

Can the Threshold Performance of Maximum Likelihood DOA Estimation be Improved by Tools from Random Matrix Theory?

Yuri I. Abramovich¹, Ben A. Johnson^{2,*}

¹Institute for Telecommunications Research
University of South Australia, Mawson Lakes, SA 5045 Australia

²Colorado School of Mines, Golden, CO 80401

*Corresponding Author: ben.a.johnson@ieee.org

Copyright ©2014 Horizon Research Publishing All rights reserved.

Abstract For direction of arrival (DOA) estimation in the threshold region, it has been shown that use of Random Matrix Theory (RMT) eigensubspace estimates provides significant improvement in MUSIC performance. Here we investigate whether these RMT methods can also improve the threshold performance of unconditional (stochastic) maximum likelihood DOA estimation (MLE).

Keywords Signal Detection and Estimation, Maximum-likelihood Estimation, Random Matrix Theory

1 Introduction

The “pre-asymptotic” performance of traditional stochastic (unconditional) MLE of DOAs for m closely-spaced Gaussian signals immersed in white Gaussian noise using an M -sensor array and a finite number T independent identically distributed (i.i.d.) training samples continues to be the subject of investigations [3–5, 15] despite being a relatively “old” problem [16]. The main reason, as formulated in [4], is that in the pre-asymptotic domain, “*no general non-asymptotic results are [currently] available for the performance evaluation of the ML method, and each problem requires a special investigation*”.

As an alternative, instead of focusing on the accurate non-asymptotic analysis, one may consider the asymptotic model where both quantities M and T grow without bound, while their quotient converges to a fixed finite quantity:

$$M, T \rightarrow \infty, \quad \frac{M}{T} \rightarrow c, \quad 0 < c < \infty \quad (1)$$

This condition is known as the Kolmogorov asymptotic condition [6], and underpins a field of analysis referred to as Generalized Statistical Analysis (GSA) or Random Matrix Theory (RMT). Of course, in any practical situation, one deals with finite M and T values, and therefore generalized “G-asymptotic” (i.e. asymptotic per the conditions in (1)) results may still not be sufficiently accurate. However, in numerous studies, it has been demonstrated that estimators that are consistent G-asymptotically are more robust in the presence of finite samples T than other estimators which are only consistent for $T \rightarrow \infty$, $M = \text{constant}$ [6, 10, 11, 18].

Specifically, let us suppose that i.i.d. observations $\mathbf{x}_1, \dots, \mathbf{x}_T$ of random vector ξ with dimension M are given, and we wish to estimate some value $\varphi(R_M)$, where φ is a continuous function of the entries of the (true) covariance matrix R_M of vector ξ . If T is large and M is fixed and does not change as T grows (i.e. the standard asymptotic case), then as an estimator of $\varphi(R_M)$, we may take $\varphi(\hat{R}_M)$, where

$$\text{plim}_{T \rightarrow \infty} \varphi(\hat{R}_M) = \varphi(R_M) \quad (2)$$

and \hat{R}_M is the standard sample covariance matrix for zero-mean data ($= \frac{1}{T} \sum_{j=1}^T \mathbf{x}\mathbf{x}^H$). Moreover, if \mathbf{x}_j , $j = 1, \dots, T$ is a set of i.i.d. complex (circular) Gaussian samples (i.e. $\mathbf{x}_j \sim \mathcal{CN}(0, R_M)$), then \hat{R}_M and $\varphi(\hat{R}_M)$ are not only consistent (in the standard asymptotic framework), but also the maximum likelihood estimate of R_M and $\varphi(R_M)$ respectively [17]. However, this familiar assertion is not in general true under the G-asymptotic condition

(1). There, for a wide range of functions $\varphi(R_M)$, one can find a measurable function ψ of the entries of the matrix R_M for which

$$\text{plim}_{M,T \rightarrow \infty} [\varphi(\hat{R}_M) - \psi(R_M)] = 0. \tag{3}$$

But in general, the functions $\varphi(R_M)$ and $\psi(R_M)$ do not coincide, in which case $\varphi(\hat{R}_M)$ is not necessarily a G-consistent estimate of $\varphi(R_M)$. However, with the help of function $\psi(R_M)$, one may try to find a measurable function $g(\hat{R}_M)$ such that

$$\text{plim}_{M,T \rightarrow \infty} [g(\hat{R}_M) - \varphi(R_M)] = 0. \tag{4}$$

and where the distribution of normalized difference $g(\hat{R}_M) - \varphi(R_M)$ is asymptotically normal. The function $g(\hat{R}_M)$ is then called a G-estimator [6], and is consistent under condition (1).

In [10,11], this RMT methodology has been used to find G-estimates of eigenvalues for a covariance matrix with known multiplicity of its eigenvalues. Specifically, for an M -variate covariance matrix R_M , let $\gamma_1 < \gamma_2 < \dots < \gamma_{\bar{M}}$ be the set of \bar{M} distinct eigenvalues ($1 \leq \bar{M} \leq M$) after accounting for individual multiplicity K_m of the \bar{M} true eigenvalues (i.e. $\sum_{m=1}^{\bar{M}} K_m = M$). Associated with each eigenvalue γ_m , there is a complex subspace of dimension K_m . This subspace is determined by an $M \times K_m$ matrix of corresponding eigenvectors, denoted by E_m , such that $E_m^H E_m = I_{K_m}$. Note that this specification is unique up to right multiplication by the orthogonal matrix, and therefore the problem of eigendecomposition for the matrix R_M

$$R_M = \sum_{j=1}^{\bar{M}} \gamma_j E_j E_j^H \tag{5}$$

is more convenient to formulate as a problem of estimation of orthogonal projection matrices, defined as

$$P_m = E_m E_m^H, \quad m = 1, \dots, \bar{M}, \tag{6}$$

given the sample covariance matrix \hat{R} that has the eigendecomposition

$$\hat{R} = \sum_{j=1}^M \hat{\lambda}_j \hat{U}_j \hat{U}_j^H \tag{7}$$

To use this eigendecomposition to estimate the orthogonal projection matrices P_m , let \mathcal{K}_m be a set of indices

$$\mathcal{K}_m = \left\{ \sum_{j=1}^{m-1} K_j, \sum_{j=1}^{m-1} K_j + 1, \dots, \sum_{j=1}^m K_j \right\} \tag{8}$$

with the cardinality of \mathcal{K}_m equal to the multiplicity of the eigenvalue γ_m , namely K_m .

The classical (and indeed maximum likelihood for multivariate Gaussian observations) estimator of the m -th eigenvalue and orthogonal projection matrices P_m in (6) is given by

$$\hat{\gamma}_m^{ML} = \frac{1}{K_m} \sum_{k \in \mathcal{K}_m} \hat{\lambda}_k; \quad \hat{P}_m^{ML} = \sum_{k \in \mathcal{K}_m} \hat{U}_k \hat{U}_k^H \tag{9}$$

One can see that $\hat{\gamma}_m^{ML}$ and \hat{P}_m^{ML} are specified as $\varphi(\hat{R}_m)$ in (2) and indeed are strongly T -consistent (and ML) estimates of γ_m and P_m .

