

Comments on "Multiple Antenna Spectrum Sensing in Cognitive Radios"

Erik Axell and Erik G. Larsson

Linköping University Post Print

N.B.: When citing this work, cite the original article.

©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Erik Axell and Erik G. Larsson, Comments on "Multiple Antenna Spectrum Sensing in Cognitive Radios", 2011, IEEE Transactions on Wireless Communications, (10), 5, 1678-1680.

<http://dx.doi.org/10.1109/TWC.2011.030911.101111>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-65761>

Comments on “Multiple Antenna Spectrum Sensing in Cognitive Radios”

Erik Axell, *Student Member, IEEE*, and Erik G. Larsson, *Senior Member, IEEE*

Abstract—We point out an error in a derivation in the recent paper [1], and provide a correct and much shorter calculation of the result in question. In passing, we also connect the results in [1] to the literature on array signal processing and on principal component analysis, and show that the signal detection methods proposed in [1] follow as special cases of standard results in these fields.

I. INTRODUCTION

The recent paper [1] deals with the following hypothesis test:

$$\begin{aligned} H_0 : \mathbf{y}_k &\sim N(\mathbf{0}, \sigma_n^2 \mathbf{I}) \quad \text{i.i.d, } k = 1, \dots, L \\ H_1 : \mathbf{y}_k &\sim N(\mathbf{0}, \sigma_s^2 \mathbf{h}\mathbf{h}^H + \sigma_n^2 \mathbf{I}) \quad \text{i.i.d, } k = 1, \dots, L, \end{aligned} \quad (1)$$

where the vectors $\mathbf{y}_k \in \mathbb{C}^{M \times 1}$ contain observed data, and $\mathbf{h} \in \mathbb{C}^{M \times 1}$ is a channel vector. The cited paper derives the generalized likelihood ratio test (GLRT) under the three different assumptions that

- (i) \mathbf{h} is unknown, but σ_s and σ_n are known;
- (ii) \mathbf{h} and σ_s are unknown, but σ_n is known; and
- (iii) all of \mathbf{h} , σ_s and σ_n are unknown.

In all three cases (i)–(iii) it is assumed that $\|\mathbf{h}\| = 1$. Without this assumption the decision problem is not well defined, owing to the ambiguity between a scaling of \mathbf{h} and σ_s . Replacing (σ_s, \mathbf{h}) with $(\sigma_s/c, c\mathbf{h})$ yields precisely the same hypothesis test, for any constant $c > 0$. This means in particular, that σ_s and $\|\mathbf{h}\|$ would not be identifiable from $\{\mathbf{y}_k\}$, not even if $L \rightarrow \infty$. It should be noted, however, that the constraint that $\|\mathbf{h}\| = 1$ does not impose any knowledge about the norm of the effective channel gain $\sigma_s \mathbf{h}$. The uncertainty about $\|\sigma_s \mathbf{h}\|$ is instead contained in the parameter σ_s . However, knowledge of σ_s as in case (i) effectively means knowledge of $\|\sigma_s \mathbf{h}\|$.

Note that similar problems as (1) were also dealt with recently in the context of spectrum sensing for cognitive radio in [2], [3], using a slightly different approach.

II. FLAW OF THE DERIVATION IN [1, SECTION III.A]

The calculation of the GLRT for case (i) is performed in Section III.A of [1]. It requires, among others, the maximization of Equation (7) in [1], reproduced here for completeness

