A BAYESIAN DERIVATION OF THE GENERALIZED COHERENCE DETECTOR

S. Sirianunpiboon[1], *S. D. Howard*[1] *and D. Cochran*[2]

[1]Defence Science and Technology Organisation
PO Box 1500, Edinburgh 5111, Australia.
[2]School of Mathematical and Statistical Sciences
Arizona State University, Tempe AZ 85287-5706 USA.

ABSTRACT

The generalized coherence (GC) estimate is a well studied statistic for detection of a common but unknown signal on several noisy channels. It was developed as a generalization of the magnitude-squared coherence estimate, a popular detection statistic in the two-channel setting, specifically to address applications in which an unknown signal source is to be detected and localized from measurements made at several sensors. In this paper, it is shown that the GC detector arises naturally from a Bayesian perspective. Specifically, it is derived as a test of the hypothesis that the signals in the channels are independent Gaussian processes against the hypothesis that the processes have some arbitrary correlation. This is achieved by introducing suitable non-informative priors for the covariance matrices across the channels under the two hypotheses. Subsequently, reduced likelihoods are obtained by marginalizing the joint distribution of the data and the covariance matrix in each case. The likelihood ratio is then shown to be a monotonic function of the GC detection statistic.

Index Terms— generalized coherence detector, reduced likelihood, non-informative priors

1. INTRODUCTION

The problem of detecting a common but unknown signal on two or more noisy channels has a rich history. Pioneering work in the context of sonar signal processing led to widespread use of the magnitude-squared coherence (MSC) estimate as a detection statistic in the two-channel setting [1]. The properties of the MSC estimate and the performance of detectors based upon it were studied extensively in the 1970s and 1980s [2, 3]. Extending the MSC concept to multiple-channel scenarios received considerable attention in the 1980s, leading to study of estimators of multiple coherence [4] and, somewhat later, generalized coherence (GC) as multi-channel detection statistics. The GC estimate was introduced in [5] in the context of multiple-channel detection and the performance of the GC estimator as a detection statistic was documented in several works throughout the 1990s and into the early 2000s. Its symmetries and invariance properties were well studied during this period [6, 7], and the extent to which the GC estimate is canonical with respect to a desirable set of invariances and symmetries was examined as well [8].

Despite its merits in principle, GC-based detection has remained an academic novelty for most of the two decades since its introduction. Deep water sonar applications, which originally provided much of the impetus for research in multiple-channel detection, became less important with the changing nature of antisubmarine warfare. And even in deep-water scenarios, the low SNRs available by the time multi-channel methods were becoming theoretically mature made simultaneous passive detection on more than two channels highly unlikely. Multi-channel methods have remained of modest interest in sonar and in a few other application regimes, such as fault detection and localization in machinery. But this level of application interest has not been sufficient to fuel the amount of activity once enjoyed by this research area, though good theoretical work continues to emerge at a slow pace (e.g., [9] extends the GC idea to the case of time-correlated signals and derives the GLRT in the time and frequency domains).

Today, distributed networks of sensors of various (and mixed) modalities are being examined in connection with a wide variety of applications. In some envisioned uses, the sensors would be closely spaced relative to the propagation distances of the signals they seek to detect and localize, thus suggesting a likely re-emergence of the importance of multi-channel passive methods.

In this paper we take a fresh look at the GC detector from a Bayesian perspective. Specifically, we show that it arises naturally as the statistic for a test of the hypothesis that the signals in the channels are independent Gaussian processes against the hypothesis that the processes have some arbitrary correlation. This is achieved by introducing suitable non-informative priors for the covariance matrices across the channels under the two hypotheses. Subsequently, reduced likelihood are obtained by marginalizing the joint distribution of the data and the covariance matrix in each case. The likelihood ratio is then shown to be equivalent to the GC detector statistic. This perspective directly connects the GC statistic to the Bayesian (as opposed the Neyman-Pearson) detector for a precisely defined multi-channel problem.

