On the behaviour of the singular values of empirical autocovariance matrices in the high-dimensional case

Daria TIEPLOVA[1,2], Philippe LOUBATON[1], Leonid PASTUR[2]

[1]Laboratoire d'Informatique Gaspard Monge, UMR 8049
Université Paris-Est Marne la Vallée
5 Bd. Descartes, Champs sur Marne, 77454 Marne la Vallée Cedex 2 France

[2]Dept. of Theoretical Physics, Institute for Low Temperature Physics and Engineering
47 Nauky Ave., Kharkiv, 61103, Ukraine

daria.tieplova@u-pem.fr, philippe.loubaton@u-pem.fr, pastur@ilt.kharkov.ua

Résumé – Dans cet article, nous nous intéressons au comportement des valeurs singulières de l'estimateur empirique de l'autocovariance entre le passé et le futur d'une série temporelle gaussienne multivariable décorrélée temporellement. Nous nous intéressons au régime des grandes dimensions dans lequel la dimension de l'observation est du même ordre de grandeur que la taille de l'échantillon disponible. Alors que dans le régime traditionnel, la matrice d'autocovariance empirique tend vers 0, nous montrons dans notre cadre que les valeurs singulières au carré tendent se répartir selon une distribution déterministe caractérisée par sa transformée de Stieltjes. Nous identifions le support de cette distribution, et montrons que les valeurs singuliès au carré sont situées presque surement au voisinage du support. Ce travail est une première étape permettant d'analyser le comportement d'algorithmes classiques d'estimation de modèles d'état.

Abstract – In this paper, we study the behaviour of the singular values of the empirical estimate of the autocovariance matrix between the future and the past of an uncorrelated multivariate Gaussian time series. We consider the high-dimensional regime in which the dimension of the observation and the sample size both converge towards ∞ at the same rate. While in the traditional regime the estimated autocovariance matrix converges towards 0, we establish in this paper that the empirical distribution of the singular values to the square has a limit deterministic behaviour that is characterized through its Stieltjes transform. We identify the support of this limit distribution, and show that the singular values to the square are located in a neighbourhood of this support. This work is a first step towards an analysis of classical state space estimation algorithms in the high-dimensional regime.

1 Introduction

It is well known that a M–dimensional time series $(u_n)_{n\in\mathbb{Z}}$ with rational spectrum admits causal state space representations. In the mutivariate case, these representations are recognized to be more convienient that the vector-valued ARMA ones because they are more parsimonious. In particular, while the minimal order of an ARMA representation cannot be clearly evaluated, the minimal dimension P of the states that represent u coincides with the rank of the autocovariance matrix $R_{f|p}^{(L)}$ between the augmented vectors $u_{n+L}^L = (u_{n+L}^T, u_{n+L+1}^T, \ldots, u_{n+2L-1}^T)^T$ (the future) and $u_n^L = (u_n^T, u_{n+1}^T, \ldots, u_{n+L-1}^T)^T$ (the past) where L should be chosen greater than P (see e.g. the monograph [7]). Moreover, the left factor of any minimal rank factorization of $R_{f|p}^{(L)}$ coincides with the observability matrix of a pair (C, A) of a minimal state space representation of u, from which it is easy to retrieve matrices C and A. When a noisy version $y_n = u_n + v_n$ is observed, and that the noise $(v_n)_{n\in\mathbb{Z}}$ is temporally uncorrelated, the autocovariance matrix between the future and the past of $(y_n)_{n\in\mathbb{Z}}$ remains equal to $R_{f|p}^{(L)}$. A crucial problem is to estimate C and A from N consecutive samples $y_1, y_2, \ldots, y_N$. For this, a standard approach, known as the principal component algorithm (see e.g. [1] and [7]), consists first in estimating empirically $R_{f|p}^{(L)}$ from the N observations. If the number of samples N is much larger than the dimension ML of the augmented vectors, the empirical estimate $\hat{R}_{f|p}^{(L)}$ is close to $R_{f|p}^{(L)}$, and is nearly a rank P matrix. This allows one to evalute P as well as C and A from the truncated singular decomposition of $\hat{R}_{f|p}^{(L)}$. The approach provides reasonable performance when the ratio $\frac{ML}{N}$ is small enough. When the dimension M of the observation is large, this condition needs the number of sample N to be very large. In a number of contexts, the number of samples is however not unlimited, and the ratio $\frac{ML}{N}$ may not be small enough. It is thus important to be able to analyse the behaviour of the above mentionned estimation procedure when ML and N are of the same order of magnitude. In order to simplify the problem, we assume in this paper that parameter L cannot be too large, and consider the high-dimensional regime where M and N both converge towards $+\infty$ in such a way that M/N converges towards a constant, while parameter L remains fixed. Therefore, in our analysis, parameter M depends on N, and we denote by

