Circulant singular spectrum analysis: A new automated procedure for signal extraction☆

Juan Bógalo [a,*], Pilar Poncela [a,b], Eva Senra [c]

[a] *Universidad Autónoma de Madrid, Spain*
[b] *UC3M-Santander Big Data Institute (IBiDat), Univ. Carlos III de Madrid, Spain*
[c] *Universidad de Alcalá, Alcalá de Henares, Spain*

ABSTRACT

Sometimes, it is of interest to single out the fluctuations associated to a given frequency. We propose a new variant of SSA, Circulant SSA (CiSSA), that allows to extract the signal associated to any frequency specified beforehand. This is a novelty when compared with other SSA procedures that need to identify ex-post the frequencies associated to the extracted signals. We prove that CiSSA is asymptotically equivalent to these alternative procedures although with the advantage of avoiding the need of the subsequent frequency identification. We check its good performance and compare it to alternative SSA methods through several simulations for linear and nonlinear time series. We also prove its validity in the nonstationary case. We apply CiSSA in two different fields to show how it works with real data and find that it behaves successfully in both applications. Finally, we compare the performance of CiSSA with other state of the art techniques used for nonlinear and nonstationary signals with amplitude and frequency varying in time.

1. Introduction

Singular Spectrum Analysis (SSA) is a nonparametric procedure based on subspace algorithms for signal extraction [1]. The main task in SSA is to extract the underlying signals of a time series like the trend, cycle, seasonal and irregular components. It has been applied to a wide range of time series problems, besides signal processing [2], like forecasting [3], missing value imputation [4] or functional time series [5] among others. SSA builds a trajectory matrix by putting together lagged pieces of the original time series and works with the Singular Value Decomposition of this matrix. It can be viewed as applying Principal Component (PC) analysis to the columns of the trajectory matrix.

SSA has been applied in different disciplines as several authors illustrate (see [6] and the references therein). For instance, there are recent applications in biometrics [7], climatology [8], energy [9] or volcanic activity [10].

In business and economics, SSA applications focus on forecasting and business cycle analysis [11]. Applications in this field range from analyzing the effect of the 2008 recession in forecasting [12,13], to predicting inflation dynamics [14] or the industrial production with multivariate SSA [15]. Related to the business cycle, SSA has also been used to track the US cycle [16], to analyze the real time nowcasting of the output gap [17] and the economic cycles and their synchronization in three European countries [18]. SSA has also been applied to estimate stochastic volatility models [19] and intraday data forecasting [20].

The common practice when applying SSA is to extract the principal components of the trajectory matrix and to identify afterwards the frequencies associated to them, by analyzing their estimated periodogram [17,21,22] or frequency response [23,24] just to cite a few methods. Though there are fast computing algorithms for the eigenvalues and eigenvectors of Toeplitz matrices [25,26], the use of circulant matrices has a great advantage as their eigenvalues and eigenvectors have a closed form. Circulant matrices have also been used in a different context, within the MUSIC algorithm, restricted to signals that are approximately periodic and deterministic [27].

* Corresponding author.
E-mail addresses: juan.bogalo@telefonica.net (J. Bógalo), pilar.poncela@uam.es (P. Poncela), eva.senra@uah.es (E. Senra).

We propose a new SSA methodology (CiSSA), that can be applied to any time series signal, based on circulant matrices that, once the user has decided beforehand the frequency of interest, it automatically matches this frequency with specific principal components. Circulant matrices become relevant in this setup, as their eigenstructure can be obtained as a function of the frequency and, therefore, we can automatically match their eigenvalues and eigenvectors with any particular frequency. Our approach, CiSSA, valid in a general setting, automatically identifies the eigenvalues and eigenvectors associated to any particular frequency using circulant matrices. Moreover, we obtain an easy way to evaluate the power spectral density since the eigenvalues approximate it at the matched frequencies.

CiSSA seems to perform and compare well with previous versions of SSA, like Basic or Toeplitz SSA, despite introducing its automatization. In order to show this, first, we have proved that CiSSA is asymptotically equivalent to these alternative procedures. Second, we have checked its performance in practice through several sets of simulations for linear and nonlinear models. Finally, we have extended its validity for nonstationary time series. Although SSA has been successfully used in nonstationary time series previously, e.g., [28], our value added is that we apply it in an automated way and also provide a theoretical background overcoming the assumption of stationarity.

