THE ROLE OF SUBSPACE SWAP IN MUSIC PERFORMANCE BREAKDOWN

*Ben A. Johnson** and *Yuri I. Abramovich†* and *Xavier Mestre‡*

* RLM, Pty. Ltd. Edinburgh, Australia and ITR/Univ. South Australia, e-mail: `ben.a.johnson@ieee.org`
†DSTO, ISR Division, Edinburgh, Australia, e-mail: `yuri.abramovich@dsto.defence.gov.au`
‡Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain, e-mail: `xavier.mestre@cttc.cat`

ABSTRACT

Direction-of-arrival estimation performance of MUSIC in the so-called "threshold" area is often attributed to the "subspace swap" phenomena. We show that the subspace swap condition can be accurately predicted using recent results from Random Matrix Theory (RMT) analysis, but that subspace "leakage" rather than full subspace swap is associated with the onset of performance degradation in closely spaced multiple source scenarios. Prediction of the "subspace swap" phenomena is examined analytically and by the use of Monte-Carlo simulation.

Index Terms— Direction of arrival estimation, adaptive arrays, error analysis, maximum likelihood estimation.

1. INTRODUCTION

It has also been known for a long time that when the sample support T (and/or signal-to-noise ratio) is insufficient, MUSIC performance "breaks down" and rapidly departs from the CRB ([1, 2]. The typical manifestation of this breakdown is the appearance of severely erroneous DOA estimates ("outliers") that dramatically degrade the overall estimation accuracy. Studies of MUSIC performance breakdown have focused on the so-called "subspace-swap" phenomena, whereby "the measured data is better approximated by some components of the orthogonal subspace than by some components of the signal subspace" [3, 4].

It has also been demonstrated [5] that, at least in multiple source scenarios, there typically is a significant "gap" in required sample support and/or SNR between the MUSIC-specific and ML-intrinsic threshold conditions. Thus, potentially different mechanisms are responsible for MLE and MUSIC "breakdowns", and the relevance of subspace swap needs to be investigated.

2. SUBSPACE SWAP AND PERFORMANCE BREAKDOWN

Since MLE and MUSIC performance distinctions disappear in classic asymptotic studies [6], it is necessary to consider other analysis approaches. In [7], an improvement in MUSIC "threshold performance" has been derived by one of us (Mestre), based on recent findings of the General Statistical Analysis (GSA) approach (also known as Random Matrix Theory (RMT)) that considers different asymptotic conditions:

$$\lim_{M,T\to\infty} M/T \to \text{constant} < \infty. \tag{1}$$

ie. where both the array dimension M and the number of snapshots T grow without bound, but at the same rate.

This work was partially conducted under DSTO/RLM Collaboration Agreement 290905

It is important to note that the G-MUSIC derivations and Monte-Carlo simulations were conducted under conditions which *"guarantee separation of the noise and first signal eigenvalue cluster of the asymptotic eigenvalue distribution of* $\hat{R}$ [7]. Therefore, G-asymptotically the "subspace swap" phenomenon is precluded by these conditions, and yet MUSIC and G-MUSIC breakdown was regularly observed in the conducted Monte-Carlo trials under these conditions.

These distinctions between MUSIC/G-MUSIC and MLE threshold conditions are supported with further simulation studies introduced in Fig. 2 and based on a scenario with a $M = 20$-element uniform linear array (ULA), $T = 15$ training samples, and $m = 4$ independent equal power Gaussian sources (stochastic source model) located at azimuth angles

$$\theta_m = \{-20^o, -10^o, 35^o, 37^o\}, \tag{2}$$

and immersed in white noise.

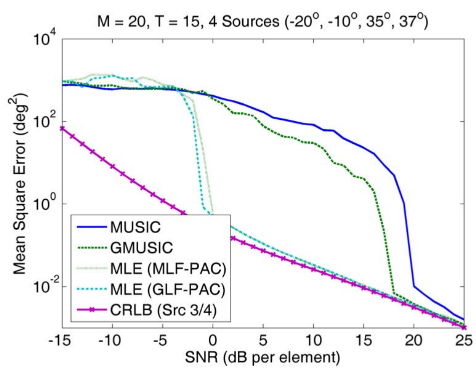

Fig. 1. Multiple-source estimation on a 20-element uniform linear array with $T = 15$ training samples for MUSIC, G-MUSIC and MLE. The SNR breakpoint (the "threshold") decreases from around 20 dB for MUSIC to 17 dB for G-MUSIC, but is still dramatically greater than the MLE-proxy (LF-PAC) threshold observed at around 0 dB.