Notation: In this paper we use bold lowercase symbols to represent vectors of complex value, whereas bold uppercase symbols for complex matrices. † denotes Hermitian transpose, $|\cdot|$ denotes the determinant, tr denotes the trace operator and $\mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ represents diagonal matrix of size $p \times p$ with $\lambda_1, \ldots, \lambda_p$ along the diagonal. The two distributions mainly used in this paper are :

1. If $\boldsymbol{x}$ is complex Gaussian of dimension $p \times 1$, i.e., if $\boldsymbol{x} \sim \mathcal{CN}_p(\boldsymbol{\mu}, \boldsymbol{R})$, its density is defined as

$$\mathcal{CN}_p(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{R}) = \frac{1}{\pi^p |\boldsymbol{R}|} e^{-\mathrm{tr}\{\boldsymbol{R}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^\dagger\}}. \tag{1}$$

2. If $\boldsymbol{R}$ is complex Wishart of dimension $p \times p$, i.e., if $\boldsymbol{R} \sim \mathcal{CW}_p(\Sigma, n)$, its density is defined by [10]

$$\mathcal{CW}_p(\boldsymbol{R}|\Sigma, n) = \frac{1}{I(\Sigma)} |\boldsymbol{R}|^{n-p} e^{-\mathrm{tr}\{\Sigma^{-1}\boldsymbol{R}\}}, \tag{2}$$

where $I(\Sigma) = \pi^{p(p-1)/2} \prod_{i=1}^{p} \Gamma(n-p+i) |\Sigma|^n$.

2. PROBLEM FORMULATION

Consider M complex random processes representing the outputs of M spatially distributed sensors. For $m = 1, \ldots, M$, let $\boldsymbol{x}_m = (x_{m,1}, \ldots, x_{m,N})$ be a vector of N complex measurements obtained from sensor m and $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M)^{\mathrm{T}}$. The generalized coherence (GC) estimate [5, 8] is defined as

$$\gamma^2 = 1 - \frac{|\boldsymbol{X}\boldsymbol{X}^\dagger|}{\|\boldsymbol{x}_1\|^2 \ldots \|\boldsymbol{x}_M\|^2} \tag{3}$$

where $|\boldsymbol{X}\boldsymbol{X}^\dagger|$ is the determinant of an $M \times M$ *Gram matrix*.

The goal here is to derive the GC detector from the Bayesian perspective. The detector is able to determine whether a common signal is present on two or more noisy channels, that is it can distinguish between hypotheses H_0, in which the vector $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M$ contain independent samples of independent zero-mean Gaussian processes with unknown covariance and its complement H_1 in which the vector $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M$ contains unknown and correlated signals.

$$H_1 : \boldsymbol{X} \sim \mathcal{CN}(0, \boldsymbol{R}_1), \text{ for some covariance } \boldsymbol{R}_1 \tag{4}$$

$$H_0 : \boldsymbol{X} \sim \mathcal{CN}(0, \boldsymbol{R}_0), \text{ for some diagonal covariance } \boldsymbol{R}_0 \tag{5}$$

Let $\boldsymbol{x}(n) = (x_{1,n}, x_{2,n}, \ldots, x_{M,n})^{\mathrm{T}}$ be a vector of sample at time n from M sensors. The likelihood function of H_0 is given by

$$\begin{aligned} p(\boldsymbol{X}|H_0, \boldsymbol{R}_0^{-1}) &= \prod_{n=1}^{N} p(\boldsymbol{x}(n)|H_0, \boldsymbol{R}_0) \\ &= \pi^{-MN} |\boldsymbol{R}_0^{-1}|^N e^{-\mathrm{tr}\{\boldsymbol{R}_0^{-1}\boldsymbol{W}\}} \end{aligned} \tag{6}$$

and similarly, the likelihood function of H_1

$$p(\boldsymbol{X}|H_1, \boldsymbol{R}_1^{-1}) = \pi^{-MN} |\boldsymbol{R}_1^{-1}|^N e^{-\mathrm{tr}\{\boldsymbol{R}_1^{-1}\boldsymbol{W}\}} \tag{7}$$

where $\boldsymbol{W} = \sum_{n=1}^{N} \boldsymbol{x}(n)\boldsymbol{x}(n)^\dagger = \boldsymbol{X}\boldsymbol{X}^\dagger$. Note that $\boldsymbol{X}\boldsymbol{X}^\dagger$ is a *Gram matrix* and its determinant is known as a *Gram determinant*.