c_N the ratio $c_N = \frac{ML}{N}$.

The relevance of the principal component algorithm lies on the observation that, in the classical asymptotic regime where $N \to +\infty$ while M and L remain fixed, $ML \times ML$ matrix $\hat{R}_{f|p}^{(L)}$ is close to a rank P matrix for each $L \geq P$. In this paper, we focus on the study of the behaviour of the singular values of this matrix, or equivalently of the eigenvalues of $\hat{R}_{f|p}^{(L)}(\hat{R}_{f|p}^{(L)})^*$, in the above mentioned high-dimension regime. Due to the lack of space, we only address the case where signal u is absent. Equivalently, we assume from now on that for each $n = 1, \ldots, N$, y_n coincides with v_n. In the standard asymptotic regime, matrix $\hat{R}_{f|p}^{(L)}$ converges towards 0 in the spectral norm sense, but, as shown below, this is no longer the case in our high-dimensional regime where the eigenvalues of $\hat{R}_{f|p}^{(L)}(\hat{R}_{f|p}^{(L)})^*$ are localized in a non vanishing interval depending on c_N and on the eigenvalues of the covariance matrix $R = \mathbb{E}(y_n y_n^*)$. When the signal u is present, it is possible, under some assumptions on u, to use a perturbation approach taking benefit of the results presented here, and to find conditions under which the presence of u produces in matrix $\hat{R}_{f|p}^{(L)}(\hat{R}_{f|p}^{(L)})^*$ some eigenvalues that are larger than when the observation is reduced to the noise. This approach is rather standard in the large random matrix theory literature (see e.g. [2], [4]). The corresponding results will be reported elsewhere.

We finally mention the very few papers that addressed recently the problem considered in this paper. [5] considers the case where the covariance matrix R is reduced to I and where $L = 1$, a quite specific case. However, y is possibly non Gaussian in [5], so that the method of [5] is completely different from the approach developed in the present paper. Moreover, in [5] the analysis of the limit eigenvalue distribution of $\hat{R}_{f|p}^{(L)}(\hat{R}_{f|p}^{(L)})^*$ is less complete than in the present paper. We also notice that [6] took benefit of the results of [5] to consider the presence of a very particular useful signal u. We also mention [3] which derived some of the results of [5] using the moment method.

Acknowledgement: This work is partially funded by the Bézout Labex, funded by ANR, reference ANR-10-LABX-58, and by the ANR Project HIDITSA, reference ANR-17-CE40-0003.

2 Assumptions and notations.

In this section, we will make precise the assumptions and introduce some notations. In the following, we assume that $(y_n)_{n\in\mathbb{Z}}$ is a zero mean M–dimensional complex Gaussian uncorrelated time series with covariance matrix R, i.e. $\mathbb{E}(y_{n+k}y_n^*) = R\,\delta_k$, where R is a $M \times M$ positive definite matrix. As M depends on N, matrix R is thus not a fixed matrix, but has to be interpreted as a sequence of matrices $(R_N)_{N\geq 1}$. The eigenvalues of matrix R_N, arranged in the decreasing order, are denoted $(\lambda_{k,N})_{k=1,\ldots,M}$. In the following, it is assumed that there exist 2 positive constants a and b such that $0 < aI \leq R_N \leq bI$ for each N. In other words, $\inf_N \lambda_{M,N} > 0$ and $\sup_N \lambda_{1,N} < +\infty$. Finally, $c_N = \frac{ML}{N}$ converges towards a non zero constant c_*. In order to simplify the presentation, we assume that $c_N \leq 1$ for each N, which implies that $c_* \leq 1$. We introduce the block-Hankel "past" and "future"matrices $Y_{p,N}$ and $Y_{f,N}$ defined by

