Bounds for the estimation of matrix-valued parameters of a Gaussian random process

Lorena Leon Arencibia, Herwig Wendt, Jean-Yves Tourneret

Signal Processing 211 (2023) 109106

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Short communication

Bounds for the estimation of matrix-valued parameters of a Gaussian random process

Lorena Leon[a,1], Herwig Wendt[a,1,*], Jean-Yves Tourneret[a,1]

Université de Toulouse, IeIT/INP-ENSEEIHT/TéSA, France

ARTICLE INFO

Article history:
Received 18 November 2022
Revised 27 March 2023
Accepted 16 May 2023
Available online 20 May 2023

Keywords:
Bayesian Cramér-Rao lower bound
Wishart random matrices
Multivariate multifractal analysis

ABSTRACT

This paper derives and studies Bayesian Cramér-Rao lower bounds for the mean squared error of covariance matrices that are structured as weighted sums of symmetric positive definite matrices associated with a circularly-symmetric Gaussian statistical model. This model naturally appears in a number of important applications, including multivariate multifractal analysis and vector-valued additive Gaussian processes. As an intermediary result, we derive a novel expression for the expectation of compositions of Wishart random matrices. We provide extensive numerical simulation results for analyzing the derived bounds and their properties, and illustrate their use for the multifractal analysis of bivariate time series.

1. Introduction

The estimation performance for the parameters of a statistical model can be analyzed by establishing fundamental lower bounds for the mean squared error (MSE) of these parameters. The lower bounds for parameters that are assigned an *a priori* probability distribution are commonly referred to as Bayesian bounds [1]. In the spirit of [2], this paper derives the Bayesian Cramér-Rao bound (in short BB) for the MSE of the matrix-valued parameters of a statistical model described in the next section.

Problem statement and statistical model. Consider N independent zero mean complex circularly-symmetric Gaussian random vectors $\boldsymbol{z}_n \in \mathbb{C}^R, n = 1, \ldots, N$, such that $\mathbb{E}[\boldsymbol{z}_n] = \boldsymbol{0}$, $\mathbb{E}[\boldsymbol{z}_n \boldsymbol{z}_n^T] = \boldsymbol{0}$, and $\mathbb{E}[\boldsymbol{z}_n \boldsymbol{z}_n^H] = \boldsymbol{R}_n$, i.e., $\boldsymbol{z}_n \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_n)$, where $\mathbb{E}[\cdot]$ denotes the usual mathematical expectation and operators $(\cdot)^T$ and $(\cdot)^H$ compute the matrix/vector transpose and the matrix/vector conjugate transpose. The covariance matrix $\boldsymbol{R}_n$ is assumed to be real-valued positive definite (p.d.) and of the form $\boldsymbol{R}_n = \boldsymbol{\Sigma}_1 g_1(n) + \boldsymbol{\Sigma}_2 g_2(n)$, for $n = 1, \ldots, N$, where $g_1(\cdot), g_2(\cdot) > 0$ are known real-valued functions and $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$ are the $R \times R$ symmetric p.d. matrix-valued parameters to be estimated. Thus, the vector of the RN samples arranged as $\boldsymbol{z} = (\boldsymbol{z}_1^T, \ldots, \boldsymbol{z}_N^T) \in \mathbb{C}^{RN}$ can be modeled as a zero mean complex circularly-symmetric Gaussian random vector with the $RN \times RN$ real-valued covariance matrix $\boldsymbol{R} = \boldsymbol{\Sigma}_1 \otimes \boldsymbol{G}_1 + \boldsymbol{\Sigma}_2 \otimes \boldsymbol{G}_2$, where $\otimes$ is the Kronecker product and $\boldsymbol{G}_1, \boldsymbol{G}_2$ are known diagonal matrices whose nth diagonal entries are given by $[\boldsymbol{G}_i]_{nn} = g_i(n)$, for $i \in \{1, 2\}$.

The assumption that $\boldsymbol{R}$ is real-valued is not strictly necessary and can be relaxed to complex-valued matrices as in [2] using the same ideas and expressions as below. Here, we focus on the real-valued case because it appears in the application motivating this work, and for convenience of presentation.

Motivation and related works. Gaussian models with a zero mean vector and a covariance matrix $\boldsymbol{R}$ structured as above arise in several important contexts. An important example - and the one motivating this work - is given by the multivariate multifractal analysis for which the matrices $\boldsymbol{\Sigma}_i$, $i \in \{1, 2\}$, correspond to parameters that describe the joint geometry of the fluctuations of the *pointwise regularity* of the data components, see, e.g., [3] for definitions, intuitions and applications of multifractal analysis, [4,5] for the multivariate case and [6–8] for recent Bayesian estimation frameworks. Another important example is given by vector-valued additive Gaussian processes, in which the matrices $\boldsymbol{G}_i$ subsume the kernels for the temporal/spatial isotropic covariance models, expressed in the Fourier domain, and $\boldsymbol{\Sigma}_i$ are the associated point covariance matrices for the vector-valued variates, see, e.g., [9–11]. Note that it is straightforward to generalize the expressions derived in this paper to more than 2 summands. For ease of presentation, we treat here the case with 2 summands, without loss of generality.

* Corresponding author.
E-mail addresses: lorena.leon@irit.fr (L. Leon), herwig.wendt@irit.fr (H. Wendt), jean-Yves.tourneret@irit.fr (J.-Y. Tourneret).