The parameters $\boldsymbol{R}_0$ and $\boldsymbol{R}_1$ are nuisance parameters in the likelihood which we want to eliminate. Since the likelihood is expressed in term of $\boldsymbol{R}_0^{-1}$ and $\boldsymbol{R}_1^{-1}$, we regard these to be our nuisance parameters instead. The Bayesian treatment of the nuisance parameters $\boldsymbol{R}_0^{-1}$ and $\boldsymbol{R}_1^{-1}$ is to marginalise the posterior distribution with respect to nuisance parameters, i.e., consider the integrated likelihood

$$p(\boldsymbol{X}|H_0) = \int p(\boldsymbol{X}|H_0, \boldsymbol{R}_0^{-1}) p(\boldsymbol{R}_0^{-1}) d\boldsymbol{R}_0^{-1} \tag{8}$$

$$p(\boldsymbol{X}|H_1) = \int p(\boldsymbol{X}|H_1, \boldsymbol{R}_1^{-1}) p(\boldsymbol{R}_1^{-1}) d\boldsymbol{R}_1^{-1}. \tag{9}$$

Here

$$dR_0^{-1} = \prod_{i=1}^{M} dr^{ii} \tag{10}$$

$$dR_1^{-1} = \prod_{i=1}^{M} dr^{ii} \prod_{i<j} d\Re(r^{ij}) d\Im(r^{ij}) \tag{11}$$

where r^{ij} denotes the $(i,j)^{\text{th}}$ element of R^{-1}.

Through this procedure we have replaced the data models containing the unknown parameters $\boldsymbol{R}_0^{-1}$ and $\boldsymbol{R}_1^{-1}$, with the models $p(\boldsymbol{X}|H_0)$ and $p(\boldsymbol{X}|H_1)$ so that the detection problem compares

$$H_1 : \boldsymbol{X} \sim p(\,\cdot\,|H_1), \tag{12}$$

$$H_0 : \boldsymbol{X} \sim p(\,\cdot\,|H_0). \tag{13}$$

We show here that the GC detection statistic is the monotonic function of the likelihood ratio,

$$\frac{p(\boldsymbol{X}|H_1)}{p(\boldsymbol{X}|H_0)}. \tag{14}$$

3. PRIOR PROBABILITY DISTRIBUTION

In order to carry out the integrations in (8) and (9), we need a prior distribution on $\boldsymbol{R}_0^{-1}$ and $\boldsymbol{R}_1^{-1}$. We would like these priors to be as non-informative as possible while still leading to proper densities $p(\boldsymbol{X}|H_1)$ and $p(\boldsymbol{X}|H_0)$.

3.1. The Jeffreys Priors

Consider the parameter manifold $\mathcal{M}_1$ of positive definite matrices $\boldsymbol{R}_1$. Since every positive definite matrices has a unique inverse we can equally well parametrize this manifold in terms of inverse covariance matrices $\boldsymbol{R}_1^{-1}$. The manifold $\mathcal{M}_1$ is endowed with a Riemannian metric by the Fisher information associated with estimation of $\boldsymbol{R}_1^{-1}$ from data, based on the model H_1. That is,

$$F = \mathrm{E}_{\boldsymbol{X}} \left\{ d \log p(\boldsymbol{X}|R_1^{-1}) \otimes d \log p(\boldsymbol{X}|R_1^{-1}) \right\} \tag{15}$$

where d denotes the exterior derivative on $\mathcal{M}_1$ [11]. The Jeffreys prior on $\mathcal{M}_1$ is the volume form associated with the Fisher metric,

$$\mathrm{vol} = \sqrt{|F|} dR_1^{-1}. \tag{16}$$

This commonly used as a canonical non-informative prior, as it is invariant with respect to re-parametrization [12, 13]. For the complex data model, Jeffreys' prior for $\boldsymbol{R}_1^{-1}$ is given by

$$p_J(\boldsymbol{R}_1^{-1})\, dR_1^{-1} = |\boldsymbol{R}_1^{-1}|^{-M} dR_1^{-1}. \tag{17}$$