In summary, our contribution is to propose a new version of SSA, Circulant SSA, for signal extraction in an automated way valid for any type of signal. With this new version, we make heavy use of circulant matrices and obtain reliable components associated to any pre-specified frequency, both for stationary and nonstationary time series.

We illustrate this new procedure by applying it to the Industrial Production Index (IP) of six developed countries and to the signal produced by the word "Alleluia". IP is a relevant indicator to track the business cycle and its seasonally adjusted signal is followed in real time to monitor the economy. We check that our estimated cycles match the official dating of recessions provided by the OECD and check the strong separability of the estimated components. Regarding the application to speech processing, we find that CiSSA identifies and reproduces the main characteristics of the word under study.

Finally, we check the validity of CiSSA to represent other nonlinear and nonstationary signals in the form of varying amplitude or frequency along time that are frequent in other fields. In order to do so, first, we apply CiSSA to a synthetic example previously used in the literature of AM-FM multicomponent models and, second, we compare the performance of CiSSA and various state of the art techniques [29–31] when applied to a real data set.

The structure of this paper is as follows: Section 2 briefly describes SSA. Section 3 proposes our new SSA procedure, named after Circulant SSA, proves its asymptotic equivalence to Basic and Toeplitz SSA and extends its use for nonstationary time series. Section 4 presents a set of simulations to check the properties of the proposed methodology. Section 5 applies it to the estimation of the business cycle of the industrial production index in six countries and to the signal produced by the word "Alleluia". Section 6 compares CiSSA with other models applied to nonlinear and nonstationary signals with varying amplitude and frequency. Finally, Section 7 concludes.

2. SSA Methodology

The origin of SSA dates back to 1986 with the publication of the papers by Broomhead and King [32,33] and Fraedrich [34]. In 1989, Vautard and Ghil [35] introduce Toeplitz SSA for stationary time series and, three years later, Vautard et al. [22] derive the algorithm called diagonal averaging to obtain the extracted compo-

nents with the length of the original series. At the same time, and independently, the so-called Caterpillar technique was developed in the former Soviet Union [36]. As pointed out by Golyandina and Zhigljavsky [1], SSA is also related to subspace methods as ESPRIT, MUSIC or Min-Norm, and all the literature that started with the seminal work of Pisarenko [37]. See, for instance Ortigueira and Lagunas [38] that compare the eigendecomposition procedures applied to second moments of the data versus the singular value decomposition of the data matrix, being the two methods used in different versions of SSA.

In this section we briefly describe the steps used in SSA to decompose a time series in its unobserved components (trend, cycles,...). Basically, SSA is a technique in two stages: decomposition and reconstruction. In the first stage, decomposition, we transform the original vector of data into a related trajectory matrix and perform its singular value decomposition to obtain the so called elementary matrices. This corresponds to steps 1 and 2 in the algorithm. In the second stage, reconstruction, (steps 3 and 4 of the algorithm) we classify the elementary matrices into disjoint groups associating each group to an unobserved component (trend, cycles,...). Finally, we transform every group into an unobserved component of the same size of the original time series by diagonal averaging.

To proceed with the algorithm, let $\{x_t\}$ denote a stochastic process $t \in \mathcal{T}$ and let $\{x_t\}_{t=1}^{T}$ be a realization[1] of x_t of length T, $\mathbf{x} = (x_1, \ldots, x_T)'$, where the prime denotes transpose and L a positive integer, called the window length, such that $1 < L < T/2$. The Basic SSA or Broomhead-King (BK) procedure involves the following 4 steps:

1st step: Embedding. From the original time series we will obtain an $L \times N$ trajectory matrix $\mathbf{X}$, $N = T - L + 1$, as follows

$$\mathbf{X} = (\mathbf{x}_1 | \ldots | \mathbf{x}_N) = \begin{pmatrix} x_1 & x_2 & x_3 & \ldots & x_N \\ x_2 & x_3 & x_4 & \ldots & x_{N+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_L & x_{L+1} & x_{L+2} & \ldots & x_T \end{pmatrix} \quad (1)$$

where $\mathbf{x}_j = (x_j, \ldots, x_{j+L-1})'$ indicates the $L \times 1$ vector with origin at time j. Notice that the trajectory matrix $\mathbf{X}$ is Hankel and both, by columns and rows, we obtain subseries of the original one.