One of the important results of GSA/RMT analysis relates to convergence of the eigenvalue empirical distribution of the sample covariance matrix, which tends almost surely to a deterministic probability density G-asymptotically. It has been shown [8] that for the m-th eigenvalue $\lambda_m, m = 1, \ldots, \overline{M}$ of the $\overline{M}$ distinct true eigenvalues (which occur with multiplicity K_m) to be estimated (ie. for the cluster of $\hat{\lambda}_m$ to be well separated from the neighboring $\hat{\lambda}_{m-1}$ cluster), that

$$T/M > \xi(m) \text{ where } \xi(m) = \max_{j\in(m,m-1)} \beta(j) \tag{3}$$

$$\beta_j = \frac{1}{M}\sum_{r=1}^{M} K_r \left(\frac{\lambda_r}{\lambda_r - f_j}\right)^2 \tag{4}$$

for $0 < j < \overline{M}; \beta(0) = \beta(\overline{M}) = 0$ (the "eigenvalue splitting condition"). The factor $f_j, j = 1, \ldots, \overline{M}-1$ denotes the $\overline{M}-1$ real-valued solutions of

$$\frac{1}{M}\sum_{m=1}^{\overline{M}} K_m \frac{\lambda_m^2}{(\lambda_m - f_j)^3} = 0 \tag{5}$$

ordered as $\overline{f}_1 > \overline{f}_2 > \ldots > \overline{f}_{min}$. Specifically, the ratio of the number of training samples T to the dimension of the array M necessary to guarantee that the eigenvalue cluster associated with the noise subspace eigenvalue $\lambda_{m+1} = \ldots = \lambda_M$ is separated from the rest of the eigenvalue distribution (the "subspace splitting condition") is given by

$$T/M > \frac{1}{M}\sum_{r=1}^{M}\left(\frac{\lambda_r}{\lambda_r - f_{min}}\right)^2 \tag{6}$$

where f_{min} denotes the minimum real-valued solution to equation (5), considering multiplicity.

If the number of samples T per antenna element M is greater than the right hand side of (6), one can ensure that signal and noise sample eigenvalues will be separated in the asymptotic sample eigenvalue distribution, and a subspace swap will occur with probability zero. It turns out that whenever there is asymptotic separation between signal and noise subspaces, one can effectively describe the behavior of the sample eigenvalues and eigenvectors using the following result from [7]: Let $\mathbf{x}_1, ..., \mathbf{x}_T$ be i.i.d. complex-valued column vectors from the M-variate distribution with circularly symmetric complex random variables having zero mean and covariance matrix R_M, that has the following eigen-decomposition

$$R_M = E_M \Sigma_M E_M^{\mathrm{H}} \tag{7}$$

$$\Sigma_M = diag(\lambda_1, \ldots, \lambda_m, \lambda_M, \ldots, \lambda_M) \tag{8}$$

where $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_m$ are the true signal eigenvalues and E_M is the corresponding eigenvector matrix. Let the sample matrix $\hat{\Sigma}_M$ be specified as

$$\hat{\Sigma}_M = \frac{1}{T}\sum_{j=1}^{T}\mathbf{x}_j\mathbf{x}_j^{\mathrm{H}} = \hat{E}_M\hat{\Sigma}_M\hat{E}_M^{\mathrm{H}}. \tag{9}$$

