HAL open science

Détection de changements dans les modèles gaussiens de rang 1

Rémi Beisson, Pascal Vallet, Audrey Giremus, Guillaume Ginolhac

To cite this version:

Rémi Beisson, Pascal Vallet, Audrey Giremus, Guillaume Ginolhac. Détection de changements dans les modèles gaussiens de rang 1. GRETSI, XXVIIIème Colloque Francophone de Traitement du Signal et des Images, GRETSI, Sep 2022, Nancy, France. hal-03778401

HAL Id: hal-03778401

https://hal.univ-grenoble-alpes.fr/hal-03778401

Submitted on 15 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Détection de changements dans les modèles gaussiens de rang 1

Rémi BEISSON[1], Pascal VALLET[1], Audrey GIREMUS[1], Guillaume GINOLHAC[2]

[1]Laboratoire IMS (CNRS, Univ. Bordeaux, Bordeaux INP)
351 Cours de la Libération, 33400 Talence, France

[2]LISTIC (Univ. Savoie/Mont-Blanc, Polytech Annecy)
5 chemin de Bellevue, 74940 Annecy, France

{remi.beisson, pascal.vallet, audrey.giremus}@ims-bordeaux.fr,
guillaume.ginolhac@univ-smb.fr

Résumé – Dans cet article, on étudie un test d'égalité entre les matrices de covariance de L séries temporelles gaussiennes multivariées de dimension M. On considère le cas particulier où chacune des L matrices de covariance est une perturbation de rang 1 de la matrice identité, correspondant à un modèle signal plus bruit. Une nouvelle statistique de test basée sur des estimées des valeurs propres des différentes matrices de covariance en jeu est proposée. Nous montrons que cette statistique de test est consistante et à taux de fausse alarme asymptotique constant dans le régime des grandes dimensions où le nombre d'observations disponibles $N_1, \ldots, N_L$ de chaque série ainsi que la dimension M tendent vers l'infini au même rythme. Des simulations montrent que la méthode proposée présente des performances appréciables par rapport aux approches alternatives.

Abstract – In this paper, we consider the problem of testing equality of covariance matrices of L Gaussian multivariate time series of dimension M in the special case where each of the L covariances is modeled as a rank 1 perturbation of the identity matrix, corresponding to a signal plus noise model. We propose a new test statistic based on the estimates of the eigenvalues of the different covariance matrices and which is shown to be consistant with asymptotic constant false alarm rate, in the regime where the observations $N_1, \ldots, N_L$ of each time series and the dimension M both converge to infinity at the same rate. Simulations show that the proposed method outperforms other relevant approaches.

1 Introduction

Considérons L séries temporelles $(\mathbf{y}_{n,1})_{n\in\mathbb{Z}}, \ldots, (\mathbf{y}_{n,L})_{n\in\mathbb{Z}}$ mutuellement indépendantes telles que :

$$\forall \ell \in \{1, \ldots, L\}, \quad (\mathbf{y}_{n,\ell})_{n\in\mathbb{Z}} \overset{\text{i.i.d.}}{\sim} \mathcal{N}_{\mathbb{C}^M}(\mathbf{0}, \mathbf{R}_\ell)$$

On s'intéresse ici à la détection de changements à travers la loi de $(\mathbf{y}_{n,\ell})_{n\in\mathbb{Z}}$ pour $\ell = 1, \ldots, L$, qui se formalise par le test d'hypothèse binaire suivant, portant sur l'égalité de matrices de covariance :

$$\begin{aligned} \mathcal{H}_0 : &\quad \mathbf{R}_1 = \ldots = \mathbf{R}_L \\ \mathcal{H}_1 : &\quad \exists (i,j) \in \{1, \ldots, L\}^2 \ / \ \mathbf{R}_i \neq \mathbf{R}_j \end{aligned} \tag{1}$$

Supposons que pour tout $\ell = 1, \ldots, L$, on dispose de N_ℓ observations $\mathbf{y}_{1,\ell}, \ldots, \mathbf{y}_{N_\ell,\ell}$, et notons $N = N_1 + \cdots + N_L$. La plupart des statistiques de test rencontrées dans la littérature [1], incluant le rapport de vraisemblances généralisé, consistent, sous réserve que $M < N_1, \ldots, N_L$, à comparer à un seuil des statistiques de la forme [1] :

$$S = \sum_{\ell=1}^{L} \frac{N_\ell}{N} \frac{1}{M} \sum_{k=1}^{M} \varphi\left(\lambda_k(\hat{\mathbf{R}}_\ell \hat{\mathbf{R}}^{-1})\right) \tag{2}$$