2nd step: Decomposition. In this step, we perform the singular value decomposition (SVD) of the trajectory matrix $\mathbf{X} = \mathbf{U}\mathbf{D}^{1/2}\mathbf{V}'$ where $\mathbf{U}$ is the $L \times L$ matrix whose columns $\mathbf{u}_k$ are the $L \times 1$ eigenvectors of the second moment matrix $\mathbf{S} = \mathbf{X}\mathbf{X}'$, $\mathbf{D} = diag(\tau_1, \ldots, \tau_L)$, $\tau_1 \geq \ldots \geq \tau_L \geq 0$, are the eigenvalues of $\mathbf{S}$ and $\mathbf{V}$ is the $N \times L$ matrix whose L columns $\mathbf{v}_k$ are the $N \times 1$ eigenvectors of $\mathbf{X}'\mathbf{X}$ associated to nonzero eigenvalues. This decomposition allows to write $\mathbf{X}$ as the sum of the so-called elementary matrices $\mathbf{X}_k$ of rank 1 or dyads,

$$\mathbf{X} = \sum_{k=1}^{r} \mathbf{X}_k = \sum_{k=1}^{r} \mathbf{u}_k \mathbf{w}_k',$$

where $\mathbf{w}_k = \mathbf{X}'\mathbf{u}_k = \sqrt{\tau_k}\mathbf{v}_k$, being $\sqrt{\tau_k}$ the singular values of the $\mathbf{X}$ matrix, and $r = \max_{\tau_k>0}\{k\} = \mathrm{rank}(\mathbf{X})$.

3rd step: Grouping. Under the assumption of weak separability given in [39], we group the elementary matrices $\mathbf{X}_k$ into G disjoint groups summing up the matrices within each group. Let $I_j, j = 1, \ldots, G$ be each disjoint group of indexes associated to the corresponding eigenvectors. The matrix $\mathbf{X}_{I_j} = \sum_{k \in I_j} \mathbf{X}_k$ is associated to the I_j group. The decomposition of the trajectory matrix into

---

[1] For simplicity, we use the same notation for the stochastic process and for the observed time series. It will be clear from the context if we are referring to the population or to the sample. If it were not, we would explicitly clarify it in the main text.

these groups is given by $\mathbf{X} = \mathbf{X}_{I_1} + \ldots + \mathbf{X}_{I_G}$. The contribution of the component coming from matrix $\mathbf{X}_{I_j}$ is given by $\sum_{k \in I_j} \tau_k / \sum_{k=1}^r \tau_k$.

4th step: Reconstruction. Let $\mathbf{X}_{I_j} = (\widetilde{x}_{ij})$. In this step, each matrix $\mathbf{X}_{I_j}$ is transformed into a new time series of the same length T as the original one, denoted as $\widetilde{\mathbf{x}}^{(j)} = (\widetilde{x}_1^{(j)}, \ldots, \widetilde{x}_T^{(j)})'$ by diagonal averaging. This is equivalent to averaging the elements of $\mathbf{X}_{I_j}$ over its antidiagonals, that is, the hankelization of this matrix with the operator $H(\cdot)$ as follows

$$\widetilde{x}_t^{(j)} = H(\mathbf{X}_{I_j}) = \begin{cases} \frac{1}{t} \sum_{i=1}^t \widetilde{x}_{i,t-i+1}, & 1 \leq t < L \\ \frac{1}{L} \sum_{i=1}^L \widetilde{x}_{i,t-i+1}, & L \leq t \leq N \\ \frac{1}{T-t+1} \sum_{i=L-N+1}^{T-N+1} \widetilde{x}_{i,t-i+1}, & N < t \leq T \end{cases}$$

