FUNDAMENTAL LIMIT OF SAMPLE EIGENVALUE BASED DETECTION OF SIGNALS IN COLORED NOISE USING RELATIVELY FEW SAMPLES

*Raj Rao Nadakuditi**

Massachusetts Institute of Technology
Department of Mathematics
Cambridge, MA 02139.
E-mail: raj@mit.edu

Jack W. Silverstein†

North Carolina State University
Department of Mathematics
Raleigh, NC 27695.
E-mail: jack@unity.ncsu.edu

ABSTRACT

Sample eigenvalue based estimators are often used for estimating the number of high-dimensional signals in colored noise when an independent estimate of the noise covariance matrix is available. We highlight a fundamental asymptotic limit of sample eigenvalue based detection that brings into sharp focus why in the large system, relatively large sample size limit, underestimation of the model order may be unavoidable for weak/closely spaced signals. We discuss the implication of these results for the detection of two weak, closely spaced signals.

1. INTRODUCTION

The observation vector, in many signal processing applications, can be modelled as a superposition of a finite number of signals embedded in additive noise. Detecting the number of signals present becomes a key issue and is often the starting point for the signal parameter estimation problem. When the signals and the noise are assumed, as we do in this paper, to be samples of a stationary, ergodic Gaussian vector process, the sample covariance matrix formed from m observations has the Wishart distribution [1]. We consider the class of estimators, inspired by the seminal work of Wax and Kailath [2], that determine the number of signals in colored noise from the generalized eigenvalues of the signal-plus-noise sample covariance matrix and the noise-only sample covariance matrix pair.

Our main objective is to explain precisely when and why, in high-dimensional, sample size limited settings underestimation of the model order is unavoidable. This is in contrast to works in the literature that use simulations, as in [3], to highlight the chronically reported symptom of sample (generalized) eigenvalue based estimators underestimating the number of signals without providing insight into whether a fundamental limit of detection is being encountered.

This paper addresses this issue using the powerful tools for analyzing large random matrices developed by Bai and Silverstein [4]. The main contribution of this paper is the introduction of the concept of *effective number of (identifiable) signals* which brings into sharp focus a fundamental limit in the identifiability, under sample size constraints, of closely spaced/low level signals using sample (generalized) eigenvalue based detection techniques. This concept explains why, in the large system relatively large sample size limit, *if the signal level is below a threshold* that depends on the noise covariance, sample size and the dimensionality of the system, *then reliable sample eigenvalue based detection is not possible*. The fundamental undetectability, due to insufficient samples, of weak/closely spaced signals using sample (generalized) eigenvalue based schemes due to insufficient samples is only exacerbated by adding more sensors since the detectability threshold is raised.

2. PROBLEM FORMULATION

We observe m samples ("snapshots") of possibly signal bearing n-dimensional snapshot vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$ where for each i, $\mathbf{x}_i \sim \mathcal{N}_n(0, \mathbf{R})$ and $\mathbf{x}_i$ are mutually independent. The snapshot vectors are modelled as

$$\mathbf{x}_i = \mathbf{A}\,\mathbf{s}_i + \mathbf{z}_i \qquad \text{for } i = 1, \ldots, m, \tag{1}$$

where $\mathbf{z}_i \sim \mathcal{N}_n(0, \Sigma)$, denotes an n-dimensional (real or complex) Gaussian noise vector where the noise covariance Σ may be known or unknown, $\mathbf{s}_i \sim \mathcal{N}_k(\mathbf{0}, \mathbf{R}_s)$ denotes a k-dimensional (real or complex) Gaussian signal vector with covariance $\mathbf{R}_s$, and $\mathbf{A}$ is a $n \times k$ unknown non-random matrix.

Since the signal and noise vectors are independent of each other, the covariance matrix of $\mathbf{x}_i$ can hence be decomposed as

$$\mathbf{R} = \mathbf{\Psi} + \mathbf{\Sigma} \tag{2}$$

*Work supported by ONR grant N00014-07-1-0269.
†Work supported by ARO grant W911NF-05-1-0244.

where

$$\mathbf{\Psi} = \mathbf{A}\mathbf{R}_s\mathbf{A}', \tag{3}$$

with $'$ denoting the conjugate transpose. Assuming that the matrix $\mathbf{A}$ is of full column rank, *i.e.*, the columns of $\mathbf{A}$ are linearly independent, and that the covariance matrix of the signals $\mathbf{R}_s$ is nonsingular, it follows that the rank of $\mathbf{\Psi}$ is k. Equivalently, the $n-k$ smallest eigenvalues of $\mathbf{\Psi}$ are equal to zero.