[1] Work supported by ANR-18-CE45-0007 MUTATION, France. The authors would like to thank Prof. G. Letac for enthusiastically sharing his expert knowledge on Wishart distributions.

https://doi.org/10.1016/j.sigpro.2023.109106

Goals, outline and contributions. This paper, we consider $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$ to be unknown with inverse Wishart (IW) prior distributions. Our goal is to derive lower bounds for the MSE of estimators of $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$. Assuming the estimation to be conducted within a Bayesian formulation, Section 2 derives the BBs of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ and analytically studies their properties. Section 3 studies the properties of the bounds in this framework using Monte Carlo simulations. The use of the proposed bounds for the parameters associated with the bivariate multifractal spectrum is finally illustrated in Section 3.3. The main contributions of this paper are i) the derivation of the BB for the above specified statistical model, which is a new theoretical result obtained from (4), (8), (9) and (11), ii) the derivation of a novel closed-form expression for computing non-trivial expectations involving Wishart random matrices, see Proposition 1, iii) the study of the analytic properties of the bounds (see Section 2) and iv) extensive numerical experiments and results that validate and illustrate the obtained theoretical expressions of the bounds (see Section 3).

2. Bayesian Cramér-Rao bound

This section derives the BB for the MSE of estimators of $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$, when $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are assigned independent IW priors, i.e., for $i \in \{1,2\}$, $\boldsymbol{\Sigma}_i \sim \mathcal{IW}(\nu_i, \boldsymbol{\Omega}_i)$, with ν_i degrees of freedom ($\nu_i \in \mathbb{R}$ and $\nu_i > R+1$), and mean matrix $(\nu_i - R - 1)^{-1}\boldsymbol{\Omega}_i$, where $\boldsymbol{\Omega}_i$ is a real-valued p.d. scale matrix. To this end, we make use of the following novel results.

Proposition 1. *Moments of the type* $\mathbb{E}[\boldsymbol{WAWBW}]$.

If $\boldsymbol{\Sigma} \sim \mathcal{IW}(\nu, \boldsymbol{\Omega})$, *then* $\boldsymbol{W} = \boldsymbol{\Sigma}^{-1}$ *has the Wishart distribution* $\mathcal{W}(\nu, \boldsymbol{\Delta} = \boldsymbol{\Omega}^{-1})$. *Then, for any pair of real-valued symmetric matrices* $(\boldsymbol{A}, \boldsymbol{B})$*:*

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\Sigma}}\left[\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\boldsymbol{B}\boldsymbol{\Sigma}^{-1}\right] &= \boldsymbol{\Delta}\boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{B}\boldsymbol{\Delta}(\nu^3+2\nu^2+\nu) + \boldsymbol{\Delta}\boldsymbol{B}\boldsymbol{\Delta}\boldsymbol{A}\boldsymbol{\Delta}(\nu^2+3\nu) \\ &\quad + [tr(\boldsymbol{\Delta}\boldsymbol{A})]\boldsymbol{\Delta}\boldsymbol{B}\boldsymbol{\Delta}(\nu^2+\nu) + [tr(\boldsymbol{\Delta}\boldsymbol{B})]\boldsymbol{\Delta}\boldsymbol{A}\boldsymbol{\Delta}(\nu^2+\nu) \\ &\quad + \boldsymbol{\Delta}\Big[(\nu^2+\nu)tr(\boldsymbol{\Delta}\boldsymbol{A}\boldsymbol{\Delta}\boldsymbol{B}) + \nu tr(\boldsymbol{\Delta}\boldsymbol{A})tr(\boldsymbol{\Delta}\boldsymbol{B})\Big]. \end{aligned} \tag{1}$$

The proof of Proposition 1 can be conducted using the approach detailed in [12]. Moreover, according to [13], for any real-valued symmetric matrix $\boldsymbol{A}$, we have

$$\mathbb{E}_{\boldsymbol{\Sigma}}\left[\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{\Sigma}^{-1}\right] = (\nu^2+\nu)\boldsymbol{\Delta}\boldsymbol{A}\boldsymbol{\Delta} + \nu tr(\boldsymbol{\Delta}\boldsymbol{A})\boldsymbol{\Delta}. \tag{2}$$

Definitions. Let $\boldsymbol{\theta} \in \mathbb{R}^p$, with $p = R^2 + R$, the vector obtained by concatenating the vectors $\text{vec}_{\text{triu}}(\boldsymbol{\Sigma}_1)$ and $\text{vec}_{\text{triu}}(\boldsymbol{\Sigma}_2)$, where the matrix operator $\text{vec}_{\text{triu}}(\boldsymbol{A})$ returns the vector of the elements of the upper triangular part of $\boldsymbol{A}$. Note that the first and the last $p/2$ elements of $\boldsymbol{\theta}$, denoted as $\boldsymbol{\theta}_{1:\frac{p}{2}}$ and as $\boldsymbol{\theta}_{\frac{p}{2}+1:p}$, correspond to the main diagonal and all elements of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ above the diagonal, respectively. In the following, the matrices $\boldsymbol{R}_n$ will be denoted as $\boldsymbol{R}_n(\boldsymbol{\theta})$ to emphasize the dependence of $\boldsymbol{R}_n$ on $\boldsymbol{\theta}$. The evaluation of the BBs of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ requires to invert the posterior Fisher information matrix (PFIM) defined as [1]