The alternative Toeplitz SSA or Vautard-Ghil (VG) relies on the assumption that $\mathbf{x}$ is stationary and zero mean and it performs the orthogonal diagonalization in step 2 from an alternative matrix $\mathbf{S}_T = (s_{ij})$ where

$$s_{ij} = \frac{1}{T - |i-j|} \sum_{m=1}^{T-|i-j|} x_m x_{m+|i-j|}, \qquad 1 \leq i, j \leq L. \tag{2}$$

In this case, the matrix $\mathbf{S}_T$ is the sample lagged variance-covariance matrix of the original series, a symmetric Toeplitz matrix. The set $(\tau_k, \mathbf{u}_k, \mathbf{w}_k)$ is named the k-th eigentriple. The rest of the algorithm remains unchanged.

3. Circulant SSA

SSA in any of its variants requires to identify the harmonic frequencies of the extracted components and this makes necessary the analysis of the periodogram. To try to automate SSA, several strategies have been proposed such as finding the correlations at different lags between the elements of two eigenvectors, associated to almost identical eigenvalues to test if they are in quadrature [40]; testing if a pair of eigenvectors are associated to the same harmonic based on the periodogram [22]; grouping eigenvectors linked to nearby frequencies in order to assign them to the same harmonic by the introduction of optimal thresholds [21,41]; performing a spectral-based Fisher g test to asses certain principal components to the business cycle frequency [17]; considering eigenvectors as filters [23] and grouping the outputs according to their frequency response [24]; and even applying cluster techniques for grouping the elementary components based on k-means [42] or hierarchical clustering [43]. Nevertheless, whatever procedure is used, the grouping of frequencies is made after the elementary components are extracted. Since the pairs of eigenvalues and eigenvectors are obtained, not as a function of the frequency, but rather on a decreasing magnitude, this means that the grouping is done with uncertainty. A partial solution is provided by computing the eigenvalues-eigenvectors as functions of the frequency for symmetric positive definite Toeplitz matrices [10]. However, the analytic form of the eigenvalues for this type of matrices is only known for heptadiagonal matrices [44]. We generalize the link between the eigenstructure of a matrix and the associated frequencies by the use of circulant matrices allowing non-periodic signals.

In this section, we propose an automated version of SSA based on circulant matrices. First, we deal with the stationary case and, later on, we will extend our proposal to the nonstationary case.

3.1. Stationary case

In this subsection we propose to apply SSA to an alternative matrix of second moments that is circulant. In this case, we have closed solutions form eigenvalues-eigenvectors that are linked to

the desirable specific frequencies. We show the asymptotic equivalence between the traditional Toeplitz matrices used in SSA and our proposed circulant matrices. Based on all the previous results we propose a new alghorithm that we name Circulant SSA (CiSSA).

Toeplitz matrices appear when considering the population second order moments of the trajectory matrix. Let $\{x_t\}$ be an infinite, zero mean stationary time series whose autocovariances are given by $\gamma_m = E(x_t x_{t-m})$, $m = 0, 1, \ldots$ and its power spectral density function, a real continuous and 2π-periodic function, denoted by f. Let

$$\boldsymbol{\Gamma}_L(f) = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{L-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{L-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{L-1} & \gamma_{L-2} & \gamma_{L-3} & \cdots & \gamma_0 \end{pmatrix} \tag{3}$$

be the $L \times L$ matrix that collects these second moments. Notice that $\boldsymbol{\Gamma}_L(f)$ is a symmetric Toeplitz matrix that depends on the power spectral density f through the covariances γ_m. Recall that $\gamma_m = \int_0^1 f(w) \exp(i2\pi mw) dw$ for any integer m where $w \in [0, 1]$ is the frequency in cycles per unit of time.