As a probability distribution this is improper. Often this is not a problem as long as the posterior distribution can be normalised. In our case, the marginalised data model distribution is

$$\begin{aligned} p(\boldsymbol{X}|H_1) &\propto \frac{1}{\pi^{MN}} \int |\boldsymbol{R}_1^{-1}|^{N-M} e^{-\mathrm{tr}\{\boldsymbol{R}_1^{-1}\boldsymbol{X}\boldsymbol{X}^\dagger\}} d\boldsymbol{R}_1^{-1} \\ &\propto \frac{\pi^{M(M-1)/2}}{\pi^{MN}} \prod_{i=1}^{M} \Gamma(N{-}M{+}i) |\boldsymbol{X}\boldsymbol{X}^\dagger|^N \end{aligned} \tag{18}$$

which is still improper, so we need to be more careful in specifying a prior distribution.

The positive definite real diagonal matrices form a submanifold $\mathcal{M}_0$ of $\mathcal{M}_1$. In this case the Jeffreys prior for $\boldsymbol{R}_0^{-1}$,

$$p_J(\boldsymbol{R}_0^{-1}) dR_0^{-1} \propto |\boldsymbol{R}_0^{-1}|^{-1} dR_0^{-1} \tag{19}$$

and the marginalised data model distributions,

$$p(\boldsymbol{X}|H_0) \propto \frac{1}{\pi^{MN}} \Gamma^M(N) \left(\prod_{i=1}^{M} \|\boldsymbol{x}_i\|^2 \right)^{-N}. \tag{20}$$

Again this is improper. Choosing the appropriate prior is crucial in Bayesian perspective, we consider a more careful choice of prior distribution in the following noting that the Jeffreys priors will be the basis of the choice.

3.2. Maximum Entropy Priors

In the previous section we found that the use of the improper Jeffreys prior lead to improper marginalised data model distributions. This implies that we must supply more informative prior distributions if we are to make progress. Our approach is to introduce enough prior information to give proper marginalised data model densities $p(\boldsymbol{X}|H_1)$ and $p(\boldsymbol{X}\ H_0)$ and then relax this as much as possible in the final result in a manner we now describe.

The maximum entropy distribution, subject to certain constraints, represents the least informative distribution [14] satisfying those constraints. Our approach is to construct proper non-informative prior distributions but constructing maximum entropy distributions relative to Jeffreys' prior under the two hypotheses subject to the following conditions

$$\mathrm{E}(\mathrm{tr}(\boldsymbol{R}^{-1})) = \alpha_1 \tag{21}$$

$$\mathrm{E}(\log(|\boldsymbol{R}|)) = \alpha_2 \tag{22}$$

First consider the case that $\boldsymbol{R}_1 \in \mathcal{M}_1$. The maximum entropy distribution of $\boldsymbol{R}_1^{-1}$ relative to the Jeffreys prior, subject to (21) and (22) is given by

$$\begin{aligned} p(\boldsymbol{R}_1^{-1}) &= Z(\tau, q)^{-1} |\boldsymbol{R}_1^{-1}|^{-M} e^{-\tau \mathrm{tr}(\boldsymbol{R}_1^{-1}) - q \log(|\boldsymbol{R}_1|)} \\ &= Z(\tau, q)^{-1} |\boldsymbol{R}_1^{-1}|^{q-M} e^{-\mathrm{tr}(\tau \boldsymbol{R}_1^{-1})} \end{aligned} \tag{23}$$

where τ and q are chosen to satisfy the conditions (21) and (22) and $Z(\tau, q)$ is the normalisation constant or partition function,

$$\begin{aligned} Z(\tau, q) &= \int_{\mathcal{M}_1} |\boldsymbol{R}_1^{-1}|^{q-M} e^{-\mathrm{tr}(\tau \boldsymbol{R}_1^{-1})} d\boldsymbol{R}_1^{-1} \\ &= \pi^{M(M-1)/2} \tau^{-Mq} \prod_{i=1}^{M} \Gamma(q - M + i). \end{aligned} \tag{24}$$

The prior (23) is proper for $\tau > 0$ and $q > M - 1$. Notice that in the limit $\tau, q \to 0$ the prior (23) becomes less informative and approaches the Jeffreys prior up to normalisation.