1. Pour une matrice hermitienne $\mathbf{A}$ de taille $(M \times M)$, on note $\lambda_1(\mathbf{A}) \geq \ldots \geq \lambda_M(\mathbf{A})$ ses valeurs propres par ordre décroissant.

où

$$\hat{\mathbf{R}}_\ell := \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} \mathbf{y}_{n,\ell} \mathbf{y}_{n,\ell}^*$$

désigne l'estimateur empirique standard de $\mathbf{R}_\ell$,

$$\hat{\mathbf{R}} := \sum_{\ell=1}^{L} \frac{N_\ell}{N} \hat{\mathbf{R}}_\ell$$

et où φ est une fonction continue sur $]0, +\infty[$. En particulier les statistiques de test de type S présentent l'intérêt d'être à taux de fausse alarme constant (*i.e.*, la loi sous $\mathcal{H}_0$ de S ne dépend pas de $\mathbf{R}_1, \ldots, \mathbf{R}_L$).

Pour des données de grande dimension, *i.e.* $M \gg 1$, il est courant dans de nombreuses applications que les changements potentiels dans les covariances $\mathbf{R}_\ell$ ne portent que sur une composante de rang réduit, et on considère alors le modèle :

$$\mathbf{R}_\ell = \mathbf{\Gamma}_\ell + \sigma^2 \mathbf{I} \tag{3}$$

avec $\mathbf{\Gamma}_\ell$ représentant la matrice de covariance de rang $K < M$ d'un signal utile et $\sigma^2\mathbf{I}$ la matrice de covariance d'un bruit additif spatialement blanc. Dans le régime asymptotique des *grandes dimensions* où $M, N_1, \ldots, N_L$ tendent vers l'infini au même rythme tandis que K, L restent fixes, il est possible de montrer, sous le modèle (3), et à la fois sous $\mathcal{H}_0$ et $\mathcal{H}_1$, que :

$$S \xrightarrow[M, N_1, \ldots, N_L \to \infty]{p.s.} \int_{\mathbb{R}} \varphi(x) f(x) \mathrm{d}x,$$

où f désigne la densité de la loi de *Wachter* [2], ce qui montre que les statistiques de test construites à partir de S ne sont pas consistantes.

Dans cet article, nous proposons une nouvelle statistique de test, non plus basée sur S, mais qui compare en un certain sens les plus grandes valeurs propres des matrices $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_L, \hat{\mathbf{R}}$. Dans le cas du modèle (3), la statistique proposée présente l'avantage d'être à la fois consistante et à taux de fausse alarme asymptotique constant. Pour ce faire, nous présentons une extension des résultats de [3] sur le comportement conjoint des plus grandes valeurs propres de $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_L, \hat{\mathbf{R}}$. Par souci de place, la plupart des résultats seront présentés pour $K = 1$, mais peuvent être étendus pour $K > 1$.

2 Spectre de $\hat{\mathbf{R}}$

Dans cette section, on étudie le comportement asymptotique aux 1er et 2nd ordres des plus grandes valeurs propres de $\hat{\mathbf{R}}$ lorsque les matrices $\mathbf{R}_1, \ldots, \mathbf{R}_L$ suivent le modèle (3).

Considérons en premier lieu les deux hypothèses suivantes, qui décrivent le régime des grandes dimensions et spécifient le comportement asymptotique de $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_L$ respectivement.

Hypothèse 1. *$N_1 = N_1(M), \ldots, N_L = N_L(M)$ sont fonctions de M telles que :*

$$\frac{M}{N_\ell} = c_\ell + o\left(\frac{1}{\sqrt{M}}\right)$$

où $c_1, \ldots, c_L \in]0, 1]$ avec L indépendant de M.