$$Y_p^{(L)} = \begin{pmatrix} y_1 & y_2 & \cdots & y_{N-1} & y_N \\ y_2 & y_3 & \cdots & y_N & y_{N+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_L & y_{L+1} & \cdots & y_{N+L-2} & y_{N+L-1} \end{pmatrix} \quad (1)$$

and

$$Y_f^{(L)} = \begin{pmatrix} y_{L+1} & y_{L+2} & \cdots & y_{N-1+L} & y_{N+L} \\ y_{L+2} & y_{L+3} & \cdots & y_{N+L} & y_{N+L+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{2L} & y_{2L+1} & \cdots & y_{N+2L-2} & y_{N+2L-1} \end{pmatrix} \quad (2)$$

We notice that both matrices depend on samples $y_{N+1}, \ldots, y_{N+2L-1}$ that are not supposed to be available. In order to simplify, we do not take into account the corresponding end effects. We also consider the normalized past and future matrices $\Sigma_{p,N} = \frac{1}{\sqrt{N}} Y_{p,N}$ and $\Sigma_{f,N} = \frac{1}{\sqrt{N}} Y_{f,N}$. Matrix $\Sigma_{f,N}\Sigma_{p,N}^*$ is the standard empirical estimate $\hat{R}_{f|p}^{(L)}$ of the autocovariance matrix $R_{f|p}^{(L)}$ between vectors $y_{n+L}^L = (y_{n+L}^T, \ldots, y_{n+2L-1}^T)^T$ and $y_n^L = (y_n^T, \ldots, y_{n+L-1}^T)^T$, which, in the present context is $R_{f|p}^{(L)} = 0$. We denote by $\mathbb{C}^+$ the set of all complex numbers such that $\mathrm{Im}(z) > 0$. Finally, if α is a positive measure whose support is $\mathcal{S}$, its Stieltjes transform $s_\alpha(z)$ is the function holomorphic on $\mathbb{C} - \mathcal{S}$ defined by $s_\alpha(z) = \int_{\mathcal{S}} \frac{d\alpha(\lambda)}{\lambda - z}$.

3 The main results.

We denote by $(\hat{\lambda}_{k,N})_{k=1,\ldots,M}$ the eigenvalues of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ arranged in the decreasing order. The empirical eigenvalue distribution of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ is the probability measure $\hat{\nu}_N = \frac{1}{M}\sum_{k=1}^M \delta_{\hat{\lambda}_{k,N}}$. As in a number of large random matrix models, measure $\hat{\nu}_N$ has a deterministic behaviour, i.e. there exists a sequence of deterministic probability measures $(\nu_N)_{N\geq 1}$ carried by $\mathbb{R}^+$ such that $\hat{\nu}_N - \nu_N \to 0$ weakly almost surely, a condition which is equivalent to the almost sure convergence of $\frac{1}{M}\sum_{k=1}^M \psi(\hat{\lambda}_{k,N}) - \int \psi(\lambda) d\nu_N(\lambda)$ towards 0 for each bounded continuous function ψ. Intuitively, this means that if N is large enough, then the histograms of the eigenvalues of various realizations of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ tend to accumulate around the graph of the probability density of ν_N. Measure ν_N is characterized by its Stieltjes transform, denoted $s_N(z)$. The following theorem holds.