This work was supported in part by the Swedish Research Council (VR), the Swedish Foundation for Strategic Research (SSF) and the ELLIIT. E. Larsson is a Royal Swedish Academy of Sciences (KVA) Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

of the exposition:

$$\begin{aligned} \hat{\mathbf{h}} = \operatorname{argmax}_{\mathbf{h}: \|\mathbf{h}\|=1} &\left(-\frac{\operatorname{tr}(\mathbf{Y}\mathbf{Y}^H)}{\sigma_n^2} + \frac{\|\mathbf{h}^H \mathbf{Y}\|^2}{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right) \sigma_n^2} \right. \\ &\left. - ML \log(\pi) - L \log\left(\frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{h}\|^2 + 1\right) - LM \log(\sigma_n^2) \right), \end{aligned} \quad (2)$$

where $\mathbf{Y} = [\mathbf{y}_1 \ \mathbf{y}_2 \ \dots \ \mathbf{y}_L] \in \mathbb{C}^{M \times L}$ is a matrix of observed signals. Section III.A of [1] finds $\hat{\mathbf{h}}$ by computing the gradient of the objective function in (2) and setting it equal to zero. While this calculation, coincidentally, gives the correct final result, it is flawed.

Fundamentally, one cannot treat the problem of maximizing a function subject to a set of constraints by maximizing the function without the constraints, and then applying the constraint afterwards. The point is that the optimization constraints need *not* be satisfied at the point(s) where the gradient of the objective function is zero, and there is no guarantee that the stationary points can be normalized to obtain a maximum of the constrained optimization problem. We provide a counterexample to illustrate this point, using the cost function in (2) for the special case when $\mathbf{h} = h$ and $\mathbf{Y} = y$ are complex-valued scalars and $\sigma_n^2 = \sigma_s^2 = 1$. Then, the problem simplifies to

$$\hat{h} = \operatorname{argmax}_{h: |h|^2=1} \left(\frac{|hy|^2}{1 + |h|^2} - \log(|h|^2 + 1) \right). \quad (3)$$

Taking the derivative of the objective function in (3) (as shown in the Appendix), and setting it equal to zero shows that the optimal h must satisfy either $h = 0$ or $|h|^2 = |y|^2 - 1$. None of these equations are compatible with the constraint $|h|^2 = 1$. The correct solution in this special case can be found by inspection:

$$\hat{h} = \operatorname{argmax}_{h: |h|^2=1} \left(\frac{|hy|^2}{(1 + 1)} - \log(1 + 1) \right) = \operatorname{argmax}_{h: |h|^2=1} |hy|^2,$$

which is satisfied by any h on the unit circle.

The point is, of course, that by setting the derivative of the objective function equal to zero, one finds stationary points. However, there is no guarantee that the optimal solution to the constrained problem can be attained by normalizing these stationary points. This is illustrated in Figure 1, which shows the objective function in the example (3) for some sample values of y . It is clear that there is a stationary point on the unit circle only if $|y|^2 - 1 = 1$. In all other cases, it is impossible to satisfy the constraint $|h|^2 = 1$ at any stationary point.

If attempting to find optima of the objective function in (2) by taking derivatives, then a proper way of handling the constraint $\|\mathbf{h}\|^2 = 1$ would be to include a Lagrange