Le régime asymptotique décrit dans l'hypothèse 1 sera symbolisé par la notation $M \to \infty$ dans ce qui suit. On définit également :

$$c := \left(\sum_{\ell=1}^{L} \frac{1}{c_\ell}\right)^{-1} \quad \text{et} \quad \mathbf{\Gamma} := \sum_{\ell=1}^{L} \frac{N_\ell}{N} \mathbf{\Gamma}_\ell \tag{4}$$

Hypothèse 2. *Pour tout $k \in \{1, \ldots, K\}$, $\ell \in \{1, \ldots, L\}$,*

$$\lambda_k(\mathbf{\Gamma}_\ell) = \gamma_{k,\ell} + o\left(\frac{1}{\sqrt{M}}\right) \tag{5}$$

et $\forall k \in \{1, \ldots, KL\}$

$$\lambda_k(\mathbf{\Gamma}) = \gamma_k + o\left(\frac{1}{\sqrt{M}}\right) \tag{6}$$

Sous les hypothèses précédentes, on peut montrer que l'information liée au spectre de $\mathbf{\Gamma}$ se retrouve dans les L plus grandes valeurs propres de $\hat{\mathbf{R}}$.

Théorème 1. *Sous les hypothèses 1 et 2, $\forall k \in \{1, \ldots, KL\}$:*

$$\lambda_k\left(\hat{\mathbf{R}}\right) \xrightarrow[M\to\infty]{p.s.} \phi_c\left(\gamma_k, \sigma^2\right)$$

avec

$$\phi_c(\gamma, \sigma^2) := \begin{cases} \frac{(\gamma+\sigma^2)(\gamma+\sigma^2 c)}{\gamma} & \text{si } \gamma > \sigma^2\sqrt{c} \\ \sigma^2(1+\sqrt{c})^2 & \text{si } \gamma \leq \sigma^2\sqrt{c} \end{cases},$$

De plus, $\lambda_{KL+1}\left(\hat{\mathbf{R}}\right) \to \sigma^2\left(1+\sqrt{c}\right)^2$ p.s. quand $M \to \infty$.

Le théorème 1, qui est une extension des résultats de [3, 4] au cas $L > 1$, montre en particulier que les plus grandes valeurs propres de $\hat{\mathbf{R}}$ convergent vers des limites dépendant directement des valeurs propres de $\mathbf{\Gamma}$, sous réserve que les ratios $\frac{\gamma_k}{\sigma^2}, k = 1, \ldots, KL$, soient suffisamment grands.

Le résultat suivant présente, sous l'hypothèse $\mathcal{H}_0$, un théorème centrale limite conjoint sur les plus grandes valeurs propres de $\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_L, \hat{\mathbf{R}}$. Il constitue une extension des résultats de [3] pour $L > 1$. Nous le présentons dans le cas $K = 1$.

Théorème 2. *Sous l'hypothèse nulle $\mathcal{H}_0$, et si*

$$\gamma_1 > \sigma^2 \max\left\{\sqrt{c_1}, \ldots, \sqrt{c_L}\right\},$$

on a :

$$\sqrt{M}\begin{pmatrix} \lambda_1\left(\hat{\mathbf{R}}\right) - \phi_c(\gamma_1, \sigma^2) \\ \lambda_1\left(\hat{\mathbf{R}}_1\right) - \phi_{c_1}(\gamma_1, \sigma^2) \\ \vdots \\ \lambda_1\left(\hat{\mathbf{R}}_L\right) - \phi_{c_L}(\gamma_1, \sigma^2) \end{pmatrix} \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}^{L+1}}\left(\mathbf{0}, \mathbf{\Theta}\right),$$

où la matrice de covariance $\mathbf{\Theta}$ a la structure suivante :

$$\mathbf{\Theta} := \begin{pmatrix} \theta_0^2 & \vartheta_1 & \ldots & \vartheta_L \\ \vartheta_1 & \ddots & & \\ \vdots & & \ddots & \\ \vartheta_L & & & \theta_L^2 \end{pmatrix}$$

avec, en notant $c_0 = c$:

$$\theta_\ell^2 = \frac{c_\ell(\gamma_1^2 - \sigma^4 c_\ell)(\gamma_1 + \sigma^2)^2}{\gamma_1^2}, \quad \ell \geq 0$$

$$\vartheta_\ell = \frac{c_0(\gamma_1^2 - \sigma^4 c_\ell)(\gamma_1 + \sigma^2)^2}{\gamma_1^2}, \quad \ell \geq 1.$$

Le théorèmes 1 et 2 sont exploités en section suivante afin de construire une nouvelle statistique de test consistante et à taux de fausse alarme asymptotique constant.