Yet in Theorem 2 of [10], X. Mestre proved that under certain conditions (As1-As4), the traditional estimators of eigenvalues and eigensubspaces ($\hat{\gamma}_m^{ML}$ and $\hat{\eta}_j^{ML} = S_1^H \sum_{k \in \mathcal{K}_m} \hat{U}_k \hat{U}_k^H S_2$, respectively, where S_1 and S_2 are two $M \times 1$ deterministic vectors) can be shown to converge under condition (1) (i.e. $M, T \rightarrow \infty$ with $M/T = c$ constant and finite) to:

$$|\hat{\gamma}_m^{ML} - \bar{\gamma}_m^{ML}| \rightarrow 0 \text{ and } |\hat{\eta}_j^{ML} - \bar{\eta}_m^{ML}| \rightarrow 0 \tag{10}$$

where the (non-random) values $\bar{\gamma}_m^{ML}$ and $\bar{\eta}_m^{ML}$ are defined as

$$\bar{\gamma}_m^{ML} = \gamma_m \left(1 - \frac{c}{M} \sum_{\substack{r=1 \\ r \neq m}}^{\bar{M}} K_r \frac{\gamma_r}{\gamma_r - \gamma_m} \right) \tag{11}$$

and

$$\bar{\eta}_m^{ML} = \sum_{k=1}^{\bar{M}} w_m(k) S_1^H E_k E_k^H S_2 \tag{12}$$

$$w_m(k) = \begin{cases} 1 - \frac{1}{K_m} \sum_{\substack{r=1 \\ r \neq m}}^{\bar{M}} K_r \left(\frac{\gamma_m}{\gamma_r - \gamma_m} - \frac{\mu_m}{\gamma_r - \mu_m} \right) & k = m \\ \frac{\gamma_m}{\gamma_r - \gamma_m} - \frac{\mu_m}{\gamma_r - \mu_m} & k \neq m \end{cases} \tag{13}$$

where μ_m is the m -th solution to the following equation in μ :

$$\frac{1}{M} \sum_{r=1}^{\overline{M}} K_r \frac{\gamma_r}{\gamma_r - \mu} = \frac{1}{c} \quad (14)$$

under the convention $\mu_1 < \mu_2 < \dots < \mu_{\overline{M}}$. Thus, the traditional ML estimators for $\hat{\gamma}_m^{ML}$ and $\hat{\eta}_m^{ML}$ are not G-consistent estimates, and therefore could potentially be improved within the RMT methodology¹. “Improved” G-consistent estimates for γ_m and P_m under the conditions (As1-As4) have been derived by X. Mestre, Theorem 3 in [10]:

$$\begin{aligned} \hat{\gamma}_m^G &= \frac{T}{K_m} \sum_{k \in \mathcal{K}_m} (\hat{\lambda}_k - \hat{\mu}_k), \\ \hat{\eta}_m^G &= \sum_{k=1}^M \rho_m(k) S_1^H \hat{U}_k \hat{U}_k^H S_2; \end{aligned} \quad (15)$$

where

$$\begin{aligned} \rho_j(k) &= \begin{cases} -\varphi_m(k), & k \notin \mathcal{K}_m \\ 1 + \psi_m(k), & k \in \mathcal{K}_m \end{cases} \\ \varphi_m(k) &= \sum_{r \in \mathcal{K}_m} \left(\frac{\hat{\lambda}_r}{\hat{\lambda}_k - \hat{\lambda}_r} - \frac{\hat{\mu}_r}{\hat{\lambda}_k - \hat{\mu}_r} \right) \\ \psi_m(k) &= \sum_{r \notin \mathcal{K}_m} \left(\frac{\hat{\lambda}_r}{\hat{\lambda}_k - \hat{\lambda}_r} - \frac{\hat{\mu}_r}{\hat{\lambda}_k - \hat{\mu}_r} \right) \end{aligned} \quad (16)$$

and $\hat{\mu}_1 \leq \hat{\mu}_2 \leq \dots \leq \hat{\mu}_M$ are the real-valued solutions to the following equation in $\hat{\mu}$:

$$\frac{1}{M} \sum_{k=1}^M \frac{\hat{\lambda}_k}{\hat{\lambda}_k - \hat{\mu}} = \frac{1}{c}; \quad c = \frac{M}{T} \quad (17)$$

One can prove (see [10]) that for any fixed M , as $T \rightarrow \infty$, we have

$$\hat{\mu}_k \rightarrow \hat{\lambda}_k, \quad T(\hat{\lambda}_k - \hat{\mu}_k) \rightarrow \hat{\lambda}_k \quad (18)$$

and therefore

$$\hat{\gamma}_m^G \rightarrow \frac{1}{K_m} \sum_{k \in \mathcal{K}_m} \hat{\lambda}_k = \hat{\gamma}_m^{ML} \quad (19)$$

On the other hand, $\varphi_m(k) \rightarrow 0$ and $\psi_m(k) \rightarrow 0$, which implies that $\hat{\eta}_m^G \rightarrow \hat{\eta}_m^{ML}$, as $T \rightarrow \infty$.

Therefore, the RMT methodology provides a *unique set* of G-consistent estimates for $\hat{\gamma}_m^G$ and $\hat{\eta}_m^G$ that tend to the classical MLE estimates under traditional asymptotic assumptions ($M = \text{constant}$, $T \rightarrow \infty$). Specifically, the G-consistent MUSIC (or G-MUSIC) pseudo-spectrum estimate

$$\hat{\eta}_1^G(\theta) = \sum_{k=1}^M \rho_1(k) S^H(\theta) \hat{U}_k \hat{U}_k^H S(\theta) \quad (20)$$

has been suggested in [9] and demonstrated significant performance improvement compared with the traditional MUSIC function $\hat{\eta}_1^{ML}(\theta)$.

Recently in [8], we showed that for closely spaced sources, despite the substantial improvements demonstrated by G-MUSIC (20) with respect to conventional MUSIC, the “threshold” performance of G-MUSIC (where the estimator’s mean squared error departs rapidly from the Cramér-Rao lower bound as SNR and/or sample support is reduced) remains significantly worse than performance of the rigorously implemented (via global search) ML DOA estimation. Since the RMT-methodology, and specifically the G-consistent estimates in (15), are able to improve conventional MUSIC performance and differ from the traditional ML eigenvalue and subspace estimates, it is quite legitimate to investigate whether these estimates may be used to improve the threshold performance of ML DOA estimation itself. In this correspondence, we try to address this question.

¹Note that for $S_1 = S_2 = S(\theta)$ where $S(\theta)$ is an M -element antenna steering (manifold) vector specified by the DOA θ , and for $m = 1$, $\hat{\eta}_1(\theta)$ is the traditional MUSIC pseudo-spectrum function.