Consider the matrix,

$$\mathbf{R}_\Sigma := \mathbf{\Sigma}^{-1}\mathbf{R} = \mathbf{\Sigma}^{-1}\mathbf{\Psi} + \mathbf{I}, \tag{4}$$

whose eigenvalues we denote by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Then assuming that the rank of $\mathbf{\Sigma}^{-1}\mathbf{\Psi}$ is also k, it follows that the smallest $n-k$ eigenvalues of $\mathbf{R}_\Sigma$ are all equal to 1 so that

$$\lambda_{k+1} = \lambda_{k+2} = \ldots = \lambda_n = \lambda = 1, \tag{5}$$

while the remaining k eigenvalues $\mathbf{R}_\Sigma$ of will be strictly greater than one. Thus, if the true signal-plus-noise covariance matrix $\mathbf{R}$ and the noise-only covariance matrix $\mathbf{\Sigma}$ were known apriori, the number of signals k could be determined from the multiplicity of the smallest eigenvalue of $\mathbf{R}_\Sigma$. The problem in practice is that the signal-plus-noise and the noise covariance matrices $\mathbf{R}$ are unknown so that such a straight-forward algorithm cannot be used.

Instead, one estimates the signal-plus-covariance matrix using m snapshots as

$$\widehat{\mathbf{R}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{x}_i\mathbf{x}_i' \tag{6}$$

and the noise-only sample covariance matrix

$$\widehat{\mathbf{\Sigma}} = \frac{1}{N}\sum_{j=1}^{N}\mathbf{z}_j\mathbf{z}_j' \tag{7}$$

where $\mathbf{x}_i$ for $i = 1, \ldots, m$ are (possibly) signal-bearing snapshots and $\mathbf{z}_j$ for $j = 1, \ldots, N$ are noise-only snapshots. We then form the matrix

$$\widehat{\mathbf{R}}_{\widehat{\Sigma}} = \widehat{\mathbf{\Sigma}}^{-1}\widehat{\mathbf{R}} \tag{8}$$

and perform inference on the eigenvalues of $\widehat{\mathbf{R}}_{\widehat{\Sigma}}$. There are many techniques (e.g., [5, 6, 7] in the literature for inferring the number of signals from the eigenvalues of $\widehat{\mathbf{R}}_{\widehat{\Sigma}}$. There is no mathematically rigorous explanation in the signal processing literature of why, when $m = O(n)$, and $m = O(N)$ which is increasingly the case in many state-of-the-art sonar and radar array processing systems, underestimation of the model order is unavoidable. We fill this void in this paper using results from large random matrix theory [4].

3. ASYMPTOTIC IDENTIFIABILITY CRITERION AND THE EFFECTIVE NUMBER OF SIGNALS

A central object in the study of large random matrices is the empirical distribution function (e.d.f.) of the eigenvalues, which for an arbitrary matrix $\mathbf{A}$ with n real eigenvalues (counted with multiplicity), is defined as

$$F^{\mathbf{A}}(x) = \frac{\text{Number of eigenvalues of } \mathbf{A} \leq x}{n}. \tag{9}$$

For a broad class of random matrices, the sequence of e.d.f.'s can be shown to converge in the $n \to \infty$ limit to a non-random distribution function [8]. Of particular interest is the convergence of the e.d.f. of $\widehat{\mathbf{R}}_{\widehat{\Sigma}}$ in the signal-free case, which is described next.