3 Une nouvelle statistique de test

Commençons par reformuler le test (1) à l'aide de la propriété suivante :

Lemme 1. *Les assertions ci-dessous sont équivalentes :*

(1) $\mathbf{\Gamma}_1 = \ldots = \mathbf{\Gamma}_L$

(2) $\forall k \in \{1, \ldots, K\}, \ell \in \{1, \ldots, L\}, \lambda_k\left(\mathbf{\Gamma}_\ell\right) = \lambda_k\left(\mathbf{\Gamma}\right)$.

Le lemme 1 nous permet de réécrire de manière équivalente le test (1) sous la forme :

$$\begin{aligned} \mathcal{H}_0 : &\quad \forall k, \ell,\ \lambda_k\left(\mathbf{\Gamma}_\ell\right) = \lambda_k\left(\mathbf{\Gamma}\right) \\ \mathcal{H}_1 : &\quad \exists k, \ell : \lambda_k\left(\mathbf{\Gamma}_\ell\right) \neq \lambda_k\left(\mathbf{\Gamma}\right) \end{aligned} \tag{7}$$

En conséquence, il est possible de discriminer les hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$ en exploitant uniquement les valeurs propres des matrices $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_L, \mathbf{\Gamma}$, dont nous pouvons par ailleurs construire des estimateurs consistants grâce au Théorème 1.
Considérons l'estimateur $\hat{\sigma}^2$ de σ^2 suivant :

$$\hat{\sigma}^2 := \sum_{\ell=1}^{L} \frac{N_\ell}{N} \sum_{k=1}^{K} \frac{1}{M-K} \sum_{k=K+1}^{M} \lambda_k \left(\hat{\mathbf{R}}_\ell\right).$$

et posons $\hat{\gamma}_k$ la plus grande solution de l'équation $\phi_c\left(\gamma, \hat{\sigma}^2\right) = \lambda_k\left(\hat{\mathbf{R}}\right)$ si $\lambda_k(\hat{\mathbf{R}}) > \hat{\sigma}^2(1+\sqrt{c})^2$ ou $\hat{\gamma}_k = \hat{\sigma}^2\sqrt{c}$ sinon. De même, posons $\hat{\gamma}_{k,\ell}$ la plus grande solution de l'équation $\phi_{c_\ell}\left(\gamma, \hat{\sigma}^2\right) = \lambda_k\left(\hat{\mathbf{R}}_\ell\right)$ si $\lambda_k(\hat{\mathbf{R}}_\ell) > \hat{\sigma}^2(1+\sqrt{c_\ell})^2$ ou $\hat{\gamma}_k = \hat{\sigma}^2\sqrt{c_\ell}$ sinon. Le théorème 1 montre alors que $\hat{\gamma}_k \to \gamma_k$ et $\hat{\gamma}_{k,\ell} \to \gamma_{k,\ell}$ p.s. quand $M \to \infty$, sous réserve que $\gamma_k > \sigma^2\sqrt{c}$ et $\gamma_{k,\ell} > \sigma^2\sqrt{c_\ell}$ pour tout $k = 1, \ldots, K$.

Muni de ces estimateurs, nous proposons dans ce qui suit une nouvelle statistique à taux de fausse alarme asymptotique constant pour le test (7). Pour la suite, on se place dans le cas particulier $K = 1$. Posons $\hat{\boldsymbol{\gamma}} = (\hat{\gamma}_1 - \hat{\gamma}_{1,1}, \ldots, \hat{\gamma}_1 - \hat{\gamma}_{1,L})^T$. En exploitant les théorèmes 1 et 2 avec une delta-méthode, il est possible de montrer que sous l'hypothèse nulle $\mathcal{H}_0$ et sous réserve que $\gamma_1 > \sigma^2 \max\{\sqrt{c_1}, \ldots, \sqrt{c_L}\}$,

$$\sqrt{M}\left(\mathbf{H}\hat{\mathbf{\Omega}}\mathbf{H}^T\right)^{-\frac{1}{2}} \hat{\boldsymbol{\gamma}} \xrightarrow[M\to\infty]{\mathcal{L}} \mathcal{N}_{\mathbb{R}^L}\left(\mathbf{0}, \mathbf{I}\right)$$

où $\mathbf{H}$ est la matrice $L \times (L+1)$ donnée par

$$\mathbf{H} = \begin{pmatrix} 1 & -1 & 0 & \ldots & \ldots & 0 \\ 1 & 0 & -1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ 1 & 0 & \ldots & \ldots & 0 & -1 \end{pmatrix},$$