2 G-asymptotic ML DOA Estimation (G-MLE) Formulation

For the stochastic (unconditional) ML model under consideration, the set of T i.i.d. data $x_j \in \mathcal{CN}(0, R_0)$, $j = 1, \dots, T$ is described by the true/exact covariance matrix R_0 , comprised of m independent planewave sources with direction of arrival θ_m , embedded in white noise. Under this model, the traditional ML DOA estimates are found as the global maximum of the stochastic (unconditional) likelihood function (LF) computed for the model covariance matrix $R(\Omega_m)$, uniquely specified by $2m + 1$ parameters in Ω_m (m source DOAs and powers as well as the noise power) and denoted $\mathcal{L}[R(\Omega_m)]$:

$$\hat{\Omega}_m = \arg \max_{\Omega_m} \mathcal{L}[R(\Omega_m)],$$

$$\mathcal{L}[R(\Omega_m)] = \left[\frac{\exp[-\text{tr} R^{-1}(\Omega_m) \hat{R}]}{\pi^M \det R(\Omega_m)} \right]^T \quad (21)$$

or, equivalently, as the minimum of the (scaled)² log-likelihood function (LLF)

$$l\mathcal{L}[R(\Omega_m)] = \frac{\text{tr} R^{-1}(\Omega_m) \hat{R}}{M} + \frac{\log \det R(\Omega_m)}{M} \quad (22)$$

Note that for any given $R(\Omega_m)$, the LLF (22) may be viewed as the ML estimate of the function

$$\begin{aligned} \varphi[R(\Omega_m), R_0] &\equiv \varphi[R^{-\frac{1}{2}}(\Omega_m) R_0 R^{-\frac{1}{2}}(\Omega_m)] = \\ &= \frac{1}{M} \text{tr} [R^{-\frac{1}{2}}(\Omega_m) R_0 R^{-\frac{1}{2}}(\Omega_m)] + \frac{1}{M} \log \det R(\Omega_m) \end{aligned} \quad (23)$$

with an unknown R_0 , here represented by its generic (positive definite Hermitian, for $T \geq M$) ML estimate sample covariance matrix \hat{R} formed from the training data \mathbf{x}_j .

Since the likelihood function (22) is just $\varphi(\hat{R}_M(\Omega_m))$ (per the notation in (2)), under the RMT methodology, we need to construct the G-consistent estimate $g(\hat{R}_M(\Omega_m))$ (per the notation in (4)) of the function φ_0 (the portion of φ from (23) dependent on R_0 , which can be simplified to $\frac{1}{M} \text{tr} R^{-1}(\Omega_m) R_0$).

Yet, according to Lemma 3.1 in [6] (see eqn. 3.10, p. 182), for a very broad class of empirical covariance matrices that embraces the complex Wishart case

$$\hat{R}_M(\Omega_m) = \frac{1}{T} R^{-\frac{1}{2}}(\Omega_m) R_0^{\frac{1}{2}} \hat{C} R_0^{\frac{1}{2}} R^{-\frac{1}{2}}(\Omega_m); \hat{C} \sim \mathcal{CW}(M, T, I_M), \quad (24)$$

under the Kolmogorov asymptotic condition (1), the following property is proven:

$$\text{plim}_{T \rightarrow \infty, \frac{M}{T} \rightarrow c} \left[\frac{1}{M} \text{tr} \hat{R}_M - \frac{1}{M} \mathcal{E}(\text{tr} \hat{R}_M) \right] = 0, \quad (25)$$

where $\mathcal{E}(\cdot)$ is the expectation operator and $\mathcal{CW}(M, T, I_M)$ is the complex Wishart distribution [7].

This property means that for the function φ_0 , we have a *special case* where its G-consistent estimate $g(\hat{R}_M(\Omega_m))$ coincides with its ML estimate $\varphi_0(\hat{R})$:

$$g(\hat{R}_M(\Omega_m)) = \varphi(\hat{R}_M(\Omega_m)) = \frac{1}{M} \text{tr} [R^{-1}(\Omega_m) \hat{R}] \quad (26)$$

In essence, the property (25) means that for the multivariate complex Gaussian data case, the (scaled) unconditional log-likelihood function (22) may be treated as the ML estimate and at the same time, the G-consistent estimate of the function $\varphi[R(\Omega_m), R_0]$ given in (23).

Lemma 3.1 from [6], and more importantly, its interpretation as G-consistency of the standard log-likelihood function³, leaves an apparent contradiction which needs to be resolved. Specifically, since according to Theorem 2 from [10], individual ML estimates of eigenvalues are not G-consistent and can be improved, per (15). At the same time, the *sum* of all eigenvalues (i.e. the sample covariance matrix trace) is G-consistent, and therefore cannot be improved within the RMT methodology⁴. This contradiction could be resolved by the direct proof that the sum

²This scaling by the factor $1/M$ is immaterial for a fixed M , but is subsequently required for G-asymptotic analysis with $(M, T) \rightarrow \infty$, $M/T \rightarrow c$, for $l\mathcal{L}[R(\Omega_m)]$ to remain finite as $M \rightarrow \infty$

³It is important to emphasize that G-consistency of the likelihood function (or the G-MUSIC function, for that matter) does not imply that the resultant DOA estimates are always G-consistent as well. On the contrary, the fact that the globally optimal ML DOA estimates (as well as the G-MUSIC DOA estimates) exhibit performance breakdown in their respective threshold regions demonstrates that G-consistency of the likelihood or G-MUSIC function is only a necessary, but not at all sufficient condition for G-consistent DOA estimation.

⁴While Theorem 3 in [10], provides a unique set of G-consistent estimates (15) within the RMT methodology, it does not resolve questions regarding the broader uniqueness of G-consistent estimates. In this paper, per our title, we confine ourselves to the application of these (strongly) consistent RMT estimates for individual eigenvalues and eigenvectors. While the possibility of another methodology which provides "better" G-consistent estimates cannot be precluded, to our knowledge, such a methodology has not yet been developed.

of G-consistent individual eigenvalue estimates $\hat{\gamma}_m^G$ in (15) is strictly equal to the sum of sample eigenvalues $\hat{\lambda}_k$ of $\hat{R}_M(\Omega_m)$, ie.

$$\sum_{m=1}^{\bar{M}} K_m \hat{\gamma}_m^G = \sum_{k=1}^M \hat{\lambda}_k = \text{tr}[\hat{R}_M(\Omega_m)] \quad (27)$$

According to (15), we therefore have to prove that

Theorem 1 *Under conditions As1-As3 in [10], the estimate*

$$\sum_{m=1}^{\bar{M}} K_m \hat{\gamma}_m^G = T \left(\sum_{k=1}^M \hat{\lambda}_k - \sum_{k=1}^M \hat{\mu}_k \right) = \sum_{k=1}^M \hat{\lambda}_k \quad (28)$$

is the G-consistent estimate of the trace of R_M , regardless of the ‘‘eigenvalue splitting condition’’ As4.