Analytic expressions for the eigenvalues of Toeplitz matrices are only known up to heptadiagonal matrices. To be able to have closed solutions of the eigenvalues and eigenvectors for any dimension, we use a special case of Toeplitz matrices that are the circulant ones. In a circulant matrix every row is a right cyclic shift of the row above as follows:

$$\mathbf{C}_L(f) = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{L-1} \\ c_{L-1} & c_0 & c_1 & \cdots & c_{L-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{pmatrix}.$$

The eigenvalues and eigenvectors of a circulant matrix have a closed form [45]. The k-th eigenvalue of the $L \times L$ circulant matrix $\mathbf{C}_L(f)$ is given by

$$\lambda_{L,k} = \sum_{m=0}^{L-1} c_m \exp\left(i2\pi m \frac{k-1}{L}\right) \tag{4}$$

for $k = 1, \ldots, L$ and its associated eigenvector can be written as

$$\mathbf{u}_k = L^{-1/2} (u_{k,1}, \ldots, u_{k,L})' \tag{5}$$

where $u_{k,j} = \exp\left(-i2\pi (j-1)\frac{k-1}{L}\right)$.

In particular, if we consider the circulant matrix of order $L \times L$ with elements c_m defined as:

$$c_m = \frac{1}{L} \sum_{j=0}^{L-1} f\left(\frac{j}{L}\right) \exp\left(i2\pi m \frac{j}{L}\right), \qquad m = 0, 1, \ldots, L-1, \tag{6}$$

we have two interesting results [46]. First, the eigenvalues of this circulant matrix coincide with the power spectral density evaluated at points $w_k = \frac{k-1}{L}$,

$$\lambda_{L,k} = f\left(\frac{k-1}{L}\right). \tag{7}$$

And, second, the matrices $\boldsymbol{\Gamma}_L(f)$ and $\mathbf{C}_L(f)$ are asymptotically equivalent as $L \to \infty$, $\boldsymbol{\Gamma}_L(f) \sim \mathbf{C}_L(f)$, in the sense that both matrices have bounded eigenvalues [47] and $\lim_{L \to \infty} \frac{\|\boldsymbol{\Gamma}_L(f) - \mathbf{C}_L(f)\|_F}{\sqrt{L}} = 0$, where $\| \cdot \|_F$ is the Frobenius norm. Moreover, the eigenvalues of both matrices $\boldsymbol{\Gamma}_L(f)$ and $\mathbf{C}_L(f)$ are asymptotically equally distributed in the sense of Weyl[2] as a consequence of the fundamental theorem of Szegö [48, p. 64] as it is shown in [49].

---

[2] Two sets of bounded real numbers $\{a_{n,k}\}_{k=1}^n$ and $\{b_{n,k}\}_{k=1}^n$ are asymptotically equally distributed in the sense of Weyl if for a given continuous function F on the interval $[-K, K]$, it holds that $\lim_{n \to \infty} \frac{\sum_{k=1}^n (F(a_{n,k}) - F(b_{n,k}))}{n} = 0$.

To obtain a more operational version of the procedure, we consider the circulant matrix $\mathbf{C}_L(\widetilde{f})$ whose elements $\widetilde{c}_m$ are given by [50]:

$$\widetilde{c}_m = \frac{L-m}{L}\gamma_m + \frac{m}{L}\gamma_{L-m}, \quad m = 0, 1, \ldots, L-1, \tag{8}$$

where the generating function $\widetilde{f}$ is an approximation of the power spectral density f. Besides that, [50] shows that $\boldsymbol{\Gamma}_L(f)$ is asymptotically equivalent to $\mathbf{C}_L(\widetilde{f})$. By the transitivity property, the three matrices $\boldsymbol{\Gamma}_L(f)$, $\mathbf{C}_L(\widetilde{f})$ and $\mathbf{C}_L(f)$ are asymptotically equivalent.

Therefore, our proposal will consist on using the eigenstructure of a circulant matrix $\mathbf{C}_L(\widetilde{f})$ with elements given by (8) and, by (7), associate the kth eigenvalue and corresponding eigenvector to the frequency $w_k = \frac{k-1}{L}$. Moreover, again by (7) the spectral density is easily evaluated at frequencies w_k by the eigenvalues of the matrix $\mathbf{C}_L(\widetilde{f})$.