Consider the jth signal sample eigenvector $\hat{\mathbf{e}}_j$ (assumed to be associated with a sample eigenvalue with multiplicity one) and a deterministic column vector $\mathbf{s}$. One can try to analyze the behavior of the sample eigenvector $\hat{\mathbf{e}}_j$ by studying the behavior of the scalar product $\langle \mathbf{s}^{\mathrm{H}}\hat{\mathbf{e}}_j\rangle$, and relate it somehow to the deterministic quantity $\langle \mathbf{s}^{\mathrm{H}}\mathbf{e}_j\rangle$. It turns out that, as $M, T \to \infty$ at the same rate under a satisfied "eigenvalue splitting condition" (4) for all eigenvalues, we get

$$\left| \left|\langle \mathbf{s}^{\mathrm{H}}\hat{\mathbf{e}}_j\rangle\right|^2 - \sum_{k=1}^{M} w_j(k)\left|\langle \mathbf{s}^{\mathrm{H}}\mathbf{e}_k\rangle\right|^2 \right| \to 0 \tag{10}$$

almost surely, where the weights $w_j(k)$ are defined as ([9], Theorem 2)

$$w_j(k) = \begin{cases} 1 - \frac{1}{K_j}\sum_{\substack{r=1\\ r\neq j}}^{M} K_r\left(\frac{\lambda_j}{\lambda_r-\lambda_j} - \frac{\mu_j}{\lambda_r-\mu_j}\right) & k = j \\ \frac{\lambda_j}{\lambda_k-\lambda_j} - \frac{\mu_j}{\lambda_k-\mu_j} & k \neq j \end{cases} \tag{11}$$

and where μ_j are the real-valued solutions to

$$\frac{1}{M}\sum_{j=1}^{M} K_j \frac{\lambda_j}{\lambda_j - \mu} = \frac{T}{M} \tag{12}$$

repeated according to the multiplicity K_j of the corresponding λ_j. This result is powerful, but allows for little interpretation. In order to simplify the analysis it is common practice to consider the particular case of the so-called "spiked population covariance matrix model". This class of covariance matrix was introduced by Johnstone [10], and it describes the asymptotic behavior of a class of covariance matrices obtained from a limited number of plane waves in noise. Under this simplification of the original model (which implies letting $M \to \infty$ for m fixed in the above formulas), we see that the asymptotic subspace splitting condition in (6) becomes

$$T/M > \left(\frac{\lambda_M}{\lambda_M - \lambda_m}\right)^2 \tag{13}$$

which can also be expressed as

$$\lambda_m > \lambda_M\left(1 + \sqrt{\gamma}\right) \tag{14}$$

where $\gamma = M/T$.

Let us now investigate the behavior of the solutions to (12) under the spiked population covariance model. Note first that (12) can equivalently be written as

$$\frac{1}{M}\left[\sum_{j=1}^{m}\frac{\lambda_j}{\lambda_j - \mu} + (M-m)\frac{\lambda_M}{\lambda_M - \mu_{min}}\right] = \frac{1}{\gamma}. \tag{15}$$

Now, let us first consider μ_{min}. By definition, we have $\mu_{min} < \lambda_M = \ldots = \lambda_{m+1} < \lambda_m \leq \ldots \leqslant \lambda_1$. Hence $\lambda_j - \mu_{min}$ can never go to zero for any $1 \leqslant j \leqslant m$. Consequently, the first term of (15) will go to zero as $M \to \infty$ for a fixed m, and μ_{min} will converge to the solution of

$$\frac{\lambda_M}{\lambda_M - \mu_{min}} = \frac{1}{\gamma} \tag{16}$$

namely,

$$\mu_{min} \to \lambda_M\left(1 - \gamma\right). \tag{17}$$

Let us now consider the convergence of μ_i, $i \leq m$. We observe that $\lambda_{i+1} \leq \mu_i \leq \lambda_i$, so that, by examining the first term in (15), the only possibility is that $\mu_i \to \lambda_i, \quad i \leq m$ (otherwise, the first term of (15) would go to zero, and we would end up with the solution to (16), which is not in the interval of interest). Furthermore, by expressing (15) in the following way

$$\frac{\lambda_j}{M(\lambda_j - \mu_i)} + \frac{1}{M}\sum_{\substack{j=1\\ j\neq i}}^{m}\frac{\lambda_j}{\lambda_j - \mu_i} + \frac{M-m}{M}\frac{\lambda_M}{\lambda_M - \mu_{min}} = \frac{1}{\gamma} \tag{18}$$