$$\boldsymbol{F} = \mathbb{E}_{\boldsymbol{z},\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2}\left[-\frac{\partial^2 L(\boldsymbol{z},\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right], \tag{3}$$

where $L(\boldsymbol{z},\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2)$ is the joint log-likelihood of the model, which is twice differentiable with respect to (w.r.t.) $\boldsymbol{\theta}$ and has a bounded support independent of $\boldsymbol{\theta}$. These required regularity conditions ensure the existence of the BB. Equation (3) can be rewritten as

$$\begin{aligned}\boldsymbol{F} &= \mathbb{E}_{\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2}\left[\mathbb{E}_{\boldsymbol{z}|\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2}\left[-\frac{\partial^2 L(\boldsymbol{z}\mid\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right] - \frac{\partial^2\pi_1(\boldsymbol{\Sigma}_1)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T} - \frac{\partial^2\pi_2(\boldsymbol{\Sigma}_2)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right], \\ &= \mathbb{E}_{\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2}\left[\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{F}_{\boldsymbol{\Sigma}_1} + \boldsymbol{F}_{\boldsymbol{\Sigma}_2}\right], \end{aligned} \tag{4}$$

where $L(\boldsymbol{z}\mid\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2)$ is the log-likelihood of $\boldsymbol{z}$ given $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, which can be expressed as

$$L(\boldsymbol{z}\mid\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2) = K - \ln\det\boldsymbol{R}(\boldsymbol{\theta}) - \boldsymbol{z}^H\boldsymbol{R}(\boldsymbol{\theta})^{-1}\boldsymbol{z}, \tag{5}$$

with $K = -N\ln\pi$. Moreover, $\pi_i(\boldsymbol{\Sigma}_i)$ is the log-prior of $\boldsymbol{\Sigma}_i$ defined as

$$\pi_i(\boldsymbol{\Sigma}_i) = -((\nu_i+R+1)/2)\ln\det\boldsymbol{\Sigma}_i - \frac{1}{2}tr(\boldsymbol{\Omega}_i\boldsymbol{\Sigma}_i^{-1}) + \text{constant}, \tag{6}$$

where $tr(\cdot)$ denotes the trace operator. Note that functions (5) and (6) also satisfy the regularity conditions ensuring the existence of the BBs of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$.

Computation of $\boldsymbol{F}_{\boldsymbol{\theta}}$, $\boldsymbol{F}_{\boldsymbol{\Sigma}_1}$ and $\boldsymbol{F}_{\boldsymbol{\Sigma}_2}$. The matrix $\boldsymbol{F}_{\boldsymbol{\theta}}$ is known as the Fisher information matrix of $\boldsymbol{\theta}$. Since $\boldsymbol{z}$ is a zero mean complex circularly-symmetric Gaussian vector, the element of $\boldsymbol{F}_{\boldsymbol{\theta}}$ located at the kth row and lth column, denoted as $[\boldsymbol{F}_{\boldsymbol{\theta}}]_{kl}$ for $k,l \in \{1,\ldots,p\}$, can be calculated as [14,15]

$$[\boldsymbol{F}_{\boldsymbol{\theta}}]_{kl} = \text{tr}\left\{\boldsymbol{R}^{-1}(\boldsymbol{\theta})\frac{\partial\boldsymbol{R}(\boldsymbol{\theta})}{\partial\theta_k}\boldsymbol{R}^{-1}(\boldsymbol{\theta})\frac{\partial\boldsymbol{R}(\boldsymbol{\theta})}{\partial\theta_l}\right\}. \tag{7}$$

Note that in general the computation of (7) needs $O(R^3N^3)$ operations because it requires to invert the matrix $\boldsymbol{R}(\boldsymbol{\theta})$. This computation can lead to a high computational cost for large values of N and R. We propose to overcome this limitation by exploiting the block diagonal structure of $\boldsymbol{R}(\boldsymbol{\theta})$. This structure allows the inverse of the matrix $\boldsymbol{R}(\boldsymbol{\theta})$ to be computed using the inverses of the $R\times R$ diagonal blocks $\boldsymbol{R}_n(\boldsymbol{\theta})$ of $\boldsymbol{R}(\boldsymbol{\theta})$. Specifically, $\boldsymbol{R}_n^{-1}(\boldsymbol{\theta}) = (\boldsymbol{\Sigma}_1 g_1(n) + \boldsymbol{\Sigma}_2 g_2(n))^{-1}$ and

$$\frac{\partial\boldsymbol{R}_n(\boldsymbol{\theta})}{\partial\theta_l} = \boldsymbol{B}_{n,l} = \begin{cases}\boldsymbol{J}_{1,l}g_1(n) & \text{if } l\in\{1,\ldots,\frac{p}{2}\},\\ \boldsymbol{J}_{2,l}g_2(n) & \text{if } l\in\{\frac{p}{2}+1,\ldots,p\},\end{cases}$$

where $\boldsymbol{J}_{i,l} = \dfrac{\partial\boldsymbol{\Sigma}_i}{\partial\theta_l}$ does not depend on $\boldsymbol{\theta}$, hence $\dfrac{\partial^2\boldsymbol{\Sigma}_i}{\partial\theta_k\partial\theta_l} = \boldsymbol{0}$. Thus, (7) can be rewritten as

$$[\boldsymbol{F}_{\boldsymbol{\theta}}]_{kl} = \sum_{n=1}^{N}\text{tr}\left\{\boldsymbol{R}_n^{-1}(\boldsymbol{\theta})\boldsymbol{B}_{n,k}\boldsymbol{R}_n^{-1}(\boldsymbol{\theta})\boldsymbol{B}_{n,l}\right\}, \tag{8}$$

which can be computed with $O(R^3N)$ operations. Overall, the values of R considered in this paper are relatively small $R \ll N$ and lead to feasible computational times.