The proof of this equivalence is shown in Appendix A, and shows that the trace of a sample matrix \hat{R} is always the G-consistent estimate of the trace of the original (true) covariance matrix R_M , even when for some individual eigenvalues of this matrix, their G-consistent estimates may not exist due to As4 not being satisfied.

As an aside for the interested reader, Mestre’s assumptions As1 to As3 (the details of which are provided in [10]) impose general conditions on the covariance matrix and eigenvalues which are always met under our problem with the sample Wishart matrix $\hat{R}_M(\Omega_m)$, but assumption As4 may or may not be met. The physical meaning of the As4 condition, also referred to as the eigenvalue splitting condition, is that the asymptotic eigenvalue distribution of \hat{R} can be considered as a collection of clusters, each one centered around a different eigenvalue of $R_M = \mathcal{E} \hat{R}_M = R^{-\frac{1}{2}}(\Omega_m) R_0 R^{-\frac{1}{2}}(\Omega_m)$. The size of the support of these clusters depends on the ratio of the number of sensors/antennas to the number of samples (ie. the ratio $c = M/T$) from which \hat{R} is constructed. If the ratio M/T is high, the clusters with close eigenvalues γ_j of R_M may merge and be represented by a single estimate with a certain ‘‘multiplicity’’ K . If, on the contrary, the ratio M/T is low, each sample eigenvalue $\hat{\lambda}_k$ will be associated with a different distinct cluster, well separated from the rest of the sample eigenvalue distribution.

This analysis provides an answer, therefore, to the main question posed in the title of this correspondence: Within the existing RMT framework, ‘‘improved’’ G-consistent estimates of individual eigenvalues $\hat{\gamma}_m^G$ and bilinear forms $\hat{\eta}_m^G$ given in (15) cannot make the traditional log-likelihood function ‘‘more G-consistent’’, and therefore provide any improvement into the threshold performance of ML DOA estimation, in full correspondence with Lemma 3.1 from [6] shown in (25) above.

To confirm this further and support numerical evaluation, we still may consider an alternative reconstruction of the G-consistent log-likelihood function using the individual estimates (15) in a manner similar to the G-MUSIC derivation in [9]. Assuming that the ‘‘eigenvalue splitting condition’’ As4 from [10] is satisfied for all the $(m+1)$ different eigenvalues⁵, we may construct the G-LLF (log likelihood function) as

$$g_0[R(\Omega_m)] = \frac{1}{M} \text{tr} R^{-1}(\Omega_m) \hat{R}_m^G, \quad \hat{R}_m^G = \sum_{j=1}^{m+1} \hat{\lambda}_j^G \hat{P}_j^G. \quad (29)$$

where, based on (15), G-consistent estimates of the eigenspace P_j in R_0 ($R_0 = \lambda_1 P_1 + \sum_{j=2}^{m+1} \lambda_j P_j$) are given by

$$\hat{P}_j^G = \sum_{k=1}^M \rho_j(k) \hat{U}_k \hat{U}_k^H; \quad \hat{R} = \sum_{j=1}^M \hat{\lambda}_j \hat{U}_j \hat{U}_j^H \quad (30)$$

Recall that G-MUSIC was calculated as per (20), and therefore (29)-(30) resemble this construction, but with some differences. First of all, in contrast to standard eigendecomposition, we note that individual eigensubspace estimates \hat{P}_j^G in (30) are no longer mutually orthogonal, though they can be shown to sum to the identity matrix since the weights $\rho_j(k)$ in (16) sum to unity. Therefore, if we introduce the diagonal matrix $D(j) \equiv \text{diag}[\rho_j(1), \dots, \rho_j(M)]$, we immediately discover that according to (30), we get

$$\hat{R}_m^G = \hat{U} \left[\sum_{j=1}^{m+1} \hat{\lambda}_j^G D_j \right] \hat{U}^H. \quad (31)$$

While individual (G-consistent) estimates \hat{P}_j^G in (30) are different from the ML estimates $\hat{P}_j = \sum_{k \in K_j} \hat{U}_k \hat{U}_k^H$, the eigenvectors in \hat{R}_m^G are *exactly the same* as in the ML estimation \hat{R}_m^{ML} (Theorem 9.3.1 in [13]).

⁵Recall that R_0 (as opposed to the whitened covariance matrix R_M) consists of an m dimensional signal subspace and a noise subspace of dimension $M - m$

Let us demonstrate that “reconstructed” eigenvalues λ_k^R in $\sum_{j=1}^{m+1} \hat{\lambda}_j^G D_j = \text{diag}[\hat{\lambda}_1^R, \dots, \hat{\lambda}_k^R]$ (at least G-asymptotically) tend to the same limit as the (not G-consistent) MLE estimate $\hat{\gamma}_k^{ML}$:

$$\lim_{\substack{M, T \rightarrow \infty \\ M/T \rightarrow c}} \hat{\lambda}_k^R \Rightarrow \lim_{\substack{M, T \rightarrow \infty \\ M/T \rightarrow c}} \hat{\gamma}_j^{ML} = \gamma_k \left(1 + \frac{c}{M} \sum_{\substack{r=1 \\ r \neq k}}^{m+1} K_r \frac{\gamma_r}{\gamma_r - \gamma_k} \right) \quad (32)$$

Using (15)-(16), for the noise subspace eigenvalues $(\hat{\lambda}_1^R, \dots, \hat{\lambda}_{M-m}^R)$, we get

$$\hat{\lambda}_k^R = \hat{\lambda}_k^G \left(1 + \sum_{\substack{r=1 \\ r \neq k}}^M \frac{\hat{\lambda}_r}{\hat{\lambda}_k - \hat{\lambda}_r} - \frac{\hat{\mu}_r}{\hat{\lambda}_k - \hat{\mu}_r} \right) - \sum_{\substack{r=1 \\ r \neq k}}^M \hat{\lambda}_r^G \left[\frac{\hat{\lambda}_r}{\hat{\lambda}_k - \hat{\lambda}_r} - \frac{\hat{\mu}_r}{\hat{\lambda}_k - \hat{\mu}_r} \right], \quad (33)$$

for $k = 1, \dots, (M - m)$ and with $\hat{\lambda}_1^G = \hat{\lambda}_2^G = \dots = \hat{\lambda}_{M-m}^G$, which is identical to

$$\hat{\lambda}_k^R = \hat{\lambda}_k^G + \hat{\lambda}_k \sum_{\substack{r=1 \\ r \neq k}}^M \frac{(\hat{\lambda}_k^G - \hat{\lambda}_r^G)(\hat{\lambda}_r - \hat{\mu}_r)}{(\hat{\lambda}_k - \hat{\lambda}_r)(\hat{\lambda}_k - \hat{\mu}_r)}, \quad k = 1, \dots, (M - m). \quad (34)$$