Finally, going to the sample we have to work with estimated, rather than population, quantities. So, we substitute the population autocovariances $\{\gamma_m\}_{m=0}^{L-1}$, by the sample second moments $\{\widehat{\gamma}_m\}_{m=0}^{L-1}$ where $\widehat{\gamma}_m, m = 0, \ldots, L-1$ is defined as

$$\widehat{\gamma}_m = \frac{1}{T-m} \sum_{t=1}^{T-m} x_t x_{t+m}. \tag{9}$$

Since the sample autocovariances converge in probability to the population autocovariances, we define $\mathbf{S}_C$ with elements given by

$$\widehat{c}_m = \frac{L-m}{L}\widehat{\gamma}_m + \frac{m}{L}\widehat{\gamma}_{L-m}, \quad m = 0, 1, \ldots, L-1. \tag{10}$$

In what follows, we describe our new proposed algorithm, named Circulant SSA. Given the time series data $\{x_t\}_{t=1}^{T}$:

1st step: Embedding. This step is as before.

2nd step: Decomposition. Compute the circulant matrix $\mathbf{S}_C$ whose elements are given in (10). Find the eigenvalues $\widehat{\lambda}_k$ of $\mathbf{S}_C$ and based on (7), associate the k-th eigenvalue and corresponding eigenvector to the frequency $w_k = \frac{k-1}{L}, k = 1, \ldots, L$.

3rd step: Grouping. Given the symmetry of the power spectral density, we have that $\widehat{\lambda}_k = \widehat{\lambda}_{L+2-k}$. Their corresponding eigenvectors given by (5) are complex, therefore, they are conjugated complex by pairs, $\mathbf{u}_k = \mathbf{u}_{L+2-k}^*$ where $\mathbf{v}^*$ indicates the complex conjugate of a vector $\mathbf{v}$, and $\mathbf{u}_k'\mathbf{X}$ and $\mathbf{u}_{L+2-k}'\mathbf{X}$ correspond to the same harmonic period. We proceed as follows to transform them in pairs of real eigenvectors in order to compute the associated components.

To form the elementary matrices we first form the groups of 2 elements $B_k = \{k, L+2-k\}$ for $k = 2, \ldots, M$ with $B_1 = \{1\}$ and $B_{\frac{L}{2}+1} = \left\{\frac{L}{2}+1\right\}$ if L is even. Second, we compute the elementary matrix by frequency $\mathbf{X}_{B_k}$ as the sum of the two elementary matrices $\mathbf{X}_k$ and $\mathbf{X}_{L+2-k}$, associated to eigenvalues $\widehat{\lambda}_k$ and $\widehat{\lambda}_{L+2-k}$ and frequency $w_k = \frac{k-1}{L}$,

$$\begin{aligned}
\mathbf{X}_{B_k} &= \mathbf{X}_k + \mathbf{X}_{L+2-k} \\
&= \mathbf{u}_k\mathbf{u}_k^H\mathbf{X} + \mathbf{u}_{L+2-k}\mathbf{u}_{L+2-k}^H\mathbf{X} \\
&= (\mathbf{u}_k\mathbf{u}_k^H + \mathbf{u}_k^*\mathbf{u}_k')\mathbf{X} \\
&= 2(R_{\mathbf{u}_k}R_{\mathbf{u}_k}' + I_{\mathbf{u}_k}I_{\mathbf{u}_k}')\mathbf{X}
\end{aligned}$$

where $R_{\mathbf{u}_k}$ denotes the real part of $\mathbf{u}_k$, $I_{\mathbf{u}_k}$ its imaginary part and $\mathbf{v}^H$ indicates the conjugate transpose of a vector $\mathbf{v}$. Notice that the matrices $\mathbf{X}_{B_k}, k = 1, \ldots, L$, are real.

4th step: Reconstruction. As before.

Notice that the elementary reconstructed series by frequency can be automatically assigned to a component according to the goal of our analysis. Fig. 1 shows the related flowchart to better illustrate the overall automated procedure and its pseudo-code is provided in Algorithm 1.

Algorithm 1 Pseudo-code of Circulant SSA.