Theorem 1. *Let $\widehat{\mathbf{R}}_{\widehat{\Sigma}}$ denote the matrix in (8) formed from m (complex Gaussian) noise-only snapshots and N independent noise-only (complex Gaussian) snapshots. Then the e.d.f. $F^{\widehat{\mathbf{R}}_{\widehat{\Sigma}}}(x) \to F^{R_\Sigma}(x)$ almost surely for every x, as $m, n(m) \to \infty$, $m, N(m) \to \infty$ and $c_m = n/m \to c > 0$ and $c_N^1 = n/N \to c_1 < 1$ where*

$$dF^W(x) = \max\left(0, \left(1 - \frac{1}{c}\right)\right)\delta(x) + \frac{\sqrt{(x-a_-)(a_+-x)}}{2\pi(c_1x^2+cx)}\mathbb{I}_{[a_-,a_+]}(x)\,dx, \tag{10}$$

where

$$a_\pm = \frac{-c_1\,c + c + 1 + c_1 \pm 2\sqrt{c + c_1 - c_1 c}}{{c_1}^2 + 1 - 2\,c_1}, \tag{11}$$

$\mathbb{I}_{[a,b]}(x) = 1$ when $a \leq x \leq b$ and zero otherwise, and $\delta(x)$ is the Dirac delta function.

Proof. This result was proved in [9]. When $c_1 \to 0$ we recover the famous Marčenko-Pastur density [10]. □

The following result exposes when the "signal" eigenvalues are asymptotically distinguishable from the "noise" eigenvalues.

Theorem 2. *Let $\widehat{\mathbf{R}}_{\widehat{\Sigma}}$ denote the matrix in (8) formed from m (complex Gaussian) signal-plus-noise snapshots and N independent (complex Gaussian) noise-only snapshots. Denote the eigenvalues of $\mathbf{R}_\Sigma$ by $\lambda_1 \geq \lambda_2 > \ldots \geq \lambda_k > \lambda_{k+1} = \ldots \lambda_n = 1$. Let l_j denote the j-th largest eigenvalue of $\widehat{\mathbf{R}}_{\widehat{\Sigma}}$. Then as $n, m(n) \to \infty$, $n, N(n) \to \infty$ and $c_m = n/m \to c > 0$ and $c_N^1 = n/N \to c_1 < 1$ we have*

$$l_j \to \begin{cases} \lambda_j \left(1 - c - c\dfrac{-c_1\lambda_j - \lambda_j + 1 + \sqrt{{c_1}^2{\lambda_j}^2 - 2\,c_1{\lambda_j}^2 - 2\,c_1\lambda_j + {\lambda_j}^2 - 2\lambda_j + 1}}{2c_1\lambda_j} \right), & \lambda_j > \tau(c, c_1) \\ \dfrac{-c_1\,c + c + 1 + c_1 + 2\sqrt{c + c_1 - c_1 c}}{{c_1}^2 + 1 - 2\,c_1}, & \lambda_j \le \tau(c, c_1) \end{cases}$$

for $j = 1, \ldots, k$ and the convergence is almost surely and the threshold $\tau(c, c_1)$ is given by

$$\tau(c, c_1) = \frac{\left({c_1}^2 + c_1\alpha_{c,c_1} - \alpha_{c,c_1} - 1\right) c - {c_1}^2 - 2c_1\alpha_{c,c_1} - c_1}{\left((c_1 - 1)\,c - c_1\right)(c_1 - 1)^2}. \tag{12}$$

and

$$\alpha_{c,c_1} := \sqrt{c + c_1 - c_1\,c} \tag{13}$$

Proof. This follows from applying the techniques in [11]. Note that when $c_1 \to 0$ we have

$$\tau(c, 0) = 1 + \sqrt{c}, \tag{14}$$

which recovers the results in [11, 12, 13]. □

This motivates our definition of the *effective number of identifiable signals* which is equal to

$$k_{eff}(\mathbf{R}, \mathbf{\Sigma}) = \#\,\text{Eigenvalues of } \mathbf{\Sigma}^{-1}\mathbf{R} > \tau\left(\frac{n}{m}, \frac{n}{N}\right) \tag{15}$$

The theorems above suggest that when the effective number of signals is less than the true number of signals then model order underestimation is (asymptotically) unavoidable. Figure 1 shows the eigen-SNR threshold $\tau(c, c_1)$ needed for reliable detection for different values as a function of c for different values of c_1. Such an analytical prediction was not possible before the results presented in this paper. Note the fundamental limit of detection in the situation when the noise-only covariance matrix is known apriori (solid line) and increase in the threshold eigen-SNR needed as the number of snapshots available to estimate the noise-only covariance matrix decreases.