Theorem 1 *For each $z \in \mathbb{C}^+$, the equation*

$$t_N(z) = \frac{1}{M}\mathrm{Tr} R_N \left(-z \left(I_M + \frac{c_N t(z)}{1 - z(c_N t_N(z))^2} R_N \right) \right)^{-1} \quad (3)$$

has a unique solution that belongs to $\mathbb{C}^+$. Moreover, function $z \to t_N(z)$ is the Stieltjes transform of a measure μ_N carried by $\mathbb{R}^+$. The $M \times M$ matrix-valued function $T_N(z) = \left(-z \left(I_M + \frac{c_N t(z)}{1-z(c_N t_N(z))^2} R_N \right) \right)^{-1}$ is the Stieltjes transform of a positive matrix-valued measure ν_N^T such that $\nu_N^T(\mathbb{R}^+) = I_M$, and $s_N(z) = \frac{1}{M}\mathrm{Tr}(T_N(z))$ coincides with the Stieltjes transform of the probability distribution $\nu_N = \frac{1}{M}\mathrm{Tr}(\nu_N^T)$. Moroever, $\hat{\nu}_N - \nu_N \to 0$ weakly almost surely when $N \to +\infty$. Finally, if we denote $\mathcal{S}_N$ the support of measure ν_N, the following holds: Assume that there exists a positive quantity $\epsilon > 0$, two real values $e, f \in \mathbb{R}$ and an integer N_0 such that

$$(e - \epsilon,\ f + \epsilon) \cap \mathcal{S}_N = \varnothing \qquad \forall N \in \mathbb{N},\ N \geq N_0. \quad (4)$$

Then, almost surely, no eigenvalue of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ appears in $[e,\ f]$ for all N large enough.

It is interesting to notice that $t_N(z)$, and therefore $s_N(z), \mu_N$ and ν_N do not depend on parameter L. Moreover, Theorem 1 shows that in order to understand the behaviour of the eigenvalues of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ when $N \to +\infty$, it is of great interest to study measure ν_N. In particular, the last statement of Theorem 1 implies that the characterization of $\mathcal{S}_N$ allows one to infer useful informations on the localization of the eigenvalues.

$\mathcal{S}_N$ can indeed be characterized. For this, we first mention that measures ν_N and μ_N are absolutely continuous one with respect to the other. Therefore, $\mathcal{S}_N$ coincides with the support of μ_N. We first state the following preliminary result that shows that μ_N (and thus ν_N) is absolutely continuous w.r.t. the Lebesgue measure.

Proposition 1 *For each $x \in \mathbb{R}^*$, $\lim_{z \to x, z \in \mathbb{C}^+} t_N(z) = t_N(x)$ exists and is finite. Moreover, μ_N and ν_N are absolutely continuous w.r.t. the Lebesgue measure, and their densities coincide with functions $f_N(x)$ and $g_N(x)$ defined for each $x \neq 0$ by $f_N(x) = \frac{1}{\pi}\mathrm{Im}(t_N(x))$ and $g_N(x) = -\frac{1}{\pi}\frac{c_N \mathrm{Im}((t_N(x))^2)}{|1-x(c_N t_N(x))^2|^2}$ respectively. 0 belongs to $\mathcal{S}_N$, and $\lim_{x\to 0} f_N(x) = \lim_{x \to 0} g_N(x) = +\infty$. Finally, if $c_N < 1$, it holds that*

$$g_N(x) \simeq_{x \to 0} \frac{1}{\pi} \frac{1}{\sqrt{c_N\,(1 - c_N)}} \frac{1}{M}\mathrm{Tr}(R_N^{-1}) \frac{1}{\sqrt{x}} \quad (5)$$

while if $c_N = 1$,

$$g_N(x) \simeq_{x \to 0} \frac{1}{\pi} \frac{\sqrt{3}}{2} \left(\frac{1}{M}\mathrm{Tr}(R_N^{-1}) \right)^{2/3} \frac{1}{x^{2/3}} \quad (6)$$