According to (15), for signal subspace eigenvalues ($r \geq M - m$), we get $\hat{\lambda}_r - \hat{\mu}_r = \frac{1}{T} \hat{\lambda}_r^G$, and therefore we have

$$\hat{\lambda}_k^R = \hat{\lambda}_k^G - \frac{\hat{\lambda}_k}{T} \sum_{\substack{r=1 \\ r \neq k}}^M \frac{(\hat{\lambda}_k^G - \hat{\lambda}_r^G) \hat{\lambda}_r^G}{(\hat{\lambda}_k - \hat{\lambda}_r)(\hat{\lambda}_k - \hat{\mu}_r)}. \quad (35)$$

Note that for $\hat{\lambda}_{r-1} \ll \hat{\mu}_r < \hat{\lambda}_r$ and $\hat{\mu}_r \rightarrow \frac{T-1}{T} \hat{\lambda}_r$ when As4 is strongly met. Therefore the difference between (35) and the expression (see (11) and (32))

$$\hat{\lambda}_k^R = \hat{\lambda}_k^G - \frac{\hat{\lambda}_k}{T} \sum_{\substack{r=1 \\ r \neq k}}^{m+1} \frac{(\hat{\lambda}_k^G - \hat{\lambda}_r^G) \hat{\lambda}_r^G}{(\hat{\lambda}_k^G - \hat{\lambda}_r^G)(\hat{\lambda}_k^G - \hat{\lambda}_r^G)} \quad (36)$$

that strictly tends to $\bar{\gamma}_k^{ML}$ in (11) under asymptotic condition (1), is only a difference of the second order (in $\frac{1}{T}$) magnitude.

Therefore, the reconstructed noise subspace eigenvalues in \hat{R}_M^G (31) G-asymptotically are not strictly identical, but the fluctuations around its asymptotic value $\bar{\gamma}_1^{ML}$ is quite negligible. For signal subspace eigenvalues, these fluctuations are even less significant, which may be demonstrated by derivations similar to those above.

While it would be more satisfying if $\hat{\lambda}_k^R$ was found to be strictly equal to the ML version $\hat{\gamma}_k^{ML}$, this is not the case, as observed by other researchers in the field [12]. But practically speaking, it is not significant that non G-consistent estimates $\hat{\gamma}_k^{ML}$ are replaced in (29) by (slightly) different non G-consistent estimates $\hat{\lambda}_k^R$. The main point is that the reconstruction in (29) using the G-consistent estimates of eigenvalues and eigenvectors (15) should be at best as efficient as the conventional accurate MLE DOA estimate.

Therefore, the conducted analysis has demonstrated how and why the G-consistent individual estimates of eigenvalues and eigenvectors help to improve MUSIC threshold performance, but still lead to the traditional ML performance. This property is proven rigorously for direct G-consistent estimation of $\frac{1}{M} \text{tr}(R^{-1}(\Omega_m)R_0)$, and in first order (in $\frac{1}{T}$) approximation for the reconstruction approach in (29). Let us now support these theoretical findings by empirical results.

3 Numerical Evaluation

Here, we seek by simulation to demonstrate, using direct exhaustive search for the global maximum of the likelihood function and its reconstructed G-modification, that their DOA estimation performance in the threshold region is identical. For this analysis, we have selected the same scenario as in [1] with $M = 8$ sensor ULA, $T = 10$ i.i.d. training samples and two equipower sources ($p_1 = p_2$) with DOAs $\{0^\circ, 0.4^\circ\}$ in white noise power $p_0 = 1$. The number of sources $m = 2$ is limited by the computational load associated with the exhaustive m -dimensional search for the LF global maximum, but still demonstrates the relevant MLE threshold behavior. For this scenario, the CRB resolution limit (as predicted by the CRB RMSE of 0.2° , derived in more detail in [1]) occurs at an SNR of approximately 40dB ($\varepsilon_{\text{CRB}} = 40\text{dB}$).

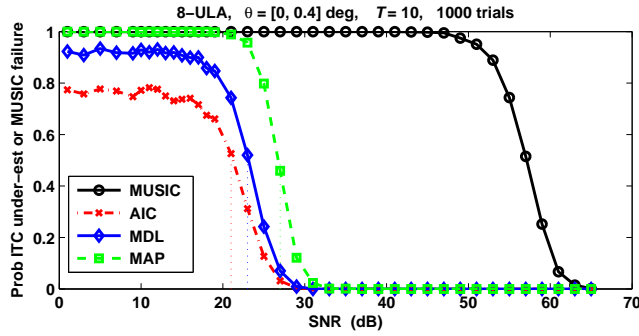


Figure 1. Sample MUSIC breakdown for closely spaced sources. Note the SNR gap between reliable detection using ITC and reliable estimation using MUSIC.

Fig. 1 plots the probability of underestimating the correct number of sources as a function of SNR for the AIC, MDL, and MAP information theoretic criteria (ITC). The same figure shows the sample probability of conventional MUSIC breakdown, which illustrates the well-documented fact that MUSIC “breaks” at SNR values significantly larger than the actual MLE threshold [8]. In this paper, we examine the performance at $\varepsilon_{ITC} = 30\text{dB}$, where Fig. 1 shows the reliable detection of the correct number of sources $m = 2$, but the ML DOA estimate itself is questionable.

It can be separately confirmed, following the methodology in [10], that for this scenario, the eigenvalue splitting condition As4, required for G-consistency of the eigenvalue estimation, are met for the 30dB case in this scenario. The fact that this condition is met here comes with no surprise. Indeed, the reliable detection of the correct number of sources $m = 2$ at this SNR already implies that the cluster of noise subspace sample eigenvalues ($\hat{\lambda}_1$ to $\hat{\lambda}_6$) is reliably separated from the cluster of the smallest signal eigenvalue $\hat{\lambda}_7$. Since $\lambda_8 \gg \lambda_7$ for our scenario with closely spaced sources, it is really the separation between the λ_7 cluster and the noise subspace one ($\lambda_1 = \dots = \lambda_6 = 1$), which is of concern for evaluation of the As4 condition.

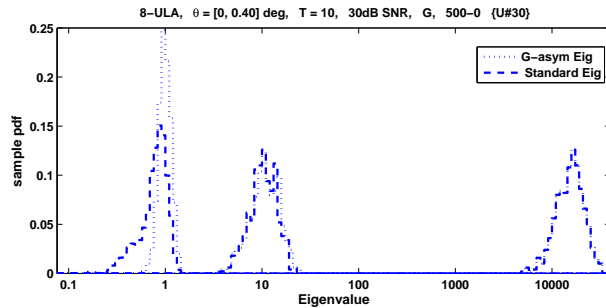


Figure 2. Sample p.d.f.s for the eigenvalues of \hat{R}_m^G

We therefore examine the eigendecomposition of the G-reconstructed covariance matrix \hat{R}^G . As predicted, the eigenvectors of \hat{R}_m and \hat{R}_m^G are exactly the same, and only the eigenvalues differ slightly (and then only in the noise subspace). The two subspace eigenvalues are also well separated both from each other and the noise subspace eigenvalues. While in \hat{R}_m the noise subspace is described by a single value ($= \frac{1}{M-m} \sum_{j=1}^{M-m} \hat{\lambda}_j$), in \hat{R}_m^G they remain close but not precisely equal to $\hat{\lambda}_1$ (i.e. $\text{eig}_1(\hat{R}_m^G) < \hat{\lambda}_1^G < \text{eig}_6(\hat{R}_m^G)$). This is demonstrated in Fig. 2, where we introduce sample p.d.f.s for $\text{eig}_7(\hat{R}_m^G)$, $\text{eig}_8(\hat{R}_m^G)$, and the noise eigenvalues $\text{eig}_n(\hat{R}_m^G) = (\text{eig}_1(\hat{R}_m^G), \dots, \text{eig}_6(\hat{R}_m^G))$, where here we treat as the noise subspace eigenvalue in \hat{R}_m^G any one of its six smallest eigenvalues. The p.d.f.s for $\hat{\lambda}_1$ and $\text{eig}_n(\hat{R}_m^G)$ practically coincide.