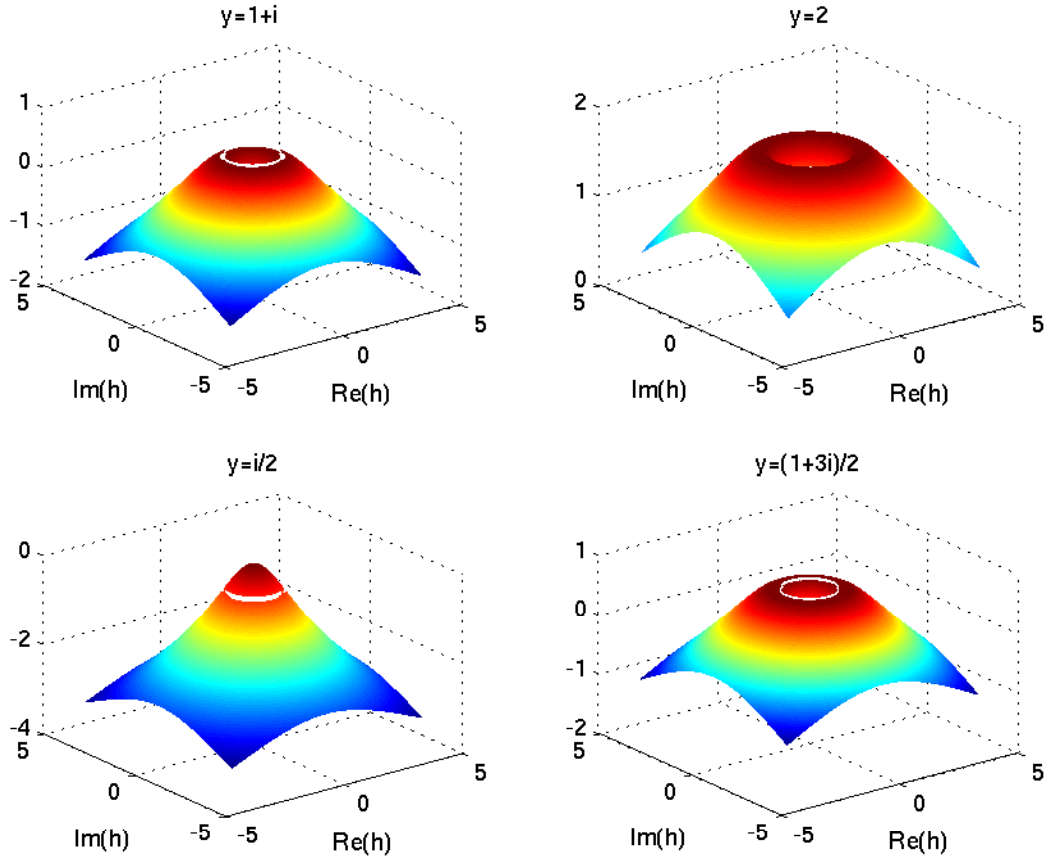


Fig. 1. Example of the objective function in (3) for $M = 1$, $L = 1$, $\sigma_n^2 = 1$, $\sigma_s^2 = 1$. The unit circle, $|h|^2 = 1$, is shown in white.

multiplier. However, as we show in Section III, there is no need to use a Lagrange multiplier or to take any derivatives—the solution to (2) can be found by inspection.

We end this section by pointing out that the expression for the gradient of the log-likelihood function given by [1, eq. (17)] is incorrect in that the constants α , β and ν given there are wrong. In the Appendix, we compute the correct values of α , β , ν (see Eq. (6)).

III. CALCULATION OF $\hat{\mathbf{h}}$ IN (2) FROM FIRST PRINCIPLES

The solution to the constrained maximization problem (2) is simple to find from first principles. The problem (2) is equivalent to

$$\begin{aligned}
 \hat{\mathbf{h}} &= \underset{\mathbf{h}: \|\mathbf{h}\|^2=1}{\operatorname{argmax}} \frac{\|\mathbf{h}^H \mathbf{Y}\|^2}{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right) \sigma_n^2} - L \log \left(\frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{h}\|^2 + 1 \right) \\
 &= \underset{\mathbf{h}: \|\mathbf{h}\|^2=1}{\operatorname{argmax}} \frac{\|\mathbf{h}^H \mathbf{Y}\|^2}{\left(\frac{\sigma_n^2}{\sigma_s^2} + 1\right) \sigma_n^2} - L \log \left(\frac{\sigma_s^2}{\sigma_n^2} + 1 \right) \\
 &= \underset{\mathbf{h}: \|\mathbf{h}\|^2=1}{\operatorname{argmax}} \mathbf{h}^H \mathbf{Y} \mathbf{Y}^H \mathbf{h}.
 \end{aligned} \tag{4}$$