et où

$$\hat{\mathbf{\Omega}} := \begin{pmatrix} \hat{\omega}_0^2 & \hat{\xi} & \ldots & \hat{\xi} \\ \hat{\xi} & \ddots & & \\ \vdots & & \ddots & \\ \hat{\xi} & & & \hat{\omega}_L^2 \end{pmatrix},$$

avec

$$\hat{\omega}_\ell^2 = \frac{c_\ell \hat{\gamma}_1^2 (\hat{\gamma}_1 + \hat{\sigma}^2)^2}{\hat{\gamma}_1^2 - \hat{\sigma}^4 c_\ell}, \quad \ell \geq 0$$

$$\hat{\xi} = \frac{c_0 \hat{\gamma}_1^2 (\hat{\gamma}_1 + \hat{\sigma}^2)^2}{\hat{\gamma}_1^2 - \hat{\sigma}^4 c_0}, \quad \ell \geq 1.$$

On propose alors la statistique de test suivante

$$T = \mathbb{1}_{]\epsilon, +\infty[}\left(\left\|\left(\mathbf{H}\hat{\mathbf{\Omega}}\mathbf{H}^T\right)^{-\frac{1}{2}} \hat{\boldsymbol{\gamma}}\right\|_2^2\right) \tag{8}$$

où le seuil ϵ peut être ajusté de manière à contrôler la probabilité de fausse alarme asymptotique, comme décrit dans le résultat suivant.

Théorème 3. *Posons $\epsilon = \frac{q_{1-\alpha}}{M}$ où $q_{1-\alpha}$ est le quantile d'ordre $1-\alpha$ de la loi du χ^2 à L degrés de liberté. Sous les hypothèses 1 et 2, et si $\gamma_1 > \sigma^2\sqrt{c}$, alors*

$$\mathbb{P}_0(T = 1) \xrightarrow[M\to\infty]{} \alpha$$

où $\mathbb{P}_0$ désigne la mesure de probabilité induite sous $\mathcal{H}_0$.

En analysant le comportement au 1er ordre de T à l'aide du théorème 1, il est également possible d'obtenir un résultat de consistance. Considérons l'hypothèse suivante, qui sera utilisée en complément de l'hypothèse 2.

Hypothèse 3. *On a $\gamma_1 > \sigma^2 \max\{\sqrt{c_1}, \ldots, \sqrt{c_L}\}$ et pour tout $\ell = 1, \ldots, L$, $\gamma_{1,\ell} > \sigma^2\sqrt{c_\ell}$. De plus, sous l'hypothèse alternative $\mathcal{H}_1$, il existe ℓ tel que $\gamma_1 \neq \gamma_{1,\ell}$.*

On déduit alors, sous les hypothèses 1, 2 et 3,

$$\left\|\left(\mathbf{H}\hat{\mathbf{\Omega}}\mathbf{H}^T\right)^{-\frac{1}{2}} \hat{\boldsymbol{\gamma}}\right\|_2^2 \xrightarrow[M\to\infty]{\text{p.s.}} \begin{cases} 0 & \text{sous } \mathcal{H}_0 \\ \epsilon_1 = \left\|\left(\mathbf{H}\mathbf{\Omega}\mathbf{H}^T\right)^{-\frac{1}{2}} \boldsymbol{\gamma}\right\|_2^2 > 0 & \text{sous } \mathcal{H}_1, \end{cases}$$

où $\boldsymbol{\gamma} = (\gamma_1 - \gamma_{1,1}, \ldots, \gamma_1 - \gamma_{1,L})^T$ et où $\mathbf{\Omega}$ est obtenue en remplaçant $\hat{\gamma}_1$ par γ_1 et $\hat{\sigma}^2$ par σ^2 dans l'expression de $\hat{\mathbf{\Omega}}$. On aboutit alors au résultat suivant.

Théorème 4. *Sous les hypothèses 1, 2 et 3, et $\forall \epsilon \in]0, \epsilon_1[$,*

$$\mathbb{P}_i\left(\lim_{M\to\infty} T = i\right) = 1,$$

où $\mathbb{P}_i$ est la mesure de probabilité sous l'hypothèse $\mathcal{H}_i$.