The moments (21) and (22) can be computed using the partition function $Z(\tau, q)$:

$$\begin{aligned} \mathrm{E}(\mathrm{tr}(\boldsymbol{R}_1^{-1})) &= -\frac{\partial}{\partial \tau} \log Z(\tau, q) \\ &= Mq\tau^{-1} \end{aligned} \tag{25}$$

and

$$\begin{aligned} \mathrm{E}(\log|\boldsymbol{R}_1|) &= -\frac{\partial}{\partial q} \log Z(\tau, q) \\ &= M \log \tau - \sum_{i=1}^{M} \psi(q - M + i) \end{aligned} \tag{26}$$

where ψ is the digamma function defined by

$$\psi(x) = \frac{d}{dx} \log \Gamma(x). \tag{27}$$

Now consider the prior under H_0. In this case $\boldsymbol{R}_0 \in \mathcal{M}_0$, and the maximum entropy distribution relative to Jeffreys' prior subject to the constraints (21) and (22) is

$$p(\boldsymbol{R}_0^{-1}) = Z(\tau, q)^{-1} |\boldsymbol{R}_0^{-1}|^{q-1} e^{-\mathrm{tr}(\tau \boldsymbol{R}_0^{-1})}. \tag{28}$$

In this case the partition function is

$$\begin{aligned} Z(\tau, q) &= \int_{\mathcal{M}_0} |\boldsymbol{R}_0^{-1}|^{q-1} e^{-\mathrm{tr}(\tau \boldsymbol{R}_0^{-1})} d\boldsymbol{R}_0^{-1} \\ &= \tau^{-Mq} \Gamma^M(q) \end{aligned} \tag{29}$$

implying that the prior (28) is proper for $\tau > 0$ and $q > 0$. We have the moments

$$\begin{aligned} \mathrm{E}(\mathrm{tr}(\boldsymbol{R}_0^{-1})) &= -\frac{\partial}{\partial \tau} \log Z(\tau, q) \\ &= Mq\tau^{-1} \end{aligned} \tag{30}$$

and

$$\begin{aligned} \mathrm{E}(\log|\boldsymbol{R}_0|) &= -\frac{\partial}{\partial q} \log Z(\tau, q) \\ &= M \log \tau - M\psi(q). \end{aligned} \tag{31}$$

We first note that (30) and (25) are consistent. Secondly, as $\tau \to 0$ we have

$$\begin{aligned} \frac{\mathrm{E}(\log|\boldsymbol{R}_1|) - \mathrm{E}(\log|\boldsymbol{R}_0|)}{\mathrm{E}(\log|\boldsymbol{R}_0|)} &\sim \frac{M\psi(q) - \sum_{i=1}^{M} \psi(q - M + i)}{M \log \tau} \\ &= \frac{\sum_{i=1}^{M-1} \frac{i}{q-M+i}}{M \log \tau} \end{aligned} \tag{32}$$

using the identity $\psi(x + 1) = \psi(x) + 1/x$. For $q = M$ this becomes

$$\frac{\mathrm{E}(\log|\boldsymbol{R}_1|) - \mathrm{E}(\log|\boldsymbol{R}_0|)}{\mathrm{E}(\log|\boldsymbol{R}_0|)} \sim \frac{M-1}{M \log \tau}. \tag{33}$$

Thus, the prior probability distributions for both hypotheses are consistent with respect to the constraints (21) and (22) in the limit $\tau \to 0$.

4. DATA MODEL

In this section, we calculate the posterior distributions for the above priors on $\boldsymbol{R}_0^{-1}$ and $\boldsymbol{R}_1^{-1}$, discussed in the previous section. The reduced data models are obtained by marginalizing with respect to these parameters.