or, equivalently, as

$$M\left(\lambda_j - \mu_i\right) = \frac{\lambda_j}{1 - \frac{\gamma}{M}\sum_{\substack{j=1\\ j\neq i}}^{m}\frac{\lambda_j}{\lambda_j-\mu_i} - \gamma\frac{M-m}{M}\frac{\lambda_M}{\lambda_M-\mu_{min}}} \tag{19}$$

we see that (using $\mu_i \to \lambda_i, \quad i \leq m$)

$$M\left(\lambda_i - \mu_i\right) \to \frac{\gamma\lambda_i}{1 - \gamma\frac{\lambda_M}{\lambda_M-\lambda_i}}, \quad i \leq m. \tag{20}$$

With all this, we are now able to investigate the behavior of the weights in (11) under the spiked population model simplification. Indeed, let us first concentrate on the case $1 \leq j \leq m$ (this corresponds to the convergence of a particular signal sample eigenvector). Expressing the weights $w_j(k)$ in the following way

$$w_j(k) = \begin{cases} 1 - \sum_{\substack{r=1 \\ r \neq j}}^{m} \left(\frac{\lambda_j}{\lambda_r - \lambda_j} - \frac{\mu_j}{\lambda_r - \mu_j} \right) \\ \qquad - \frac{(M-m)\lambda_M(\lambda_j - \mu_j)}{(\lambda_M - \lambda_j)(\lambda_M - \mu_j)} & k = j \\ \\ \frac{\lambda_j}{\lambda_k - \lambda_j} - \frac{\mu_j}{\lambda_k - \mu_j} & k \neq j \end{cases} \tag{21}$$

and using the above limits on the μ_j, $1 \leq j \leq m$, we obtain

$$w_j(k) \rightarrow \begin{cases} \frac{1 - \gamma\left(\frac{\lambda_M}{\lambda_M - \lambda_j}\right)^2}{1 - \gamma \frac{\lambda_M}{\lambda_M - \lambda_j}} & k = j \\ 0 & k \neq j \end{cases} \tag{22}$$

Hence, one can ensure that, under the spiked population covariance matrix, and assuming $\lambda_m > \lambda_M \left(1 + \sqrt{\gamma}\right)$, one has

$$\left|\langle \mathbf{s}^{\mathrm{H}} \hat{\mathbf{e}}_j \rangle\right|^2 \rightarrow \frac{1 - \gamma\left(\frac{\lambda_M}{\lambda_M - \lambda_j}\right)^2}{1 - \gamma \frac{\lambda_M}{\lambda_M - \lambda_j}} \left|\langle \mathbf{s}^{\mathrm{H}} \mathbf{e}_j \rangle\right|^2. \tag{23}$$

where $\mathbf{s}$ is a deterministic column vector. This is precisely the result introduced by Paul [11] for the specific class of spiked population covariance matrices, where again a fixed limited number of eigenvalues is greater than the smallest one, whose multiplicity grows with the number of antennas. If we replace $\mathbf{s}$ with an eigenvector of the true covariance matrix $\mathbf{e}_k$, we observe that (under the spiked population covariance matrix model, and assuming asymptotic subspace separation) the projection of a sample eigenvector onto the linear space spanned by an eigenvector associated with a different eigenvalue converges to zero, i.e.

$$|\langle \mathbf{e}_k^{\mathrm{H}} \hat{\mathbf{e}}_j \rangle|^2 \rightarrow \frac{1 - \gamma\left(\frac{\lambda_M}{\lambda_M - \lambda_j}\right)^2}{1 - \gamma \frac{\lambda_M}{\lambda_M - \lambda_j}} \delta_{j-k}. \tag{24}$$

Additionally, Paul also studied the convergence of the sample eigenvector $\hat{\mathbf{e}}_j$ when there is no asymptotic separation between signal and noise subspaces, namely $\lambda_m < \lambda_M(1 + \sqrt{\gamma})$. In particular, he established that, under the spiked population covariance matrix model,

$$\left|\langle \mathbf{e}_j^{\mathrm{H}} \hat{\mathbf{e}}_j \rangle\right|^2 \rightarrow 0 \tag{25}$$

almost surely as $M, T \rightarrow \infty$ at the same rate (in fact, Paul proved this for real-valued Gaussian observations with a diagonal covariance matrix, but we conjecture that the result is also valid for the observation model considered here).