4. ARRAY PROCESSING IMPLICATIONS

Suppose there are two uncorrelated (hence, independent) signals so that $\mathbf{R}_s = \text{diag}(\sigma_{S1}^2, \sigma_{S2}^2)$. In (1) let $\mathbf{A} = [\mathbf{v}_1 \mathbf{v}_2]$. In a sensor array processing application, we think of $\mathbf{v}_1 \equiv \mathbf{v}(\theta_1)$ and $\mathbf{v}_2 \equiv \mathbf{v}_2(\theta_2)$ as encoding the array manifold vectors for a source and an interferer with powers σ_{S1}^2 and σ_{S2}^2, located at θ_1 and θ_2, respectively. The signal-plus-noise covariance matrix is given by

$$\mathbf{R} = \sigma_{S1}^2\mathbf{v}_1\mathbf{v}_1' + \sigma_{S2}^2\mathbf{v}_2\mathbf{v}_2' + \mathbf{\Sigma} \tag{16}$$

where $\mathbf{\Sigma}$ is the noise-only covariance matrix. The matrix $\mathbf{R}_\Sigma$ defined in (4) can be decomposed as

$$\mathbf{R}_\Sigma = \mathbf{\Sigma}^{-1}\mathbf{R} = \sigma_{S1}^2\mathbf{\Sigma}^{-1}\mathbf{v}_1\mathbf{v}_1' + \mathbf{\Sigma}^{-1}\sigma_{S2}^2\mathbf{v}_2\mathbf{v}_2' + \mathbf{I}$$

so we that we can readily note that $\mathbf{R}_\Sigma$ has the $n - 2$ smallest eigenvalues $\lambda_3 = \ldots = \lambda_n = 1$ and the two largest eigenvalues respectively, where $\mathbf{u}_1 := \mathbf{\Sigma}^{-1/2}\mathbf{v}_1$ and $\mathbf{u}_2 := \mathbf{\Sigma}^{-1/2}\mathbf{v}_2$. Applying the result in Proposition 2 allows us to express the effective number of signals as

$$k_{\text{eff}} = \begin{cases} 2 & \text{if} \quad \tau\left(\frac{n}{m}, \frac{n}{N}\right) < \lambda_2 \\ 1 & \text{if} \quad \lambda_2 \le \tau\left(\frac{n}{m}, \frac{n}{N}\right) < \lambda_1 \\ 0 & \text{if} \quad \lambda_1 \le \tau\left(\frac{n}{m}, \frac{n}{N}\right) \end{cases} \tag{18}$$

Equation (18) captures the tradeoff between the identifiability of two closely spaced signals, the dimensionality of the system, the number of available snapshots and the cosine of the angle between the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. Note that since the effective number of signals depends on the structure of the noise covariance matrix (via the eigenvalues of $\mathbf{R}_\Sigma$), different assumed noise covariance structures (AR(1) versus white noise, for example) will impact the signal level SNR needed for reliable detection in different ways.

ACKNOWLEDGEMENTS

Raj Rao was supported by an Office of Naval Research Post-Doctoral Fellowship Award under grant N00014-07-1-0269. Jack Silverstein was supported by the U.S. Army Research Office under Grant W911NF-05-1-0244. R. R. thanks Arthur Baggeroer for encouragement and invaluable feedback. This material was based upon work supported by the National Science Foundation under Agreement No. DMS-0112069. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

$$\lambda_1 = 1 + \frac{(\sigma_{S1}^2 \| \mathbf{u}_1 \|^2 + \sigma_{S2}^2 \| \mathbf{u}_2 \|^2)}{2} + \frac{\sqrt{(\sigma_{S1}^2 \| \mathbf{u}_1 \|^2 - \sigma_{S2}^2 \| \mathbf{u}_2 \|^2)^2 + 4\sigma_{S1}^2 \sigma_{S2}^2 |\langle \mathbf{u}_1, \mathbf{u}_2 \rangle|^2}}{2} \tag{17a}$$

$$\lambda_2 = 1 + \frac{(\sigma_{S1}^2 \| \mathbf{u}_1 \|^2 + \sigma_{S2}^2 \| \mathbf{u}_2 \|^2)}{2} - \frac{\sqrt{(\sigma_{S1}^2 \| \mathbf{u}_1 \|^2 - \sigma_{S2}^2 \| \mathbf{u}_2 \|^2)^2 + 4\sigma_{S1}^2 \sigma_{S2}^2 |\langle \mathbf{u}_1, \mathbf{u}_2 \rangle|^2}}{2} \tag{17b}$$