While Fig. 2 does demonstrate minor variations in the noise subspace eigenvalues in \hat{R}_m^G around the single noise subspace eigenvalue $\hat{\lambda}_1^G$ in \hat{R}_m , caused by finite T and M values, these variations are very unlikely to result in any noticeable changes in DOA estimation performance when \hat{R}_m^G is used instead of \hat{R}_m or \hat{R} in the likelihood function, simply because even larger variations of noise subspace eigenvalues in \hat{R} around a single ML estimate ($= \frac{1}{M-m} \sum_{j=1}^{M-m} \hat{\lambda}_j$) in \hat{R}_m are also not associated with any DOA ML estimation performance degradation.

Yet, since all the conditions (As1-As4) of the Theorem 3 [10] are met, we can expect that these minor variations do result in improved individual eigenvalue $\hat{\lambda}_j^G$ and eigensubspace \hat{P}_j^G estimation accuracy. To confirm this, we repeat the analysis conducted in [10], where to assess the quality of the eigensubspace estimates, Mestre introduced an *orthogonality factor*:

$$O(j) = \frac{|\text{tr}[P_j \hat{P}_j]|}{|\text{tr}[I_n - P_j] \hat{P}_j|}. \quad (37)$$

Here $P_j, j = 1, 2, 3$ are the eigenvectors/subspaces of the true covariance matrix R_0 and \hat{P}_j are the estimates

(either standard ML ones or the G-consistent \hat{P}_j^G ones). This orthogonality factor is the ratio of the total power of the estimated subspace \hat{P}_j that resides in the true subspace P_j to the power of its “leakage” into the true orthogonal subspace $[I_M - P_j]$.

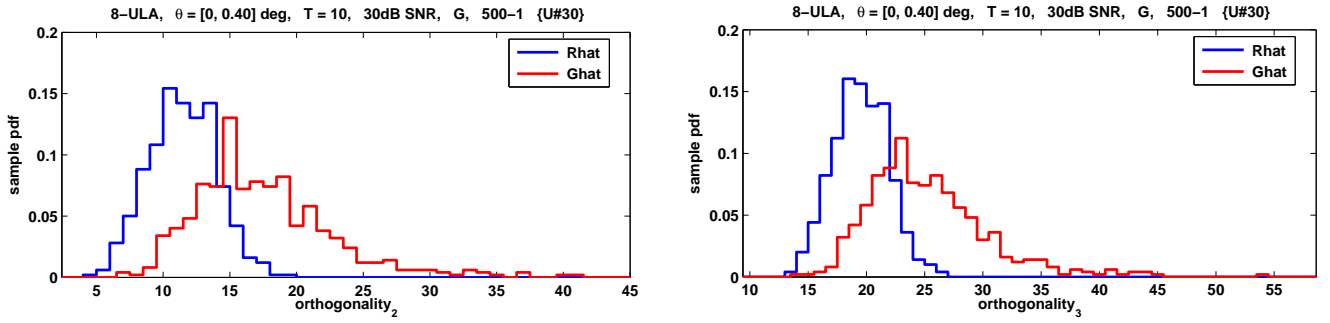


Figure 3. Sample p.d.f.s for MLE and GMLE orthogonality factors for the two signal eigenvectors.

The estimate \hat{P}_j^G (as defined in (30)) is “better” than the MLE \hat{P}_j , defined as $\hat{P}_1 = \sum_{j=1}^6 \hat{U}_j \hat{U}_j^H$; $\hat{P}_2 = \hat{U}_7 \hat{U}_7^H$; $\hat{P}_3 = \hat{U}_8 \hat{U}_8^H$ if, with high probability, the respective orthogonality factor is greater. In Fig. 3, we introduce sample p.d.f.s for the orthogonality factor $O(j)$ and $O(j)^G$, averaged over the same 500 Monte-Carlo trials for the SNR=30dB case. Similarly to [10], we observe that the G-consistent subspace estimate achieves a much higher orthogonality factor than the traditional ML sample estimator, indicating a better capacity to estimate subspaces, and therefore better accuracy for subspace-based DOA estimation techniques such as MUSIC.

To demonstrate, we compare the results of exhaustive global maximum search for the conventional (normalized) likelihood function (ratio):

$$LR[R(\Omega_m)] = \frac{\det R^{-1}(\Omega_m) \hat{R} e^M}{e^{\text{tr} R^{-1}(\Omega_m) \hat{R}}} \quad (38)$$

and its G-asymptotic counterpart $GLR[R(\Omega_m)]$ constructed using \hat{R}_m^G instead of \hat{R} .

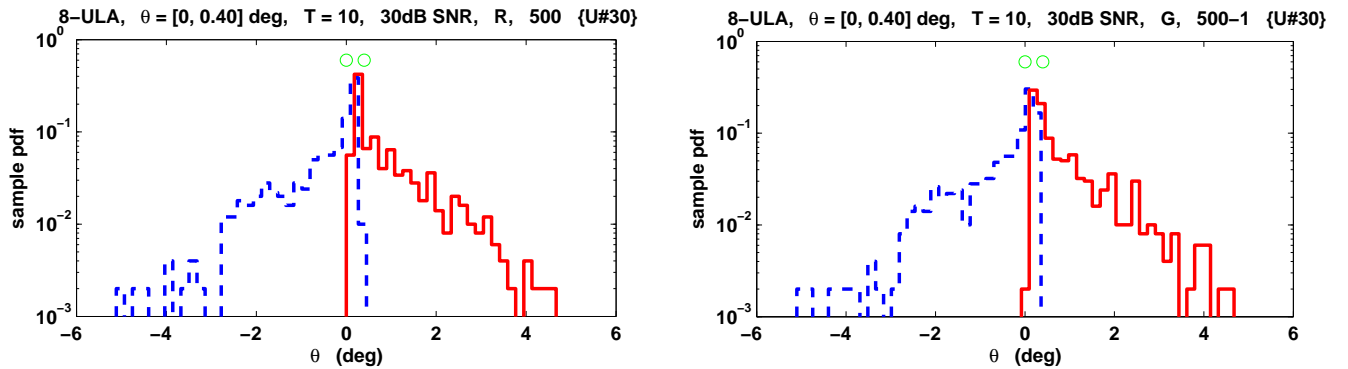


Figure 4. Sample p.d.f.s for MLE and GMLE DOA estimation at 30dB SNR; a- \hat{R}_m , b- \hat{R}_m^G .