Paul refers to the change in convergence above and below the condition (14) as a "phase transition phenomenon", which is clearly analogous to the subspace swap phenomena known in the signal processing literature for twenty years [12]. Here we have shown that the condition (6), or the simplified one for the spiked population matrix (14), which asymptotically prevents the phase transition phenomenon from occurring, is in fact the condition which guarantees the separability of the signal and noise subspaces in the asymptotic sample eigenvalue distribution. But rather than focus narrowly on the conditions under which the norm of the scalar product between the true and estimated eigenvectors in the signal subspace fall below 0.5, we wish to examine subspace swap, which implies

$$\mathbf{e}_4^{\mathrm{H}} \hat{E}_N \hat{E}_N^{\mathrm{H}} \mathbf{e}_4 > \mathbf{e}_4^{\mathrm{H}} \hat{E}_S \hat{E}_S^{\mathrm{H}} \mathbf{e}_4. \tag{26}$$

(or equivalently $\mathbf{e}_4^{\mathrm{H}} \hat{E}_S \hat{E}_S^{\mathrm{H}} \mathbf{e}_4 < 0.5$), i.e. the last signal eigenvector is better represented by the noise subspace than the signal subspace. In order to predict the value of $\mathbf{e}_j^{\mathrm{H}} \hat{E}_S \hat{E}_S^{\mathrm{H}} \mathbf{e}_j$, let us consider the following Theorem 2 of Mestre [9].

Theorem 2 *If the splitting condition* (6) *for the smallest signal subspace eigenvalue λ_m is satisfied, the random value*

$$\hat{\eta} = \mathbf{s}^{H} \sum_{j=m+1}^{M} \hat{\mathbf{e}}_j \hat{\mathbf{e}}_j^{H} \ \mathbf{s} \tag{27}$$

where $\mathbf{s}$ is a deterministic column vector, asymptotically ($M, T \rightarrow \infty, M/T \rightarrow \gamma$) tends to the non-random value $\overline{\eta}_G$, ie.

$$|\hat{\eta} - \overline{\eta}_G| \xrightarrow{a.s.} 0 \text{ as } M, T \rightarrow \infty, M/T \rightarrow \text{ const.} \tag{28}$$

where

$$\overline{\eta}_G = \mathbf{s}^{H} \sum_{j=m+1}^{M} w(j) \mathbf{e}_j \mathbf{e}_j^{H} \ \mathbf{s} \tag{29}$$

and $\mathbf{e}_j, j = 1, \ldots, M$ are the eigenvectors of the matrix R_M arranged in descending order and

$$w(j) = \begin{cases} 1 - \frac{1}{M-m} \sum_{k=1}^{M} \left(\frac{1}{\lambda_k - 1} + \frac{\mu_{min}}{\lambda_k - \mu_{min}} \right), & j > m \\ \frac{1}{\lambda_j - 1} - \frac{\mu_{min}}{\lambda_j - \mu_{min}}, & j \leqslant m \end{cases} \tag{30}$$

where μ_{min} is the minimal (potentially negative) real-valued solution to

$$\frac{1}{M} \sum_{j=1}^{M} \frac{\lambda_j}{\lambda_j - \mu} = \frac{T}{M} \tag{31}$$

assuming that $\lambda_M = 1$.