This result shows that 0 is a singular point of the densities of μ_N and ν_N. This, in practice, means that a number of eigenvalues of matrix $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ are close from 0. Moreover, the rate of convergence of g_N towards $+\infty$ is higher if $c_N = 1$, showing that in this case, the proportion of eigenvalues close to 0 is even larger than if $c_N < 1$. While we do not address the case $c_N > 1$ in this paper, we notice that if $c_N > 1$, then measures μ_N and ν_N both contain a Dirac mass with weight $1 - \frac{1}{c_N}$. Moreover, the densities $f_N(x)$ and $g_N(x)$ are zero in a neighborhood of 0. Fig. 1 illustrates Theorem 1 and Proposition 1. Here, $M = 500$, $N = 1500$ and $L = 2$ so that $c_N = 2/3$. The eigenvalues of matrix R_N are defined by $\lambda_{k,N} = 1/2 + \frac{\pi}{4}\cos\left(\frac{\pi(k-1)}{2M}\right)$ for $k = 1, \ldots, M$. Matrix R_N verifies $\frac{1}{M}\mathrm{Tr}(R_N) \simeq 1$. The histogram of the eigenvalues of a realization of $\Sigma_{f,N}\Sigma_{p,N}^*\Sigma_{p,N}\Sigma_{f,N}^*$ is represented as well as the graph of the density $g_N(x)$.

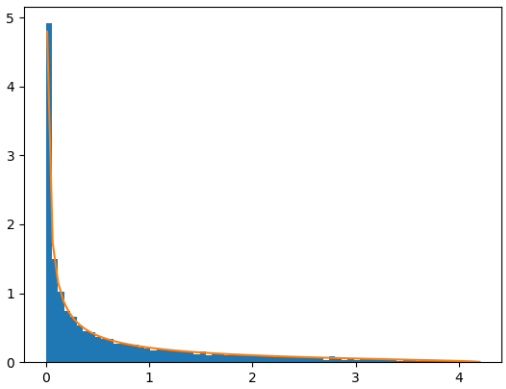

Figure 1: Histogram of the eigenvalues and graph of $g_N(x)$ for $M = 500, N = 1500, L = 2$

In order to characterize $\mathcal{S}_N$ more explicitely, we introduce function $w_N(z)$ defined by $w_N(z) = z c_N t_N(z) - \frac{1}{c_N t_N(z)}$, and remark that (3) leads immediately to $\phi_N(w_N(z)) = z$ for each z where $\phi_N(w)$ is the function defined

$$\phi(w) = cw^2 \frac{1}{M}\mathrm{Tr} R(R - wI)^{-1}(c\,\frac{1}{M}\mathrm{Tr} R(R - wI)^{-1} - 1)$$

Moreover, it can be checked that the interior $\mathcal{S}_N^\circ$ of $\mathcal{S}_N$ coincides with $\mathcal{S}_N^\circ = \{x \in \mathbb{R}^{+*}, \mathrm{Im}(w_N(x)) > 0\}$. Therefore, if $x \in \mathbb{R}^+ - \mathcal{S}_N$, $w_N(x)$ is real, and thus coincides with one of the real solutions of the equation $\phi_N(w) = x$. It is in fact possible to establish that if $x \in \mathbb{R}^+ - \mathcal{S}_N$, then, the equation $\phi_N(w) = x$ admit a unique real solution verifying $\phi_N'(w) > 0$ and $\frac{1}{M}\mathrm{Tr}(R(R - wI)^{-1}) < 0$ and that this real solution is precisely $w_N(x)$. Therefore, $\mathbb{R}^+ - \mathcal{S}_N$ coincides with the image by function ϕ_N of the intervals on which $\phi_N'(w) > 0$ and $\frac{1}{M}\mathrm{Tr}(R(R - wI)^{-1}) < 0$. It is easy to check that function $w \to \phi_N(w)$ converges towards $+\infty$ when $w \to \lambda_{1,N}, w > \lambda_{1,N}$, decreases until the point $w_{+,N}$ defined as the unique solution of equation $\phi_N'(w) = 0$ on the interval $]\lambda_{1,N}, +\infty[$, and then increases on $]w_{+,N}, +\infty[$. It is clear that $\frac{1}{M}\mathrm{Tr}(R(R - wI)^{-1}) < 0$ if $w \geq w_{+,N}$. Therefore, if we denote by $x_{+,N}$ the point $x_{+,N} = \phi_N(w_{+,N})$, then, the image by ϕ_N of $]w_{+,N}, +\infty[$, i.e. the interval $]x_{+,N}, +\infty[$, is included into $\mathbb{R}^+ - \mathcal{S}_N$. Depending on the existence of local extrema of ϕ_N located in $[0, x_{+,N}]$, satisfying $\frac{1}{M}\mathrm{Tr}(R(R - wI)^{-1}) < 0$,