On the other hand, the following result is obtained

$$\begin{aligned}[\boldsymbol{F}_{\boldsymbol{\Sigma}_i}]_{kl} &= -\frac{\partial^2\pi_i(\boldsymbol{\Sigma}_i)}{\partial\theta_k\partial\theta_l} \\ &= ((\nu_i+R+1)/2)tr(\boldsymbol{\Sigma}_i^{-1}\boldsymbol{J}_{i,k}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{J}_{i,l}) \\ &\quad -(1/2)tr(\boldsymbol{\Omega}_i[\boldsymbol{\Sigma}_i^{-1}\boldsymbol{J}_{i,k}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{J}_{i,l}\boldsymbol{\Sigma}_i^{-1} \\ &\quad +\boldsymbol{\Sigma}_i^{-1}\boldsymbol{J}_{i,l}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{J}_{i,k}\boldsymbol{\Sigma}_i^{-1}]). \end{aligned} \tag{9}$$

Expectations. This section explains how to compute the expectations in (4). Calculating the expectation $\mathbb{E}_{\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2}\left[\boldsymbol{F}_{\boldsymbol{\Sigma}_1} + \boldsymbol{F}_{\boldsymbol{\Sigma}_2}\right]$ reduces to determining the expectation of $-\dfrac{\partial^2\pi_1(\boldsymbol{\Sigma}_1)}{\partial\boldsymbol{\theta}_{1:\frac{p}{2}}\partial\boldsymbol{\theta}_{1:\frac{p}{2}}^T}$ w.r.t. $\boldsymbol{\Sigma}_1$ and the expectation of $-\dfrac{\partial^2\pi_2(\boldsymbol{\Sigma}_2)}{\partial\boldsymbol{\theta}_{\frac{p}{2}+1:p}\partial\boldsymbol{\theta}_{\frac{p}{2}+1:p}^T}$ w.r.t. $\boldsymbol{\Sigma}_2$, since the other terms equal 0. Both expectations have a closed-form expression that can be determined using the matrix expectations (1) and (2). The challenge here is to compute the expectation $\mathbb{E}_{\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2}[\boldsymbol{F}_{\boldsymbol{\theta}}]$, which involves calculating the expectation of the expression $\boldsymbol{R}_n^{-1}\boldsymbol{B}_{n,k}\boldsymbol{R}_n^{-1}\boldsymbol{B}_{n,l}$ with respect to $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$ for all $n = 1,\ldots,N$.

This computation is possible provided we can compute the expectation

$$\mathbb{E}_{\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2}\left[(a\boldsymbol{\Sigma}_1+b\boldsymbol{\Sigma}_2)^{-1}\boldsymbol{A}(a\boldsymbol{\Sigma}_1+b\boldsymbol{\Sigma}_2)^{-1}\boldsymbol{B}\right] \tag{10}$$

for $a, b \in \mathbb{R}_+$, $\boldsymbol{\Sigma}_1 \sim \mathcal{IW}(\nu_1, \boldsymbol{\Omega}_1)$, $\boldsymbol{\Sigma}_2 \sim \mathcal{IW}(\nu_2, \boldsymbol{\Omega}_2)$ and any pair of symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. If a or b equal zero, (10) can be calculated using (2). Otherwise, we propose to approximate (10) numerically via a Monte Carlo algorithm. Given a, b, $\boldsymbol{A}$ and $\boldsymbol{B}$, we can generate a large number S of samples $\{\boldsymbol{\Sigma}_1^{(s)}, \boldsymbol{\Sigma}_2^{(s)}\}_{s=1}^S$ according to IW distributions, compute $\boldsymbol{E}^{(s)} = \left[\left(a\boldsymbol{\Sigma}_1^{(s)}+b\boldsymbol{\Sigma}_2^{(s)}\right)^{-1}\boldsymbol{A}\left(a\boldsymbol{\Sigma}_1^{(s)}+b\boldsymbol{\Sigma}_2^{(s)}\right)^{-1}\boldsymbol{B}\right]$ and approximate (10) by the average of the matrices $\{\boldsymbol{E}^{(s)}\}_{s=1}^S$.