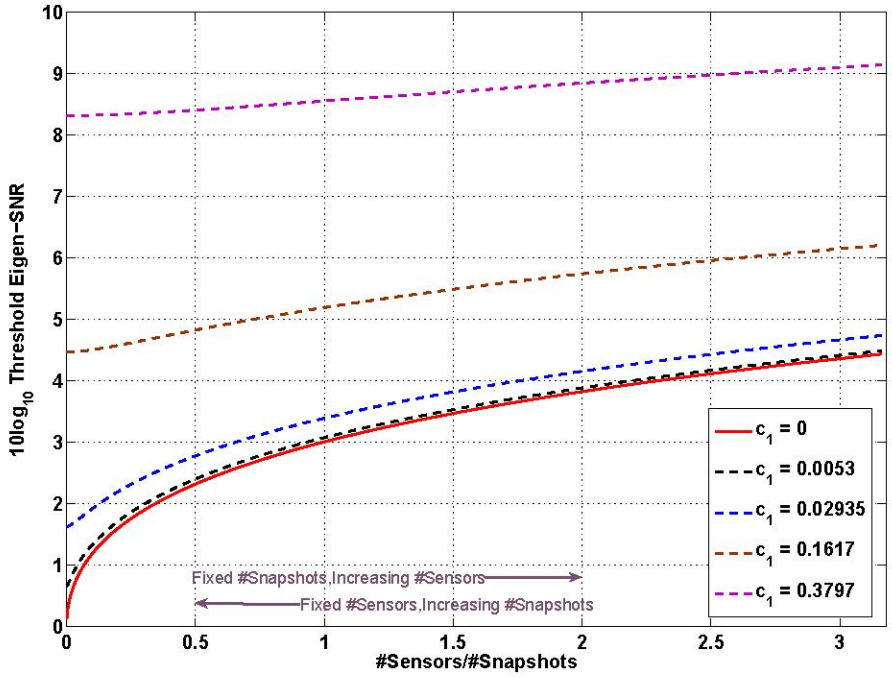

Fig. 1. Plot of the minimum (generalized) Eigen-SNR required (given by (12)) to be able to asymptotically discriminate between the "signal" and "noise" eigenvalue of the matrix $\hat{\mathbf{R}}_{\hat{\Sigma}}$ constructed as in (8) as a function of the ratio of the number of sensors to snapshots for different values of $c_1 \approx$Number of sensors/Number of noise-only snapshots.

5. REFERENCES

[1] J. Wishart. The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, 20 A:32–52, 1928.

[2] Mati Wax and Thomas Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. Acoust. Speech Signal Process.*, 33(2):387–392, 1985.

[3] A. P. Liavas and P. A. Regalia. On the behavior of information theoretic criteria for model order selection. *IEEE Trans. Signal Process.*, 49(8):1689–1695, August 2001.

[4] Zhidong Bai and Jack Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*. Science Press, Beijing, 2006.

[5] L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai. On detection of the number of signals when the noise covariance matrix is arbitrary. *J. Multivariate Anal.*, 20(1):26–49, 1986.

[6] Z. Zhu, S. Haykin, and X. Huang. Estimating the number of signals using reference noise samples. *IEEE Trans. on Aero. and Elec. Systems*, 27(3):575–579, May 1991.

[7] P. Stoica and M. Cedervall. Detection tests for array procesing in unknown correlated noise fields. *IEEE Trans. Signal Process.*, 45:2351–2362, September 1997.

[8] A. Edelman and N. R. Rao. Random matrix theory. In *Acta Numerica*, volume 14, pages 233–297. Cambridge University Press, 2005.

[9] J. W. Silverstein. The limiting eigenvalue distribution of a multivariate F matrix. *SIAM Journal on Math. Anal.*, 16(3):641–646, 1985.

[10] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.

[11] J. Baik and J. W. Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *Journal of Multivariate Analysis*, (6):1382–1408, 2006.

[12] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005.

[13] D. Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. Technical report, Stanford University, 2005. (To appear in 2007 in *Statistica Sinica*).