}} 0.$$

Note that Theorem 2 provides another set of consistent estimators for the same quantities q_k and $\eta_N(\theta)$ as in Theorem 1 but that do not rely on the knowledge of ν, as the τ_i are individually estimated. There is however no saying whether knowing ν or not improves the quality of the estimates.

III-B. Application to array processing

In this section, we use Theorem 1 and Theorem 2 to perform source detection, power estimation, and source localization in array processing.

We consider the model (1) in which s_{lj}, $l = 1, \ldots, L$ are $L \geq 0$ signal sources impinging at time instant j an N-antenna array with angles $\theta_1, \ldots, \theta_L$. The signal y_i received at time i at the array is corrupted by the background noise $\sqrt{\tau_i} w_i$ with w_i unitary and τ_i having unit-mean distribution ν. In particular, if $\nu = \nu' / \int t d\nu'(t) \triangleq \nu_G$ with $\nu' \sim \chi_{2N}^2$

(the chi-square distribution with $2N$ degrees of freedom), the noise is standard Gaussian. We however consider here scenarios where the noise may have a more impulsive behavior. In particular, one may take $\nu = \nu'' / \int t d\nu''(t)$ with $\nu'' \sim (1 - \varepsilon) \nu_G + \varepsilon \delta_A$ for some ε small and a large A value, introducing then outliers of high amplitude, which simulates strong but rare noise impulses. One may alternatively take ν continuous with heavier than Gaussian tails, leading in particular to scenarios where the limiting spectrum of the sample covariance matrix has unbounded support, translating here the large impulsiveness of the background noise.

Note now that $\eta_N(\theta)$ (in Theorem 1) can be written $\eta_N(\theta) = a(\theta)^* \Pi_W a(\theta)$ with $\Pi_W = I_N - \sum_{k=1}^{L} u_k u_k^*$ a projector on the population noise subspace. This is then the localization function found in the standard MUSIC algorithm [4]. Item 2) of Theorems 1 and 2 thus provides asymptotically consistent estimates for this localization function, under knowledge of ν or not. From Theorems 1 and 2, we may then define two novel DoA estimates $\hat{\theta}_1, \ldots, \hat{\theta}_L$ and $\hat{\theta}_1^{\text{emp}}, \ldots, \hat{\theta}_L^{\text{emp}}$ for the angles $\theta_1, \ldots, \theta_L$, which we hereafter call the robust G-MUSIC and empirical robust G-MUSIC estimates, respectively, and which we define as the L deepest minima of the functions $\theta \mapsto \hat{\eta}_N(\theta)$ and $\theta \mapsto \hat{\eta}_N(\theta)^{\text{emp}}$, respectively. For comparison, we recall that the standard MUSIC algorithm assumes as an estimate for $\eta_N(\theta)$ the function $\theta \mapsto a(\theta)^* \Pi_W^{(\text{SCM})} a(\theta)$ with $\Pi_W^{(\text{SCM})}$ the noise subspace of the sample covariance matrix for the vectors $y_1, \ldots, y_n$. Robust (non-G) MUSIC is similar but with $\hat{C}_N$ in place of the sample covariance matrix. The (non-robust) G-MUSIC algorithm is similar to the robust G-MUSIC method but with the function v set equal to one, which is somewhat similar to the G-MUSIC approach developed in [6].[2] In addition to G-MUSIC, we introduce the empirical version of G-MUSIC which, similar to the empirical robust G-MUSIC approach, uses estimates of τ_i instead of using the information of ν.

As is standard in array processing, we may assume that $a(\theta)$ is such that, for each $i \neq j$, $a(\theta_i)^* a(\theta_j) \to 0$. This is in particular the case if $[a(\theta)]_k = N^{-\frac{1}{2}} \exp(2\pi \iota (k - 1)\theta)$, corresponding to a uniform linear array. In this scenario, it is easily seen that, as $N \to \infty$, $q_i \to p_i$ for each $i \in \{1, \ldots, L\}$. The item 1) of both Theorems 1 and 2 can then be used here to estimate $p_1, \ldots, p_L$ under knowledge or not of ν. Similar to the localization function estimate, for some ν, classical sample covariance matrix-based techniques to estimate the p_i such as [3], [14] cannot be adapted to account for the τ_i.