The MSE of any estimator $(\widehat{\boldsymbol{\Sigma}}_1, \widehat{\boldsymbol{\Sigma}}_2)$ of $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$ is defined as the trace of the error covariance matrix $\text{MSE} = tr(\mathbb{E}_{\boldsymbol{\theta}}[(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^T])$, where $\hat{\boldsymbol{\theta}}$ is the vector obtained by concatenating the vectors $\text{vec}_{\text{triu}}(\widehat{\boldsymbol{\Sigma}}_1)$ and $\text{vec}_{\text{triu}}(\widehat{\boldsymbol{\Sigma}}_2)$. Finally, the inverse of the PFIM (3) denoted as $[\boldsymbol{F}]^{-1}$ yields the desired lower bound for the MSE of any estimator $(\widehat{\boldsymbol{\Sigma}}_1, \widehat{\boldsymbol{\Sigma}}_2)$ of $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$, when these matrices are IW distributed:

$$\text{MSE} \geq \text{BB} = tr([\boldsymbol{F}]^{-1}). \tag{11}$$

Analytic properties. Assuming that a or b are zero and $\boldsymbol{\Delta}$ is diagonal, (11) can be computed in closed-form. In particular, the entries of $\boldsymbol{F}$ are given by

$$[\boldsymbol{F}]_{kl} = \gamma_1 tr(\boldsymbol{\Delta}\boldsymbol{J}_k)tr(\boldsymbol{\Delta}\boldsymbol{J}_l) + \gamma_2 tr(\boldsymbol{\Delta}\boldsymbol{J}_k\boldsymbol{\Delta}\boldsymbol{J}_l) \tag{12}$$

where $\gamma_1 = \frac{1}{2}(\nu^2 + \nu(2N-R+3))$ and $\gamma_2 = \frac{1}{2}(\nu^3 + \nu^2(2N-R+6) + \nu(2N+9-R))$. When $\boldsymbol{\Delta}$ is the identity matrix, the first term of the PFIM is a matrix with zero entries except for an $R \times R$ diagonal block with non-zero entries whereas the second term is a diagonal matrix. In this case, it can be shown that the PFIM has $R-1$ eigenvalues (ev.) equal to γ_2, $R(R+1)/2 - R = (R^2-R)/2$ ev. equal to $2\gamma_2$ and one ev. equal to $R\gamma_1 + \gamma_2$. The trace of $[\boldsymbol{F}]^{-1}$ with the above assumptions, denoted as aBB (for approximate BB), is

$$\text{aBB} = tr([\boldsymbol{F}]^{-1}) = \frac{R(R-1)}{4\gamma_2} + \frac{R((R-1)\gamma_1+\gamma_2)}{\gamma_2^2 + R\gamma_1\gamma_2}. \tag{13}$$

This shows that the bound behaves asymptotically as ν^{-3} as $\nu \to +\infty$, and as a second order polynomial in ν as $\nu \to R+1$. Moreover, the behavior is asymptotically linear in R, and a second order polynomial in R as $R \to 1$. As expected, the asymptotic decay with sample size is N^{-1}.

3. Numerical illustrations

In this section, extensive numerical simulations are used to study the properties of the BB for the MSE of any estimator of the pair of matrices $\boldsymbol{\Lambda} = (\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$ in the above probabilistic framework, and compare the bounds against the MSEs of Bayesian estimators.

3.1. Monte Carlo simulations

Estimation algorithms. We consider the maximum a posteriori (MAP) and minimum mean square (MMSE) estimators, defined by:

$$\widehat{\boldsymbol{\Lambda}}^{\text{MAP}} = \underset{\boldsymbol{\Lambda}}{\text{argmax}}\, p(\boldsymbol{\Lambda} \mid \boldsymbol{z}), \tag{14}$$

$$\widehat{\boldsymbol{\Lambda}}^{\text{MMSE}} = \mathbb{E}[\boldsymbol{\Lambda} \mid \boldsymbol{z}]. \tag{15}$$

In order to compute (14) and (15), we use a Gibbs sampler [16] to generate a large collection N_{mc} of samples distributed according the posterior distribution equation to be recalled here. After a burn-in period, where the first N_{bi} samples are discarded, the Bayesian estimators $\widehat{\boldsymbol{\Lambda}}^{\text{MAP}}$ and $\widehat{\boldsymbol{\Lambda}}^{\text{MMSE}}$ are approximated using the last generated samples.

Simulation setup. Unless otherwise stated $R=2$, $N=2^8$, $\boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_2 = \mathbb{I}_R$ ($R \times R$ identity matrix) and $\nu_1 = \nu_2 = 80$. Without loss of generality, we use the functions $g_1(n) = 2\pi\cos^2(\boldsymbol{x}[n]) + 0.1$ and $g_2(n) = 2\pi\sin^2(\boldsymbol{x}[n]) + 0.1$, where $\boldsymbol{x}$ is the vector of N components whose values have been generated in the interval [0,2], equispaced with a distance of $2/(N-1)$. In all cases, the sample MSE of the estimators is computed as the average of the trace of the error covariance matrix over 1000 independent realizations. Gibbs samplers are run with $N_{\text{mc}} = 1000$ and $N_{\text{bi}} = 500$, and (10) is approximated (when needed) as described before using $S = 200$.

3.2. Performance analysis

Performance vs. sample sizes. Fig. 1 displays the BB, its approximation aBB, and the MSEs of the MMSE and MAP estimators, for various sample sizes, where $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$ are random matrices with IW prior distributions. The following comments are appropriate:

- The BB decreases as N^{-1} when $N \to \infty$ and to a constant when $N \to 0$.
- The BB and its approximation aBB are asymptotically close but tend to different constants for small sample size.
- *BB vs. MSE:* The MSEs of both estimators are approaching the BBs when N increases - the more data, the tighter the bounds.
- *MMSE vs. MAP:* Overall, the MMSE estimator has better performance than the MAP estimator, in particular for small sample size. This result was expected since the MMSE estimator minimizes the MSE.