4 Simulations

Dans cette section, nous illustrons numériquement les performances de la statistique de test T proposée en (8). On considère $L = 2$ et un scénario de changement de sous-espace, où

$$\mathbf{\Gamma}_\ell = \mathbf{a}(\theta_\ell)\mathbf{a}(\theta_\ell)^*$$

avec $\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}(1, \mathrm{e}^{\mathrm{i}\theta} \ldots, \mathrm{e}^{\mathrm{i}(M-1)\theta})^T$. Sous $\mathcal{H}_0$, on considère $\theta_1 = \theta_2 = 0$ et sous $\mathcal{H}_1$, $\theta_1 = 0, \theta_2 = \frac{\pi}{M}$, de sorte que les modèles statistiques sous $\mathcal{H}_0$ et $\mathcal{H}_1$ ne soient pas trivialement séparables asymptotiquement.

Afin d'illustrer le théorème 3, nous traçons en figure 1 la fonction de répartition complémentaire de MT sous $\mathcal{H}_0$ (évaluée sur 50000 tirages), pour $N_1 = N_2 = 2M$, $M \in \{10, 30, 50, 70\}$ et $\sigma^2 = 0.5$. Les tracés montrent une bonne adéquation avec la loi du $\chi^2(L)$ pour des valeurs modestes de M, ce qui montre la fiabilité de l'analyse asymptotique proposée.

Nous illustrons ensuite les performances en termes de probabilité de détection par le biais de courbes ROC (*Receiver Operating Characteristic*). Afin de comparer ces performances, on propose de considérer la statistique de test $T_{\mathrm{GLRT}} = \mathbb{1}_{]\eta, +\infty[}(S)$

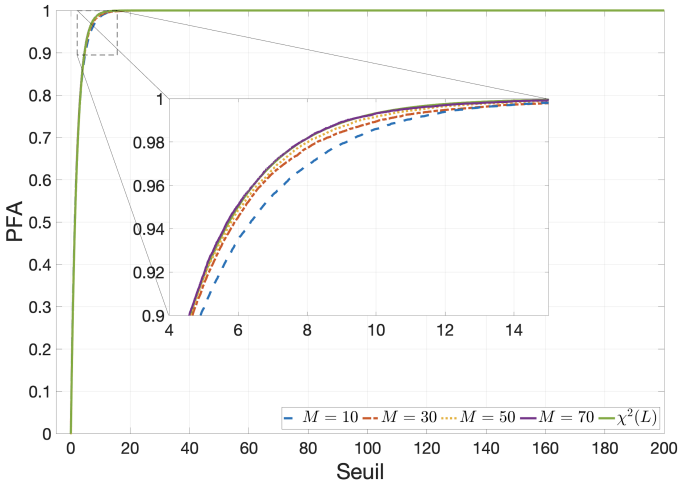

FIGURE 1 – PFA de la statistique de test MT pour $\sigma^2 = 0.5$ $M = \{10, 30, 50, 70\}$ et loi $\left(1 - \chi^2(L)\right)$

, où S est définie en (2) pour $\varphi = \log$, ainsi qu'une statistique notée $T_{\text{FISH}} = \mathbb{1}_{]\kappa,+\infty[}(F)$ et inspirée des travaux de [5] sur les spiked models pour les matrices de Fisher, avec

$$F = \sum_{\substack{\ell,\ell'=1 \\ \ell<\ell'}}^{L} \Bigg[\left(\lambda_1\left(\hat{\mathbf{R}}_\ell^{-1}\hat{\mathbf{R}}_{\ell'}\right) - \nu_{\ell,\ell'}^{+}\right)^{+} + \left(\nu_{\ell,\ell'}^{-} - \lambda_M\left(\hat{\mathbf{R}}_\ell^{-1}\hat{\mathbf{R}}_{\ell'}\right)\right)^{+} \Bigg],$$

où $\nu_{\ell,\ell'}^{\pm} = \left(\frac{1 \pm \sqrt{c_\ell + c_{\ell'} - c_\ell c_{\ell'}}}{1 - c_\ell}\right)^2$. D'après les résultats de [5], il est possible de montrer que la statistique T_{FISH} est consistante dans le régime des grandes dimensions sous réserve que le RSB soit suffisamment grand (les détails sont omis). Les seuils ϵ, η, κ des trois statistiques seront ajustés empiriquement de manière à vérifier une certaine PFA donnée.

La figure 2 montre le tracé des courbes ROC de T, T_{GLRT} et T_{FISH} pour $M = 40$, $N_1 = N_2 = 2M$ et $\sigma^2 \in \{0.1, 0.2, 0.5\}$. On remarque que la statistique T proposée présente des performances significativement meilleures, notamment grâce à une dégradation plus lente des performances à mesure que le RSB diminue.