The solution to (4), and thus to (2), is well known and is given by the eigenvector to $\mathbf{Y} \mathbf{Y}^H$ that corresponds to the largest eigenvalue, normalized to unit norm [4, pp. 176-177]. To see this using elementary algebra, let $\mathbf{Y} \mathbf{Y}^H = \mathbf{T} \mathbf{\Delta} \mathbf{T}^H$ be the eigenvalue-decomposition of $\mathbf{Y} \mathbf{Y}^H$, where \mathbf{T} is unitary

and $\mathbf{\Delta}$ is a diagonal matrix that contains the eigenvalues of $\mathbf{Y} \mathbf{Y}^H$, sorted in decreasing order. Then, the quadratic form $\mathbf{h}^H \mathbf{Y} \mathbf{Y}^H \mathbf{h} = \mathbf{h}^H \mathbf{T} \mathbf{\Delta} \mathbf{T}^H \mathbf{h}$ is maximized when $\mathbf{T}^H \mathbf{h} = [1, 0, \dots, 0]^T$ so that $\mathbf{h} = \mathbf{T}[1, 0, \dots, 0]^T$, which is by construction precisely the dominant eigenvector of $\mathbf{Y} \mathbf{Y}^H$. A similar derivation was presented in [3].

IV. RELATION TO THE ARRAY SIGNAL PROCESSING LITERATURE

Section III.A of [1] deals with the GLRT for the model (1) in case (i) and has been discussed above. Section III.B in [1] treats case (ii), and the results there follow in a similar manner.

Section III.C of [1] derives the GLRT for case (iii), when all parameters \mathbf{h} , σ_s , σ_n are unknown. This is perhaps the most interesting case in practice. We point out in passing that the GLRT for this case is well known in the literature on principal component analysis and in the array signal processing literature. In particular, it follows as a special case of the more general result showed in [5]. More precisely, [5] derived the maximum-likelihood estimate of an $n \times n$ covariance matrix of the “low-rank-plus-identity” form $\mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I}$, where \mathbf{G} is of dimension $n \times r$ and $r < n$. The result derived in Section III.C of [1] follows as a special case when $r = 1$. The results of [5], and the associated GLRT that results from inserting the ML covariance matrix estimate into the likelihood function, have been extensively used in the array signal processing literature for tasks such as detection of the number of signals impinging on a sensor array. See, for example, [6] for early work in this direction and [7] for a comprehensive summary of the field.

$$\begin{aligned}
& \frac{\partial}{\partial \mathbf{h}} \left(-\frac{\text{tr}(\mathbf{Y}\mathbf{Y}^H)}{\sigma_n^2} + \frac{\|\mathbf{h}^H \mathbf{Y}\|^2}{\left(\frac{\sigma_s^2}{\sigma_n^2} + \|\mathbf{h}\|^2\right) \sigma_n^2} - ML \log(\pi) - L \log\left(\frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{h}\|^2 + 1\right) - LM \log(\sigma_n^2) \right) \\
&= \frac{\partial}{\partial \mathbf{h}} \left(\frac{\|\mathbf{h}^H \mathbf{Y}\|^2}{\left(\frac{\sigma_s^2}{\sigma_n^2} + \|\mathbf{h}\|^2\right) \sigma_n^2} - L \log\left(\frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{h}\|^2 + 1\right) \right) \\
&= \frac{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right) \sigma_n^2 \frac{\partial}{\partial \mathbf{h}} \|\mathbf{h}^H \mathbf{Y}\|^2 - \|\mathbf{h}^H \mathbf{Y}\|^2 \frac{\partial}{\partial \mathbf{h}} \left(\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right) \sigma_n^2\right)}{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right)^2 \sigma_n^4} - \frac{L}{\frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{h}\|^2 + 1} \frac{\partial}{\partial \mathbf{h}} \left(\frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{h}\|^2 + 1\right) \quad (5) \\
&= \frac{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right) \sigma_n^2 2\mathbf{Y}\mathbf{Y}^H \mathbf{h} - \|\mathbf{h}^H \mathbf{Y}\|^2 2\sigma_n^2 \mathbf{h}}{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right)^2 \sigma_n^4} - \frac{L}{\frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{h}\|^2 + 1} 2\frac{\sigma_s^2}{\sigma_n^2} \mathbf{h} \\
&= \frac{2}{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right) \sigma_n^2} \mathbf{Y}\mathbf{Y}^H \mathbf{h} - \frac{2}{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right)^2 \sigma_n^2} \mathbf{h}\mathbf{h}^H \mathbf{Y}\mathbf{Y}^H \mathbf{h} - \frac{2\frac{\sigma_s^2}{\sigma_n^2} L}{\frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{h}\|^2 + 1} \mathbf{h}.
\end{aligned}$$