Performance vs. degrees of freedom. Fig. 2 compares the MSEs of the MMSE and MAP estimators and the BBs for various degrees of freedom, $\nu_1 = \nu_2 \in \{10, 15, 20, 25, \ldots, 120\}$. We can observe that the BB decreases when ν_1 and ν_2 increase; indeed, in that case, the priors are more informative. The approximation aBB is very similar to BB and predicts that this decay is of order ν^{-3}. Moreover, the values of the MSEs for both the MAP and MMSE estimators are observed to be significantly larger than the lower BB for small values for ν_1 and ν_2 (uninformative priors), but very close to the bound for large values of ν_1, ν_2 (informative priors).

Performance vs. number of components. Fig. 3 displays the MSEs of the MMSE and MAP estimators and the BBs when the number of components R - thus, the number of parameter p - is varied, specifically $R \in \{1, 2, 3, \ldots, 10\}$. We can observe that: *1) BB vs. R:* The BB increases with increasing values for R and is very tightly approximated by aBB, thus suggesting an asymptotically linear behavior in R. *2) BB vs. MSE:* The values taken by the MSE and the BB are very similar for a small number of components/parameters. For large values of R, the bound is slightly less tight. Since the sample size is fixed here, this behavior is coherent. Indeed, we would expect that larger sample sizes are required to converge to the BB when more parameters are estimated.

3.3. Application to a multivariate multifractal analysis

Finally, the theoretical results of this paper are applied to a practical example related to multivariate multifractal analysis [4,5,17]. In particular, the model and estimation framework of [7] are considered for bivariate time series ($R=2$). In this context, the Fourier transform of the logarithm of wavelet leaders, defined as nonlinear and nonlocal transformations of wavelet coefficients, approximately obeys the data model considered in this paper, where the elements of $\boldsymbol{\Sigma}_1$ are directly related to the multifractality of the data, i.e., $\boldsymbol{\Sigma}_1 = -[c_{20},\ c_{11};\ c_{11},\ c_{02}]$ and thus $\boldsymbol{\theta}_{1:3} = -(c_{20},\ c_{11},\ c_{02})$, where $c_{20}, c_{02} < 0$ are related to the widths

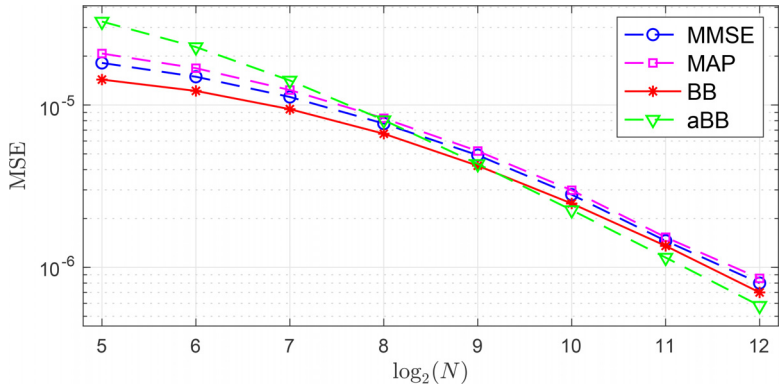

Fig. 1. Comparison between the sample MSEs of the MAP and MMSE estimators averaged over 1000 independent realizations versus the BB for sample sizes $N \in \{2^5, 2^6, \ldots, 2^{12}\}$, $\nu_1 = \nu_2 = 80$ and $R = 2$.

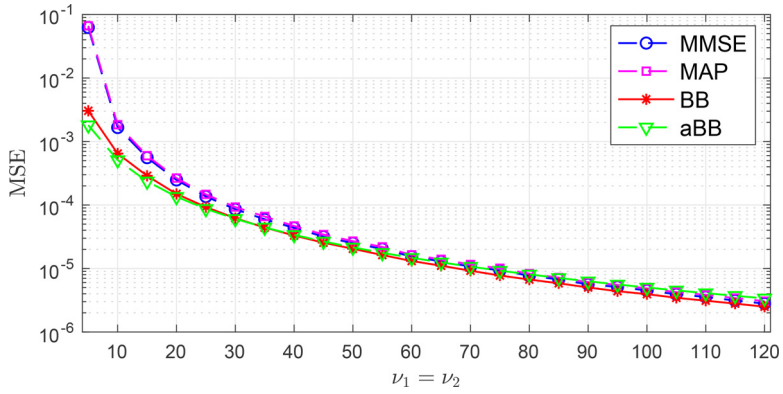

Fig. 2. Comparison between the sample MSEs of the MAP and MMSE estimators averaged over 1000 independent realizations versus the BB, varying the degrees of freedom $\nu_1 = \nu_2$, for $R = 2$ and $N = 2^8$.