This global optimization is performed in two steps. In step 1, we find the global extremum over a fine grid on a half-plane ($\theta_1 < \theta_2$), and in step 2, we conduct local optimization over the parameters $\theta_1 < \theta_2$ as well as the two source powers using the MATLAB `fmincom` routine. The white noise power is assumed known *a priori* and set to unity. More details on the optimization routine and MLE results can be found in [1, 2, 14]. Due to our convention that $\hat{\theta}_1 < \hat{\theta}_2$, the DOA sample distributions illustrated by Fig. 4 overlap, since for scenarios with very closely spaced sources, over the 500 trials, the minimum $\hat{\theta}_2$ was sometimes less than the maximum $\hat{\theta}_1$, even though on each individual trial θ_1 was less than θ_2 . One can see that the distribution for the DOAs derived using the standard LR (38) and the G-asymptotic counterpart are practically indistinguishable, as are any moments of these distributions. Moreover, in addition to the identical statistical performance of the DOA estimates, identical individual estimates were registered in a trial-by-trial basis.

In addition to the illustrated results, the same results were observed at several higher SNRs, and were also the same when the ML covariance matrix estimate $\hat{R}_m = \sum_{k=1}^M \hat{\gamma}_m^{ML} P_m^{ML}$ was used. Finally, in a completely ad-hoc “mix and match” approach, we also assigned the G-consistent eigenvalue estimates $\hat{\lambda}_1^G$ (the G-consistent noise power estimate), $\hat{\lambda}_2^G$ and $\hat{\lambda}_3^G$ (15) to the maximum likelihood (mutually orthogonal) eigensubspaces \hat{P}_1 , \hat{P}_2 , and \hat{P}_3 and once again did not record any DOA estimation improvements, showing that the minor variations in noise subspace eigenvalues in \hat{R}_m^G with respect to \hat{R}_m was not significant in MLE.

4 Concluding Remarks

The theoretical results from RMT and their examination by direct Monte-Carlo simulations has confirmed that for Gaussian sources in Gaussian noise, ML DOA estimation in the so-called “threshold” region is not improved by use of recent results from Random Matrix Theory. More specifically, we demonstrated that the Girko Lemma 3.1 from [6] implies that for the multi-variate (complex) Gaussian data case, the stochastic (unconditional) likelihood function is both a T -consistent and G-consistent estimate of the function $1/M \operatorname{tr}[R^{-1}(\Omega_m)R_0]$, where R_0 is represented by a set of T i.i.d. training vectors \mathbf{x}_j .

We demonstrated that in full accordance with this Lemma, an alternative reconstruction of the log-likelihood function using individually G-consistent estimates of eigenvalues and eigenvectors does not improve DOA estimation threshold performance, despite the fact that these individual G-consistent estimates are more accurate than their corresponding ML estimates. Thus, the demonstrated in [9–11] improvement in MUSIC performance delivered by the G-consistent estimates of the individual noise eigensubspace is not at odds with the lack of improvement in MLE performance.

While applicable to arbitrary parametric description of the covariance matrix R_0 , these results should not be extended beyond the multivariate (complex) i.i.d. Gaussian case without detailed examination. Finally, it is important to stress that our results are confined to the currently (in [10]) unique set of G-consistent eigenvalue and eigenvector estimates, and therefore the broader question on whether the threshold performance of ML DOA estimation can be improved by some other technique remains open.

Acknowledgements

The authors express their gratitude to Dr. Xavier Mestre of the Telecommunications Technological Center of Catalonia (CTTC) for his guidance on RMT results.

Appendix

Proof of Theorem 1

The G-consistent eigenvalue estimators defined in Theorem 3 in [10] ((15) in the main body of the paper) can be restructured as

$$\sum_{m=1}^{\bar{M}} K_m \hat{\gamma}_m^G = T \left(\sum_k^M \hat{\lambda}_k - \sum_k^M \hat{\mu}_k \right), \quad (4.1)$$

and therefore we have to prove that

$$T \sum_{k=1}^M \hat{\mu}_k = (T-1) \sum_{k=1}^M \hat{\lambda}_k. \quad (4.2)$$

Given eqn. (17) in the main body of the paper, we may define

$$T \sum_{k=1}^M \hat{\mu}_k = \sum_{k=1}^M \sum_{m=1}^M \frac{\hat{\lambda}_m \hat{\mu}_k}{\hat{\lambda}_m - \hat{\mu}_k}. \quad (4.3)$$

Substitution shows that instead of the form in eqn. (17), $\hat{\mu}_1, \dots, \hat{\mu}_M$ can be introduced as the roots of the following polynomial:

$$\begin{aligned} Q(\mu) &= \prod_{k=1}^M (\hat{\mu}_k - \mu) \\ &= \prod_{j=1}^M (\hat{\lambda}_j - \mu) - \frac{1}{T} \sum_{r=1}^M \hat{\lambda}_r \prod_{\substack{j=1 \\ j \neq r}}^M (\hat{\lambda}_j - \mu) \end{aligned} \quad (4.4)$$

where the right-hand side expression is obtained from eqn. (17). Evaluating $Q(\mu)$ at $\mu = \lambda_m$, we obtain

$$\prod_{k=1}^M (\hat{\lambda}_m - \hat{\mu}_k) = \frac{1}{T} \hat{\lambda}_m \prod_{\substack{j=1 \\ j \neq m}}^M (\hat{\lambda}_m - \hat{\lambda}_j). \quad (4.5)$$

On the other hand, observing the form of the derivative of $Q'(\mu)$:

$$\begin{aligned} Q'(\mu) &= - \sum_{l=1}^M \prod_{\substack{k=1 \\ k \neq l}}^M (\hat{\mu}_k - \mu) = \\ &= - \sum_{l=1}^M \prod_{\substack{j=1 \\ j \neq l}}^M (\hat{\lambda}_j - \mu) + \frac{1}{T} \sum_{r=1}^M \hat{\lambda}_r \sum_{\substack{l=1 \\ l \neq r}}^M \prod_{\substack{j=1 \\ j \neq r}}^M (\hat{\lambda}_j - \mu) \end{aligned} \quad (4.6)$$

and evaluating it at $\mu = \hat{\lambda}_m$, we obtain

$$\begin{aligned} \sum_{l=1}^M \prod_{\substack{k=1 \\ k \neq l}}^M (\hat{\mu}_k - \hat{\lambda}_m) &= - \prod_{\substack{j=1 \\ j \neq m}}^M (\hat{\lambda}_j - \hat{\lambda}_m) - \frac{1}{T} \hat{\lambda}_m \sum_{\substack{l=1 \\ l \neq m}}^M \prod_{\substack{j=1 \\ j \neq m}}^M (\hat{\lambda}_j - \hat{\lambda}_m) \\ &= - \frac{1}{T} \sum_{\substack{r=1 \\ r \neq m}}^M \hat{\lambda}_r \prod_{\substack{j=1 \\ j \neq r}}^M (\hat{\lambda}_j - \hat{\lambda}_m) \end{aligned} \quad (4.7)$$