Substituting (6) and (28) into (8), we obtain the reduced likelihood for hypothesis H_0,

$$\begin{aligned} &p(\boldsymbol{X}|H_0) \\ &= \frac{\tau^{Mq}\pi^{-MN}}{\Gamma^M(q)} \int_{\mathcal{M}_0} |\boldsymbol{R}_0^{-1}|^{N+q-1} e^{-\mathrm{tr}(\boldsymbol{R}_0^{-1}(\boldsymbol{W}+\tau I_M))} d\boldsymbol{R}_0^{-1} \\ &= \frac{\tau^{Mq}\pi^{-MN}}{\Gamma^M(q)} \prod_{i=1}^{M} \int_0^\infty \sigma_i^{-2(N+q)-1} e^{-\sigma_i^{-2}(\|\boldsymbol{x}_i\|^2+\tau)} d\sigma_i \\ &= \frac{\tau^{Mq}\Gamma^M(N+q)}{\pi^{MN}\Gamma^M(q) \prod_{i=1}^{M} (\|\boldsymbol{x}_i\|^2 + \tau)^{N+q}} \end{aligned} \tag{34}$$

For hypothesis H_1, substituting (7) and (23) into (9),

$$\begin{aligned} &p(\boldsymbol{X}|H_1) \\ &= \frac{\tau^{Mq} \int_{\mathcal{M}_1} |\boldsymbol{R}_1^{-1}|^{N+q-M} e^{-\mathrm{tr}\{\boldsymbol{R}_1^{-1}(\boldsymbol{W}+\tau I_M)\}} d\boldsymbol{R}_1^{-1}}{\pi^{MN}\pi^{M(M-1)/2} \prod_{i=1}^{M} \Gamma(q-M+i)} \\ &= \frac{\tau^{Mq} \prod_{i=1}^{M} \Gamma(N+q-M+i)}{\pi^{MN} \prod_{i=1}^{M} \Gamma(q-M+i) |\boldsymbol{W} + \tau I_M|^{N+q}}. \end{aligned}$$

5. GENERALIZED COHERENCE ESTIMATE AS LIKELIHOOD RATIO

In this section we show that the GC detection statistic is obtained from the reduced likelihood ratio of the two hypotheses. The likelihood ratio statistic for the hypothesis H_0 against H_1 is given by

$$\frac{p(\boldsymbol{X}|H_1)}{p(\boldsymbol{X}|H_0)} = \frac{K(N+q, M)}{K(q, M)} \left(\frac{\prod_{i=1}^{M} (\|\boldsymbol{x}_i\|^2 + \tau)}{|\boldsymbol{W} + \tau I_M|} \right)^{N+q} \tag{35}$$

where

$$K(x, M) = \prod_{i=1}^{M} \frac{\Gamma(x - M + i)}{\Gamma(x)} \tag{36}$$

As $\tau \to 0$, the likelihood ratio becomes

$$\frac{p(\boldsymbol{X}|H_1)}{p(\boldsymbol{X}|H_0)} = \frac{K(N+q, M)}{K(q, M)} \left(\frac{\prod_{i=1}^{M} \|\boldsymbol{x}_i\|^2}{|\boldsymbol{X}\boldsymbol{X}^\dagger|} \right)^{N+q} \tag{37}$$

The GC estimate γ^2 given in (3) [5, 8] is related to the likelihood ration (37) by

$$\frac{p(\boldsymbol{X}|H_1)}{p(\boldsymbol{X}|H_0)} = \frac{K(N+q, M)}{K(q, M)} \frac{1}{(1-\gamma^2)^{N+q}} \tag{38}$$

and so a monotonic function of the likelihood ratio. In the case $M = 2$, we have

$$\frac{p(\boldsymbol{X}|H_1)}{p(\boldsymbol{X}|H_0)} = \frac{q-1}{N+q-1} \left(\frac{|\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle|^2}{\|\boldsymbol{x}_1\|^2 \|\boldsymbol{x}_2\|^2} \right)^{-(N+q)} \tag{39}$$

6. DETECTION THRESHOLDS

The Neyman-Pearson detection thresholds for the generalized coherence detector have been explored in detail in [8]. Here we consider the Bayesian detection thresholds.