This theorem allows us to find the asymptotic MUSIC pseudospectrum if $\mathbf{s} = S(\theta)$ is an antenna steering vector. If instead $\mathbf{s} = \mathbf{e}_j$ (the j-th eigenvector of the actual covariance matrix R_M), from (10), we get

$$\mathbf{e}_j^{\mathrm{H}} \hat{E}_S \hat{E}_S^{\mathrm{H}} \mathbf{e}_j = 1 - \mathbf{e}_j^{\mathrm{H}} \hat{E}_N \hat{E}_N^{\mathrm{H}} \mathbf{e}_j \Rightarrow 1 - w(j) \tag{32}$$

where $w(j)$ is specified by (30). Therefore, we get (for $j \leq m$)

$$\mathbf{e}_j^{\mathrm{H}} \hat{E}_S \hat{E}_S^{\mathrm{H}} \mathbf{e}_j \xrightarrow{a.s.} \frac{\lambda_j}{\lambda_j - \mu_{min}} - \frac{1}{\lambda_j - 1}. \tag{33}$$

For the "spiked population covariance matrix", when $\mu_{min} \rightarrow (1 + \gamma)$ (17) (and $\lambda_M = 1$), we finally get

$$\frac{\lambda_j}{\lambda_j - (1 + \gamma)} - \frac{1}{\lambda_j - 1} = \frac{1 - \gamma \frac{1}{(\lambda_j - 1)^2}}{1 + \gamma \frac{1}{(\lambda_j - 1)}} \tag{34}$$

and for our specific scenario with the minimal signal subspace eigenvalue associated with $\mathbf{e}_4$:

$$\mathbf{e}_4^{\mathrm{H}} \hat{E}_S \hat{E}_S^{\mathrm{H}} \mathbf{e}_4 \xrightarrow{a.s.} \left(1 - \frac{\gamma}{(\lambda_4 - 1)^2}\right) / \left(1 + \frac{\gamma}{(\lambda_4 - 1)}\right). \tag{35}$$

We now get the same asymptotic expression as in (24), but for the projection onto the entire sample subspace. This means that when "intra-subspace swap" is precluded by having the eigenvalue splitting condition (4) satisfied for all signal subspace eigenvectors, the power (24) of the true eigenvector $\mathbf{e}_4$ asymptotically resides in the fourth sample subspace eigenvector $\hat{\mathbf{e}}_4$, while the remaining power resides in the sample noise subspace. If instead only the subspace splitting condition (6) is satisfied, then the same power (35) is distributed across multiple sample signal subspace eigenvectors.

As can be observed in Fig. 2, the discrepancy between the estimated mean values for $\mathbf{e}_4^{\mathrm{H}}\hat{E}_S\hat{E}_S^{\mathrm{H}}\mathbf{e}_4$ and the prediction (35) at an source SNR experiencing significant MUSIC breakdown is within 0.2% point for a set of 10^3 Monte-Carlo trials.

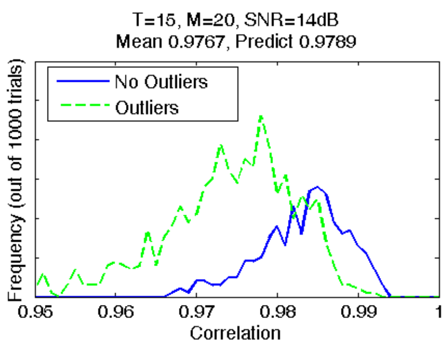

Fig. 2. Projection of 4th true eigenvector onto the sample signal subspace.

Finally, the most important observation from the MUSIC breakdown standpoint is that for both "proper" trials with no outliers and "improper" MUSIC trials with at least one outlier, the minimal signal subspace eigenvector still resides in the sample signal subspace with more than 95% of its power. This subspace "leakage" is accurately predicted by (35) for multiple SNR values, as indicated in Fig 2, and doesn't reach the 50% value associated with full subspace swap until a much lower SNR (comparable to that associated with MLE rather than MUSIC breakdown).

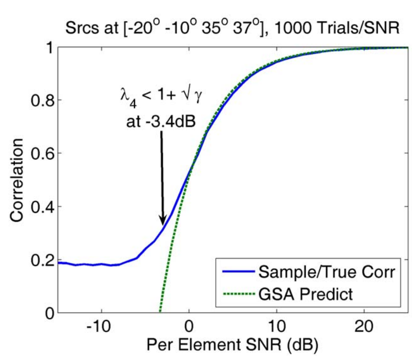

Fig. 3. Comparison of predicted and observed projection of the 4th sample eigenvector onto the true signal subspace. The correspondence between the observations and the predictions above $\lambda_4 < 1 + \sqrt{\gamma}$ is accurate even at small array sizes such as the $M = 20$ array.