the support may be equal to the whole interval $[0, x_{+,N}]$, or to a finite union of intervals included into $[0, x_{+,N}]$. In this last case, the 2 extreme intervals contain 0 and $x_{+,N}$ respectively. In fig. 2 and 3, we represent function ϕ_N when the number of distinct eigenvalues $\overline{M}$ of R_N is equal to 3, and denote $(\overline{\lambda}_i)_{i=1,\ldots,\overline{M}}$ the corresponding eigenvalues. The $(\omega_i)_{i=1,2,3}$ are the solutions of $c\,\frac{1}{M}\mathrm{Tr}R(R-wI)^{-1}-1=0$, and μ_1, μ_2 are the 2 solutions of the equation $\frac{1}{M}\mathrm{Tr}R(R-wI)^{-1}=0$. The intervals on which $\frac{1}{M}\mathrm{Tr}(R(R-wI)^{-1}) < 0$ are thus $]\overline{\lambda}_1, \mu_1[,]\overline{\lambda}_2, \mu_2[$ and $]\overline{\lambda}_3, +\infty[$. In Fig. 2, ϕ_N is decreasing on $]\overline{\lambda}_1, \mu_1[$ and on $]\overline{\lambda}_2, \mu_2[$, so that $\mathcal{S}_N = [0, x_+]$. In Fig. 3, ϕ_N is this time increasing on an interval $]x_{1,-}, x_{1,+}[$ included into $]\overline{\lambda}_1, \mu_1[$. Therefore, $\mathcal{S}_N$ coincides with $[0, x_{1,-}] \cup [x_{1,+}, x_+]$.

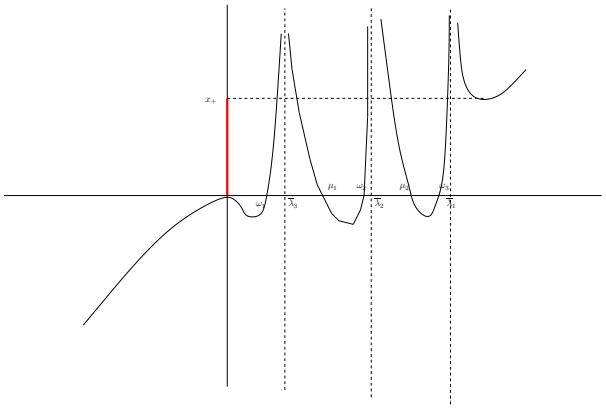

Figure 2: Typical representation of $\phi(w)$ as a function of w for $\overline{M} = 3$. There is no local maximum on $[\overline{\lambda}_1, \mu_1]$ and on $[\overline{\lambda}_2, \mu_2]$, so that $\mathcal{S} = [0, x_+]$.