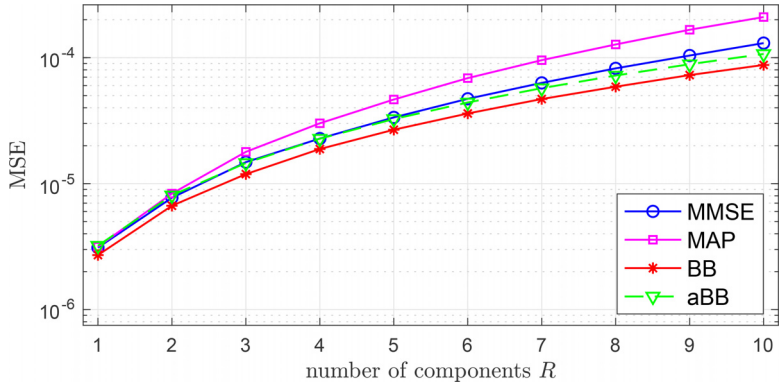

Fig. 3. Comparison between the sample MSEs of the MAP and MMSE estimators averaged over 1000 independent realizations versus the BB, varying the number of components R, for $N = 2^8$, $\mathbf{\Omega}_1 = \mathbf{\Omega}_2 = \mathbb{I}_R$ and $\nu_1 = \nu_2 = 80$.

of the marginal multifractal spectra, and c_{11} quantifies the joint multifractality. The matrix $\mathbf{\Sigma}_2$ is an adjustment parameter that essentially subsumes the short-lag autocorrelation of log-wavelet leaders. Moreover, one can define a *multifractal correlation* parameter as $\rho_{\text{mf}} = -\frac{c_{11}}{\sqrt{c_{20}c_{02}}}$ (with $-1 \leq \rho_{\text{mf}} \leq 1$), that quantifies a dependence beyond linear correlation between the time series. Note that the BB of ρ_{mf} is obtained from the BB of $\mathbf{\Sigma}_1$ using the functional invariance principle [1].

Simulation study. This section considers 2000 independent copies of $2^{10} \times 2$ time series of a canonical multifractal model process, i.e., a bivariate multifractal random walk (bMRW) [17–19], to compute the sample MSE of the MMSE estimator, and the BB, for different multifractal parameter settings, controlled by $\mathbf{\Sigma}_1$. In a first experiment, $\mathbf{\Sigma}_1$ is generated using $\mathbf{\Omega}_1 = [0.5,\ 0;\ 0,\ \omega]$, with $0.37 \leq \omega \leq 1.2$ and $\nu_1 = 10$, leading to realistic expected values for the multifractal parameters, i.e., $-c_{20} = 0.05$ and $-c_{02} = \{0.037, \ldots, 0.12\}$. In a second experiment $\mathbf{\Omega}_1 = [0.4,\ \gamma;\ \gamma,\ 0.4]$ with γ tuned such that $0 \leq \rho_{\text{mf}} \leq 0.8$ in average. The parameters of $\mathbf{\Sigma}_2$ cannot be controlled by the bMRW synthesis and are thus unknown, and we set $\mathbf{\Omega}_2 = \mathbf{\Omega}_1$ and $\nu_2 = \nu_1$. Results for the two experiments are presented in Fig. 4 (top and bottom row, respectively). They indicate that the derived BBs provide good indications for the variations of the observed MSE of the multifractal parameter estimates. In particular, they show that: 1) the MSE of the estimator of c_{20} does not depend on c_{02}, which is expected since c_{20} is a marginal parameter of the first data component that is independent of c_{02}; 2) the MSE of the estimator of c_{02} increases with c_{02}. Indeed, c_{02} controls the variance of the marginal likelihood of the second data component; 3) the MSE of the estimator of c_{11} also increases with c_{02} because ρ_{mf} is held fixed so that c_{11} and

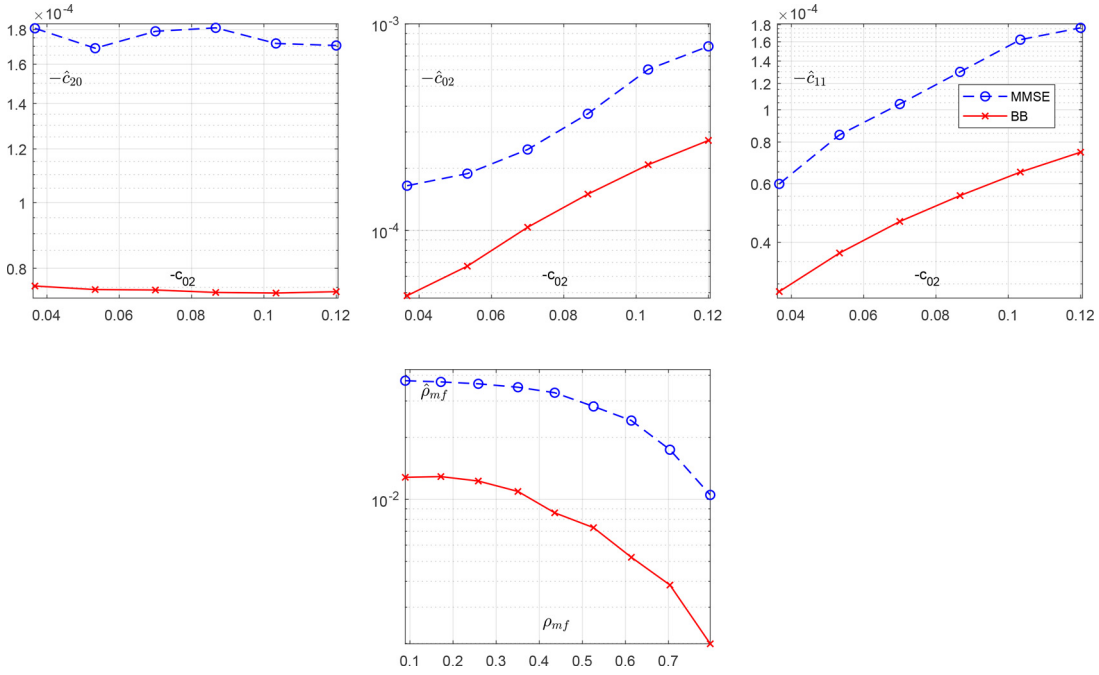

Fig. 4. Sample MSE for multifractal parameters c_{20}, c_{02}, c_{11} as a function of c_{02} (top) and sample MSE for multifractal correlation ρ_{mf} (bottom).

thus its covariance also increase; 4) the MSE of ρ_{mf} decreases in a non-trivial way when this parameter increases.