Let C_{ij} for $i, j = 0, 1$ denote the loss function for our decision problem. That is, C_{ij} is the loss incurred in deciding H_i when H_j is true. We take $C_{00} = C_{11} = 0$. The expected loss in deciding H_1 is true is given by

$$\rho(H_1) = C_{10} p(H_0|\boldsymbol{X}) \tag{40}$$

and for deciding H_0 is true is

$$\rho(H_0) = C_{01} p(H_1|\boldsymbol{X}) \tag{41}$$

The Bayesian decision minimises the expected loss, that is, we decide H_1 if

$$\frac{\rho(H_1)}{\rho(H_0)} = \frac{C_{10} p(H_0|\boldsymbol{X})}{C_{01} p(H_1|\boldsymbol{X})} < 1 \tag{42}$$

which is rewritten as

$$\frac{p(\boldsymbol{X}|H_1)}{p(\boldsymbol{X}|H_0)} > \frac{C_{10} p(H_0)}{C_{01} p(H_1)} \equiv \beta \tag{43}$$

It follows from (37) and (38) that the Bayesian detection threshold for GC is given by

$$\gamma^2 > 1 - \left(\frac{K(N+q,M)}{K(q,M)} \frac{1}{\beta} \right)^{\frac{1}{N+q}} \tag{44}$$

7. CONCLUSION

The GC estimate is a well established test statistic for detecting the presence of a common but unknown signal from measurements made at multiple sensors. We have derived the GC detector from Bayesian perspective, showing that it arises naturally when one tests the hypothesis that the signals in the channels are independent Gaussian processes against the hypothesis that the processes have some arbitrary correlation. This was achieved by introducing suitable non-informative priors for the covariance matrices across the channels under the respective hypotheses. Reduced likelihoods were then obtained by marginalizing the joint distribution of the data and the covariance matrix in each case. The likelihood ratio was then shown to be a monotonic function of the GC detection statistic.

8. REFERENCES

[1] G. C. Carter, "Coherence and time delay estimation," *Proc. IEEE*, vol. 75, no. 2, pp. 236–255, 1987.

[2] A. H. Nuttall, "Invariance of distribution of coherence estimate to second-channel statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 29, no. 2, pp. 120–122, 1981.

[3] H. Gish and D. Cochran, "Invariance of the magnitude-squared coherence estimate with respect to second-channel statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 12, pp. 1774–1776, 1987.

[4] R. D. Trueblood and D. L. Alspach, "Multiple coherence as a detection statistic," Tech. Rep. NOSC 265, Naval Ocean Systems Center, 1978.

[5] H. Gish and D. Cochran, "Generalized coherence," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing,*, Apr. 1988, pp. 2745–2748.

[6] A. Clausen and D. Cochran, "An invariance property of the generalized coherence estimate," *IEEE Trans. Signal Processing*, vol. 45, no. 4, pp. 1065–1067, 1997.

[7] A. Clausen and D. Cochran, "Asymptotic analysis of the generalized coherence estimate," *IEEE Trans. Signal Processing*, vol. 49, no. 1, pp. 45–53, 2001.

[8] D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Trans. Signal Processing*, vol. 43, no. 9, pp. 2049–2057, 1995.

[9] D. Ramírez, J. Vía, I. Santamaría, and L. L. Scharf, "detection of spatially correlated Gaussian time series," *IEEE Trans. Signal Processing*, vol. 58, no. 10, pp. 5006–5015, 2010.

[10] N. R. Goodman, "Statistical analysis based on a certain multivariate complex gaussian distribution (An Introduction)," *Ann. Math. Statist.*, vol. 34, no. 1, pp. 152–177, 1963.

[11] S. T. Smith, "Covariance, subspace, and intrinsic Cramér-Rao bounds," *IEEE Trans. Signal Processing*, vol. 53, no. 5, pp. 1610–1630, 2005.

[12] R. E. Kass and L. Wasserman, "Formal rules of selecting prior distributions: A review and annotated bibliography," *J. Amer.Statist.Assoc.*, vol. 91, pp. 1343–1370, 1996.

[13] H. Jeffreys, *Theory of Probability*, Oxford University Press, 1961.

[14] E. T. Jaynes, "Information theory and statistical mechanics," *The Physical review*, vol. 106, no. 4, pp. 620–630, 1957.