3. SUMMARY AND CONCLUSION

In this paper we investigated the well-known DOA estimation performance breakdown phenomenon, which manifests as a rapid departure of estimation accuracy from the CRB due to the increasing probability of erroneous "outlier" estimates as the SNR or number of training samples is decreased below certain threshold values.

We analyzed this phenomenon for conventional MUSIC, the recently developed G-MUSIC [7], and MLE for a multiple Gaussian source scenario with i.i.d. sample support. Rather than traditional $T \to \infty$ asymptotic analysis, we used General Statistical Analysis and specifically focused on under-sampled scenarios with the number of training samples T less than the antenna dimension M.

The most controversial observation gained was that in the presence of closely spaced sources, MUSIC performance breakdown frequently takes place for SNR and sample support conditions that (according to GSA predictions which were shown to be accurate even for the under-sampled case) should almost surely preclude the "subspace swap" phenomenon. Therefore, for multiple source scenarios, MUSIC "performance breakdown" is associated with loss of resolution accompanied with a relatively insignificant inter-subspace "leakage", rather than full subspace swap.

4. REFERENCES

[1] D. Tufts, A. Kot, and R. Vaccaro, "The threshold effect in signal processing algorithms which use an estimated subspace," in *SVD and Signal Processing II: Algorithms, Analysis and Applications*, R. Vaccaro, Ed. New York: Elsevier, 1991, pp. 301–320.

[2] H. L. van Trees, *Detection, Estimation, and Modulation Theory: Part I*. New York: John Wiley and Sons, 1968.

[3] J. Thomas, L. Scharf, and D. Tufts, "The probability of a subspace swap in the SVD," *IEEE Trans. Signal Processing*, vol. 43, no. 3, pp. 730–736, Mar 1995.

[4] M. Hawkes, A. Nehorai, and P. Stoica, "Performance breakdown of subspace-based methods: Prediction and cure," in *Proc. ICASSP*, vol. 6. Salt Lake, UT, USA: IEEE, 2001, pp. 4005–4008.

[5] Y. I. Abramovich and N. K. Spencer, "Performance breakdown of subspace-based methods in arbitrary antenna arrays: GLRT-based prediction and cure," in *Proc. ICASSP*, vol. 2. Montreal, Canada: IEEE, 2004, pp. 117–120.

[6] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramér-Rao bound," *IEEE Trans. Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.

[7] X. Mestre, "An improved weighted MUSIC algorithm for small sample size scenarios," in *Proc. 14th Workshop on Adaptive Sensor Array Processing*. Lexington, MA, USA: MIT-LL, 2006. [Online]. Available: www.ll.mit.edu/asap/asap_06/pdf/Papers/19_mestre_pa.pdf

[8] ——, "Estimating the eigenvalues and associated subspaces of correlation matrices from a small number of observations," in *Proc. 2nd Int. Symp. Comm., Control, Signal Proc.*, Marrakech (Morocco), 2006.

[9] ——, "Improved estimation of eigenvalues of covariance matrices and their associated subspaces using their sample estimates," *Submitted to IEEE Transactions on Information Theory*, 2006. [Online]. Available: http://www.cttc.cat/drafts/cttc-rc-2006-004.pdf

[10] I. M. Johnstone, "On the distribution of the largest eigenvalue in principal component analysis," *Annals of Statistics*, vol. 29, no. 2, pp. 295–327, 2001.

[11] D. Paul, "Asymptotics of the leading sample eigenvalues for a spiked covariance model," Stanford University, Tech. Rep., Dec 2004. [Online]. Available: http://anson.ucdavis.edu/∼debashis/techrep/eigenlimit.pdf

[12] D. W. Tufts, A. C. Kot, and R. J. Vaccaro, "The threshold analysis of SVD-based algorithms," in *Proc. ICASSP*, New York, NY, USA, Apr 1988, pp. 2416–2419.