4. Conclusion

This paper derived and studied Bayesian Cramér-Rao lower bounds for the MSE of estimators of two symmetric positive definite matrices whose sum is the covariance matrix of a zero mean complex circularly-symmetric Gaussian model. To calculate these Bayesian bounds, a novel closed-form expression for a non-trivial expectation involving Wishart random matrices was provided. The properties of these Bayesian bounds were studied analytically. Various numerical simulations were used to validate the theoretical results and study the properties of the proposed Bayesian bounds. The practical interest of the bounds derived in this paper was finally illustrated for the estimation of the parameters of the bivariate multifractal spectrum.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

CRediT authorship contribution statement

Lorena Leon: Conceptualization, Writing – review & editing. **Herwig Wendt:** Conceptualization, Writing – review & editing. **Jean-Yves Tourneret:** Conceptualization, Writing – review & editing.

Data availability

No data was used for the research described in the article.

References

[1] H.L.V. Trees, Detection, Estimation, and Modulation Theory, Part I, John Wiley & Sons, Inc., 2001.
[2] O. Besson, S. Bidon, J.-Y. Tourneret, Bounds for estimation of covariance matrices from heterogeneous samples, IEEE Trans. Signal Process. 56 (7) (2008) 3357–3362.
[3] S. Jaffard, P. Abry, H. Wendt, Irregularities and scaling in signal and image processing: multifractal analysis, World scientific publishing, Singapore, 2015, pp. 31–116.
[4] S. Jaffard, S. Seuret, H. Wendt, R. Leonarduzzi, S. Roux, P. Abry, Multivariate multifractal analysis, Appl. Comput. Harmon. Anal. 46 (3) (2019) 653–663.
[5] S. Jaffard, S. Seuret, H. Wendt, R. Leonarduzzi, P. Abry, Multifractal formalisms for multivariate analysis, Proc. R. Soc. A 475 (2229) (2019).
[6] H. Wendt, S. Combrexelle, Y. Altmann, J.-Y. Tourneret, S. McLaughlin, P. Abry, Multifractal analysis of multivariate images using gamma Markov random field priors, SIAM J. Imag. Sci. (SIIMS) 11 (2) (2018) 1294–1316.
[7] L. Len, H. Wendt, J.-Y. Tourneret, P. Abry, Bayesian estimation for the parameters of the bivariate multifractal spectrum, in: Proc. European Signal Process. Conf. (EUSIPCO), 2021, pp. 1930–1934. Dublin, Ireland
[8] L. Leon, H. Wendt, J.-Y. Tourneret, P. Abry, A Bayesian framework for multivariate multifractal analysis, IEEE Trans. Signal Process. 70 (2022) 3663–3675.
[9] D.K. Duvenaud, H. Nickisch, C. Rasmussen, Additive Gaussian processes, Adv. Neural Inf. Process. Syst. 24 (2011).
[10] J. Hensman, N. Durrande, A. Solin, Variational fourier features for gaussian processes, J. Mach. Learn. Res. 18 (1) (2017) 5537–5588.
[11] L. Cheng, S. Ramchandran, T. Vatanen, N. Lietzén, R. Lahesmaa, A. Vehtari, H. Lähdesmäki, An additive Gaussian process regression model for interpretable non-parametric analysis of longitudinal data, Nat. Commun. 10 (1) (2019) 1–11.
[12] P. Graczyk, G. Letac, H. Massam, The hyperoctahedral group, symmetric group representations and the moments of the real wishart distribution, J. Theor. Probab. 18 (1) (2005) 1–42.
[13] T. Holgersson, J. Pielaszkiewicz, A collection of moments of the Wishart distribution, in: Recent Developments in Multivariate and Random Matrix Analysis, Springer International Publishing, 2020, pp. 147–162.
[14] R. Frehlich, Cramér-rao bound for Gaussian random processes and applications to radar processing of atmospheric signals, IEEE Trans. Geosci. Remote Sens. 31 (6) (1993) 1123–1131.
[15] B. Porat, B. Friedlander, Computation of the exact information matrix of Gaussian time series with stationary random components, IEEE Trans. Acoust., Speech, Signal Process. 34 (1) (1986) 118–130.
[16] C.P. Robert, G. Casella, Monte Carlo statistical methods, Springer, New York, USA, 2005.
[17] H. Wendt, R. Leonarduzzi, P. Abry, S. Roux, S. Jaffard, S. Seuret, Assessing cross-dependencies using bivariate multifractal analysis, in: Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proces. (ICASSP), 2018. Calgary, Canada
[18] E. Bacry, J. Delour, J. Muzy, Multifractal random walk, Phys. Rev. E 64 (2001).
[19] J.F. Muzy, J. Delour, E. Bacry, Modelling fluctuations of financial time series: from cascade process to stochastic volatility model, Eur. Phys. J. B 17 (2000).