La Figure 3 présente quant à elle les courbes ROC des trois statistiques, pour $M = 40$, $N_1 = 3M$ et $N_2 = 1.5M$ avec $\sigma^2 \in \{0.1, 0.2, 0.5\}$. On peut noter que, lorsque c_ℓ augmente, γ_ℓ s'approche de la condition détaillée en hypothèse 3 et par conséquent les performances de la statistique de test T sont impactées. Néanmoins, comme nous l'avons observé en figure 2, les performances ainsi que la résistance au bruit de la statistique de test T restent supérieures.

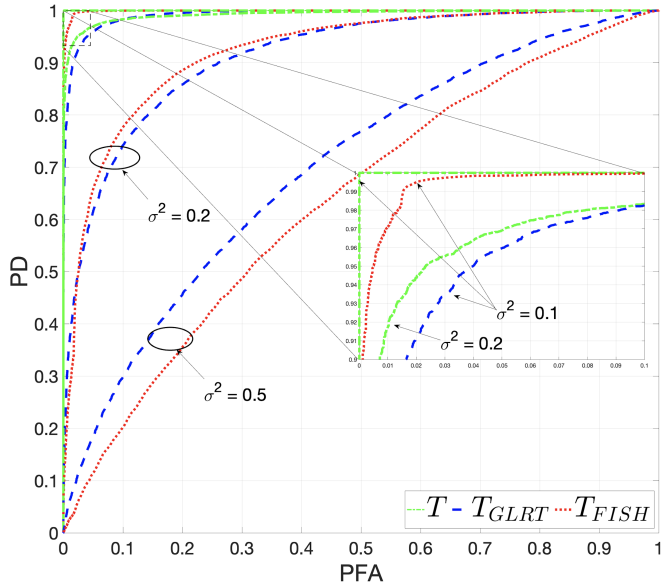

FIGURE 2 – Courbes ROC pour $M = 40$, $N_1 = N_2 = 2M$ et $\sigma^2 \in \{0.1, 0.2, 0.5\}$

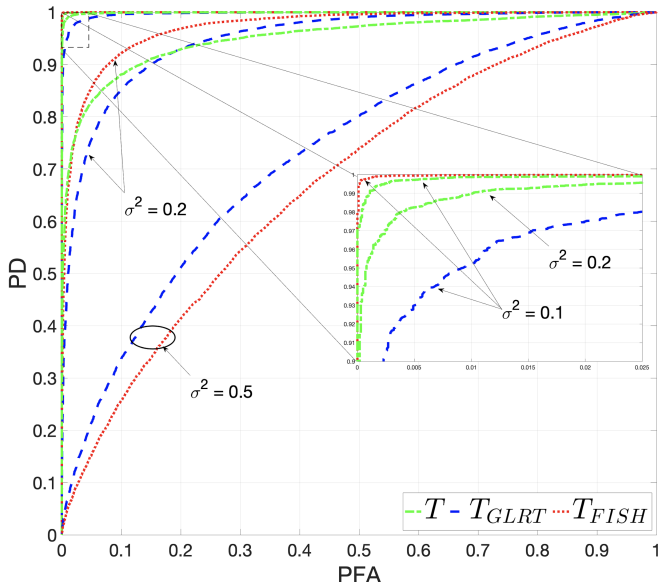

FIGURE 3 – Courbes ROC pour $M = 40$, $N_1 = 3M$, $N_2 = 1.5M$ et $\sigma^2 \in \{0.1, 0.2, 0.5\}$

Références

[1] D. Ciuonzo, V. Carotenuto, and A. De Maio. On multiple covariance equality testing with application to SAR change detection. *IEEE Trans. on Signal Process.*, 65(19) :5078–5091, 2017.

[2] K. Wachter. The limiting empirical measure of multiple discriminant ratios. *Ann. Statist.*, pages 937–957, 1980.

[3] F. Benaych-Georges and R.R. Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.*, 227(1) :494–521, 2011.

[4] J. Baik. G.Ben Arous. S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab. 33 (5) 1643 - 1697*, 2005.

[5] Q. Wang and J. Yao. Extreme eigenvalues of large-dimensional spiked fisher matrices with application. *Ann. Statist.*, 45(1) :415–460, 2017.