Dividing both sides of (4.7) by $\prod_{\substack{j=1 \\ j \neq r, m}}^M (\hat{\lambda}_j - \hat{\lambda}_m)$ and using (4.5) to simplify the left-hand side, we obtain

$$\frac{1}{T} \sum_{l=1}^M \frac{\hat{\lambda}_m}{\hat{\lambda}_m - \hat{\mu}_l} = 1 - \frac{1}{T} \sum_{\substack{r=1 \\ r \neq m}}^M \frac{\hat{\lambda}_m + \hat{\lambda}_r}{\hat{\lambda}_r - \hat{\lambda}_m} \quad (4.8)$$

(Note that (4.4)-(4.8) follow Appendix IV in in [10]). On the other hand, from (4.3) we get

$$\sum_{k=1}^M \sum_{j=1}^M \frac{\hat{\lambda}_k \hat{\mu}_j}{\hat{\lambda}_k - \hat{\mu}_j} = \sum_{k=1}^M \hat{\lambda}_k \left[\sum_{j=1}^M \frac{\hat{\lambda}_k}{\hat{\lambda}_k - \hat{\mu}_j} - 1 \right] \quad (4.9)$$

From (4.8), it then follows that

$$\sum_{k=1}^M \frac{\hat{\lambda}_k}{\hat{\lambda}_k - \hat{\mu}_j} = T - \sum_{\substack{r=1 \\ r \neq k}}^M \frac{\hat{\lambda}_k + \hat{\lambda}_r}{\hat{\lambda}_r - \hat{\lambda}_k} \quad (4.10)$$

Since

$$\sum_{\substack{r=1 \\ r \neq k}}^M \frac{\hat{\lambda}_k + \hat{\lambda}_r}{\hat{\lambda}_r - \hat{\lambda}_k} + 1 - 1 = \sum_{\substack{r=1 \\ r \neq k}}^M \left(\frac{2\hat{\lambda}_r}{\hat{\lambda}_r - \hat{\lambda}_k} - 1 \right) \quad (4.11)$$

we get

$$\sum_{k=1}^M \sum_{j=1}^M \frac{\hat{\lambda}_k \hat{\mu}_j}{\hat{\lambda}_k - \hat{\mu}_j} = (T - M + M - 1) \sum_{k=1}^M \hat{\lambda}_k - \sum_{k=1}^M \sum_{\substack{r=1 \\ r \neq k}}^M \frac{2\hat{\lambda}_k \hat{\lambda}_r}{\hat{\lambda}_r - \hat{\lambda}_k}. \quad (4.12)$$

But for $\hat{\lambda}_k > 0$, the following identity is true (see [10])

$$\sum_{k=1}^M \sum_{\substack{r=1 \\ r \neq k}}^M \frac{\hat{\lambda}_k \hat{\lambda}_r}{\hat{\lambda}_r - \hat{\lambda}_k} = 0. \quad (4.13)$$

Therefore, we get

$$T \sum_{k=1}^M \hat{\mu}_k = (T - 1) \sum_{k=1}^M \hat{\lambda}_k \quad (4.14)$$

as required by (4.2). \square

REFERENCES

- [1] Y.I. Abramovich, B.A. Johnson, and N.K. Spencer. Statistical nonidentifiability of close emitters: Maximum-likelihood estimation breakdown. In *Proc. EUSIPCO-2009*, pages 1968–1972, Glasgow, UK, 2009.
- [2] Y.I. Abramovich, B.A. Johnson, and N.K. Spencer. Statistical nonidentifiability of close emitters: Maximum-likelihood estimation breakdown and its GSA analysis. In *Proc. ICASSP-2009*, pages 2133–2136, Taipei, Taiwan, 2009.
- [3] F. Athley. Threshold region performance of maximum likelihood direction of arrival estimators. *IEEE Trans. Signal Processing*, 53(4):1359–1373, Apr 2005.
- [4] P. Forster, P. Larzabal, and E. Boyer. Threshold performance analysis of maximum likelihood DOA estimation. *IEEE Trans. Sig. Proc.*, 52(11):3183–3191, Nov 2004.
- [5] A.B. Gershman. Pseudo-randomly generated estimator banks: a new tool for improving the threshold performance of direction finding. *IEEE Trans. Sig. Proc.*, 46(5):1351–1364, May 1998.
- [6] V.L. Girko. *An Introduction to Statistical Analysis of Random Arrays*. VSP, Utrecht, Netherlands, 1998.
- [7] R.A. Janik and M.A. Nowak. Wishart and anti-Wishart random matrices. *J. Phys. A*, 36:3629–3637, 2003.
- [8] B.A. Johnson, Y.I. Abramovich, and X. Mestre. MUSIC, G-MUSIC, and maximum-likelihood performance breakdown. *IEEE Trans. Sig. Proc.*, 56(8):3944–3958, August 2008.
- [9] X. Mestre. An improved subspace based algorithm for small sample size regime. In *Proc. ICASSP-06*, Toulouse, 2006.
- [10] X. Mestre. Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates. *IEEE Trans. Info. Theory*, 54(11):5113–5129, November 2008.
- [11] X. Mestre. On the asymptotic behaviour of the sample estimates of eigenvalues and eigenvectors of covariance matrices. *IEEE Trans. Sig. Proc.*, 56(11):5353–5368, November 2008.
- [12] Xavier Mestre. E-mail communications. March 05, 2010.
- [13] R.J. Muirhead. *Aspects of Multivariate Statistical Theory*. Wiley, New York, 1982.
- [14] P. Parvazi, A.B. Gershman, and Y. Abramovich. Detecting outliers in the estimator bank-based direction finding techniques using the likelihood ratio quality assessment. In *Proc. ICASSP*, volume II, pages 1065–1068, Honolulu, Hawaii, April 2007. IEEE.
- [15] Christ D. Richmond. Mean-squared error performance prediction of maximum-likelihood signal parameter estimation. In *Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop*, Lexington, MA, 11-13 Mar 1998. MIT-LL. DTIC Ascension ADA421709.
- [16] D. C. Rife and R. R. Boorstyn. Multiple tone parameter estimation from discrete-time observations. *Bell System Technical Journal*, 55(9):1389–1410, Nov 1976.
- [17] S. Zacks. *The Theory of Statistical Inference*. Wiley, New York, 1971.
- [18] K. Zarifi and A.B. Gershman. Asymptotic performance analysis of blind minimum output energy receivers for large DS-CDMA systems. *IEEE Trans. Sig. Proc.*, 56(2):650 – 663, Feb 2008.