Figure 1 compares the performance of robust G-MUSIC and empirical robust G-MUSIC against MUSIC and G-MUSIC for $L = 2$, $\theta_1 = 10°$, $\theta_2 = 12°$, and $\nu \sim$

---

[2]The model in [6] is essentially the same as here with v set to one, up to a structural difference in the spiked model under study. It is interesting to note in particular that, setting v to one, the results of Theorem 1 are very similar in nature to the results in [6].
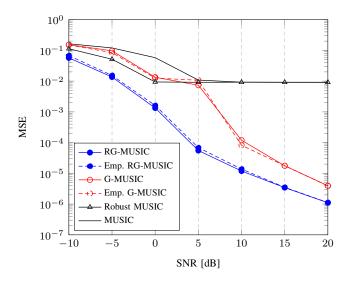
Fig. 1. Mean square error (MSE) performance of the smallest angle estimation of the MUSIC estimators for $N = 20$, $n = 100$, two sources at $10°$ and $12°$, Gamma-$(0.8, 1.25)$ impulsions. Robust G-MUSIC (RG-MUSIC) and empirical robust G-MUSIC (Emp. RG-MUSIC) are compared against MUSIC, G-MUSIC, and empirical G-MUSIC (Emp. G-MUSIC).

Gamma$(0.8, 1.25)$. This figure shows that, due to N/n not being too small, MUSIC and robust MUSIC cannot separate the two close angles, which is a well-documented behavior. The G-MUSIC approaches accommodate for this, but suffer from the impulsive noise nature, and even more so if the shape parameter of the Gamma distribution gets reduced, as can be verified by further simulations that can be found, along with additional discussions, in [13].

IV. REFERENCES

[1] P. Vallet, W. Hachem, P. Loubaton, X. Mestre, and J. Najim, "An improved music algorithm based on low-rank perturbation of large random matrices," in *Proceedings of IEEE Workshop on Statistical Signal Processing (SSP'11)*, Nice, France, 2011.

[2] R. Couillet, F. Pascal, and J. W. Silverstein, "The random matrix regime of Maronna's M-estimator with elliptically distributed samples," *Journal of Multivariate Analysis*, 2013. [Online]. Available: http://arxiv.org/abs/1311.7034

[3] X. Mestre, "Improved estimation of eigenvalues of covariance matrices and their associated subspaces using their sample estimates," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5113–5129, Nov. 2008.

[4] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

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

[6] J. Vinogradova, R. Couillet, and W. Hachem, "Statistical inference in large antenna arrays under unknown noise pattern," *arXiv preprint 1301.0306*, 2013.

[7] R. Couillet, F. Pascal, and J. W. Silverstein, "Robust Estimates of Covariance Matrices in the Large Dimensional Regime," *IEEE Transactions on Information Theory*, 2013. [Online]. Available: http://arxiv.org/abs/1204.5320

[8] R. Couillet and M. McKay, "Large dimensional analysis and optimization of robust shrinkage covariance matrix estimators," *to appear in Journal of Multivariate Analysis*, 2013.

[9] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in MISO broadcast channels with limited feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4509–4537, 2012. [Online]. Available: http://arxiv.org/abs/0906.3682

[10] R. Couillet and W. Hachem, "Analysis of the limit spectral measure of large random matrices of the separable covariance type," *Journal of Multivariate Analysis*, 2013.

[11] F. Chapon, R. Couillet, W. Hachem, and X. Mestre, "On the isolated eigenvalues of large Gram random matrices with a fixed rank deformation," *Electronic Journal of Probability*, 2012, submitted for publication.

[12] R. Couillet and W. Hachem, "Fluctuations of spiked random matrix models and failure diagnosis in sensor networks," *IEEE Transactions on Information Theory*, 2012, to appear.

[13] R. Couillet, "Robust spiked random matrices and a robust g-music estimator," *submitted to Journal of Multivariate Analysis*, 2014. [Online]. Available: http://arxiv.org/pdf/1404.7685

[14] R. Couillet, J. W. Silverstein, Z. D. Bai, and M. Debbah, "Eigen-inference for energy estimation of multiple sources," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2420–2439, 2011.