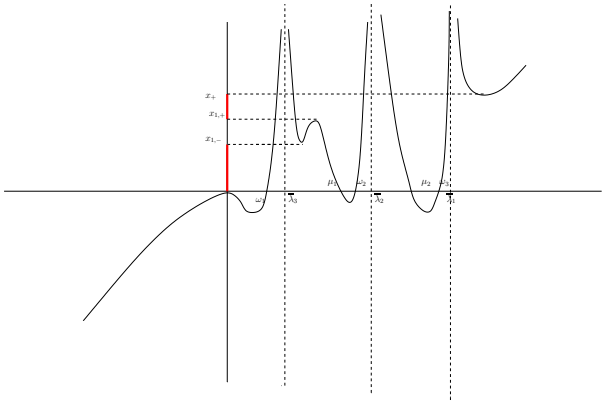

Figure 3: Typical representation of $\phi(w)$ as a function of w for $\overline{M} = 3$. There are 2 local extrema on $[\overline{\lambda}_1, \mu_1]$ and no local extremum on $[\overline{\lambda}_2, \mu_2]$, so that $\mathcal{S} = [0, x_{1,-}] \cup [x_{1,+}, x_+]$.

Intuitively, if the eigenvalues of R_N are close enough one to each other, it can be reasonnably expected that ϕ_N will not admit local maxima in $]\lambda_{M,N}, \lambda_{1,N}[$ such that $\frac{1}{M}\mathrm{Tr}R(R - w(z)I)^{-1} < 0$. Therefore, in such a context, the support is likely to be reduced to the single interval $[0, x_{+,N}]$. This intuition is confirmed by the following result.

Proposition 2 *Assume that there exist $C > 0$ and $\epsilon > 0$ such that for each sufficiently large N, the following condition*

$$|\lambda_{k,N} - \lambda_{l,N}| \leq C \left(\frac{|k-l|}{M} \right)^{(1+\epsilon)/2} \tag{7}$$

holds for each pair (k, l), $1 \leq k \leq l \leq M$. Then, for each M large enough, $\mathcal{S} = [0, x_{+,N}]$.

This result appears quite useful to address the case where a signal u is present. More details will be provided in a forthcoming paper.

4 Conclusion

In this paper, we have studied the behaviour of the eigenvalues of matrix $\hat{R}^{(L)}_{f|p}\hat{R}^{*(L)}_{f|p}$ where $\hat{R}^{(L)}_{f|p}$ represents the usual empirical estimate of the autocovariance matrix between the future and the past of an uncorrelated complex Gaussian time series. In the usual high-dimensional regime, we have established that the empirical eigenvalue distribution of $\hat{R}^{(L)}_{f|p}\hat{R}^{*(L)}_{f|p}$ behaves as a deterministic probability measure ν_N that is characterized by its Stieltjes transform. The support $\mathcal{S}_N$ of ν_N has been chacterized, and we have proved that almost surely, for N large enough, all the eigenvalues of $\hat{R}^{(L)}_{f|p}\hat{R}^{*(L)}_{f|p}$ are located in a neighbourhood of $\mathcal{S}_N$. We have finally given a condition on the covariance matrix of the observation under which $\mathcal{S}_N$ is reduced to a single interval.

References

[1] K.S. Arun, S.Y. Kung, "Balanced approximation of stochastic systems", SIAM J. Matrix Anal. Appl., 11, pp. 42-68, 1990

[2] F. Benaych-Georges and R. Nadakuditi, "The singular values and vectors of low rank perturbations of large rectangular random matrices", J. Mult. Analysis, vol. 111, pp. 120–135, 2012.

[3] A. Bose, M. Bhattacharjee, "Large Covariance and Autocovariance Matrices", CRC Press, July 2018.

[4] F. Chapon, R. Couillet, W. Hachem and X. Mestre, "The outliers among the singular values of large rectangular random matrices with additive fixed rank deformation", Markov Processes and Related Fields, 20, 183–228 (2014).

[5] Z. Li, G. Pan, J. Yao, "On singular value distribution of large-dimensional autocovariance matrices", J. Multivariate Analysis, vol. 137, pp. 119-140, May 2015.

[6] Z. Li, Q. Wang, J. Yao, "Identifying the number of factors from singular values of a large sample auto-covariance matrix", The Annals of Statistics 45 (1), 257-288 (2017).

[7] P. Van Overschee, B. de Moor, "Subspace Identification for Linear Systems: Theory, Implementation, Applications", Kluwer Academic Publishers, 1996.