APPENDIX

Let $\mathbf{h} = \mathbf{a} + j\mathbf{b}$ be a complex vector, where $\mathbf{a} = [a_1, a_2, \dots, a_n]^T$ and $\mathbf{b} = [b_1, b_2, \dots, b_n]^T$ are its real and imaginary parts, respectively. The gradient of a real-valued function $f(\mathbf{h})$ w.r.t. the complex-valued vector \mathbf{h} is defined as [8, p. 798]

$$\frac{\partial f(\mathbf{h})}{\partial \mathbf{h}} \triangleq \frac{\partial f(\mathbf{h})}{\partial \mathbf{a}} + j \frac{\partial f(\mathbf{h})}{\partial \mathbf{b}} \triangleq \begin{bmatrix} \frac{\partial f(\mathbf{h})}{\partial a_1} \\ \frac{\partial f(\mathbf{h})}{\partial a_2} \\ \vdots \\ \frac{\partial f(\mathbf{h})}{\partial a_n} \end{bmatrix} + j \begin{bmatrix} \frac{\partial f(\mathbf{h})}{\partial b_1} \\ \frac{\partial f(\mathbf{h})}{\partial b_2} \\ \vdots \\ \frac{\partial f(\mathbf{h})}{\partial b_n} \end{bmatrix}.$$

Then, the gradient of the log-likelihood function is written in (5), where in the third equality we used that $\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^H \mathbf{A} \mathbf{x} = 2\mathbf{A} \mathbf{x}$ for any Hermitian matrix \mathbf{A} [8, pp. 796-798]. Setting the expression (5) equal to the zero vector is equivalent to solving

$$(\alpha \mathbf{I} + \beta \mathbf{h}\mathbf{h}^H) \mathbf{Y}\mathbf{Y}^H \mathbf{h} = \nu \mathbf{h},$$

$$\begin{aligned}
\text{where, } \alpha &= \frac{1}{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right) \sigma_n^2}, \\
\beta &= -\frac{1}{\left(\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2\right)^2 \sigma_n^2} = -\frac{\alpha}{\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2}, \quad (6) \\
\nu &= \frac{L}{\frac{\sigma_n^2}{\sigma_s^2} + \|\mathbf{h}\|^2} = L\sigma_n^2 \alpha.
\end{aligned}$$

REFERENCES

- [1] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 814–823, Feb. 2010.
- [2] P. Bianchi, M. Debbah, M. Maida and J. Najim, "Performance of statistical tests for single source detection using random matrix theory," *IEEE Transactions on Information Theory*, to appear.
- [3] P. Wang, J. Fang, N. Han and H. Li, "Multi-antenna assisted spectrum sensing for cognitive radio," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1791–1800, May 2010.
- [4] R.A. Horn and C.R. Johnson, "Matrix Analysis," *Cambridge University Press*, 1990.
- [5] T.W. Anderson, "Asymptotic theory for principal component analysis," *The Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 122–148, Mar. 1963.
- [6] M. Wax, and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, Apr. 1985.
- [7] H. van Trees, "Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing", *Wiley*, 2001.
- [8] S. Haykin, "Adaptive Filter Theory (4th Edition)," *Prentice Hall*, September 2001.