Require: Time series x_t and window length L
Ensure: Reconstructed series associated with each disjoint group of frequencies
1: Construct the trajectory matrix $\mathbf{X}$ by (1)
2: **for** $m = 0$ **to** $L-1$ **do** ▷ *Estimated autocovariances*
3:     Compute $\widehat{\gamma}_m$ given in (9)
4: **end for**
5: **for** $m = 0$ **to** $L-1$ **do** ▷ *First row of circulant matrix* $\mathbf{S}_C$
6:     Compute $\widehat{c}_m$ given in (10)
7: **end for**
8: Build the circulant matrix $\mathbf{S}_C$
9: **for** $k = 1$ **to** L **do** ▷ *Unitary diagonalization of* $\mathbf{S}_C$ *and elementary matrices*
10:     Find the eigenvalue $\widehat{\lambda}_k$ of $\mathbf{S}_C$ based on (4)
11:     Calculate its corresponding eigenvector $\mathbf{u}_k$ by (5)
12:     The pair $(\widehat{\lambda}_k, \mathbf{u}_k)$ is associated with the frequency $w_k = \frac{k-1}{L}$
13:     Determine the contribution of the frequency w_k, $\widehat{\lambda}_k / \sum \widehat{\lambda}_k$
14:     Compute the elementary matrix $\mathbf{X}_k = \mathbf{u}_k\mathbf{u}_k^H\mathbf{X}$ associated with the frequency w_k
15: **end for**
16: Set the group $B_1 = \{1\}$ and the matrix $\mathbf{X}_{B_1} = \mathbf{X}_1$ ▷ *Elementary pairs and matrices by frequency*
17: **for** $k = 2$ **to** $M = \left\lfloor \frac{L+1}{2} \right\rfloor$ **do**
18:     Compose elementary pair by frequency $B_k = \{k, L+2-k\}$
19:     Compute elementary matrix by frequency $\mathbf{X}_{B_k} = \mathbf{X}_k + \mathbf{X}_{L+2-k}$
20: **end for**
21: **if** L is even **then**
22:     Set the group $B_{\frac{L}{2}+1} = \left\{\frac{L}{2}+1\right\}$ and the matrix $\mathbf{X}_{B_{\frac{L}{2}+1}} = \mathbf{X}_{\frac{L}{2}+1}$
23: **end if**
24: Determine the G disjoint groups I_j of the pairs B_k with the significant or interesting frequencies w_k for the non-zero contributions
25: **for** $j = 1$ **to** G **do** ▷ *Matrices associated with the disjoint groups*
26:     Compute the matrix $\mathbf{X}_{I_j}$ associated with each group I_j by $\mathbf{X}_{I_j} = \sum_{B_k \in I_j} \mathbf{X}_{B_k}$
27: **end for**
28: **for** $j = 1$ **to** G **do** ▷ *Reconstructed series*
29:     Calculate the reconstructed series $\widetilde{x}_t^{(j)}$ by the diagonal averaging or *hankelization* of matrix $\mathbf{X}_{I_j}$, $\widetilde{x}_t^{(j)} = H\left(\mathbf{X}_{I_j}\right)$
30: **end for**

3.2. Asymptotic equivalence of Basic, Toeplitz and Circulant SSA

Toeplitz and Circulant SSA are modifications of the original Basic SSA. In this section, we will prove that the three versions of SSA (Basic, Toeplitz and Circulant) are asymptotically equivalent according to the definition given in [46]. Later on, we will run some simulations to compare the performance of the three versions in finite samples.

Theorem 1. *Given the $L \times N$ trajectory matrix $\mathbf{X}$ defined in (1), let $\mathbf{S}_B = \mathbf{X}\mathbf{X}'/N$, $\mathbf{S}_T$ the Toeplitz matrix with elements defined by (2) and $\mathbf{S}_C$ the circulant matrix with elements given in (10). Consider the sequence of matrices $\{\mathbf{S}_B\}$, $\{\mathbf{S}_T\}$ and $\{\mathbf{S}_C\}$ as $L \longrightarrow \infty$. Then $\mathbf{S}_B \sim \mathbf{S}_T \sim \mathbf{S}_C$.*

Proof. The proof is given in the appendix □

This theorem gives the basis to understand the similar results obtained in practice between Basic and Toeplitz SSA when the window length is very large (the larger, the better as the result is asymptotic). This was empirically shown using stationary time series in climate and geophysics [51,52]. Here, we provide a theoretical basis for these empirical findings. Additionally, we also extend
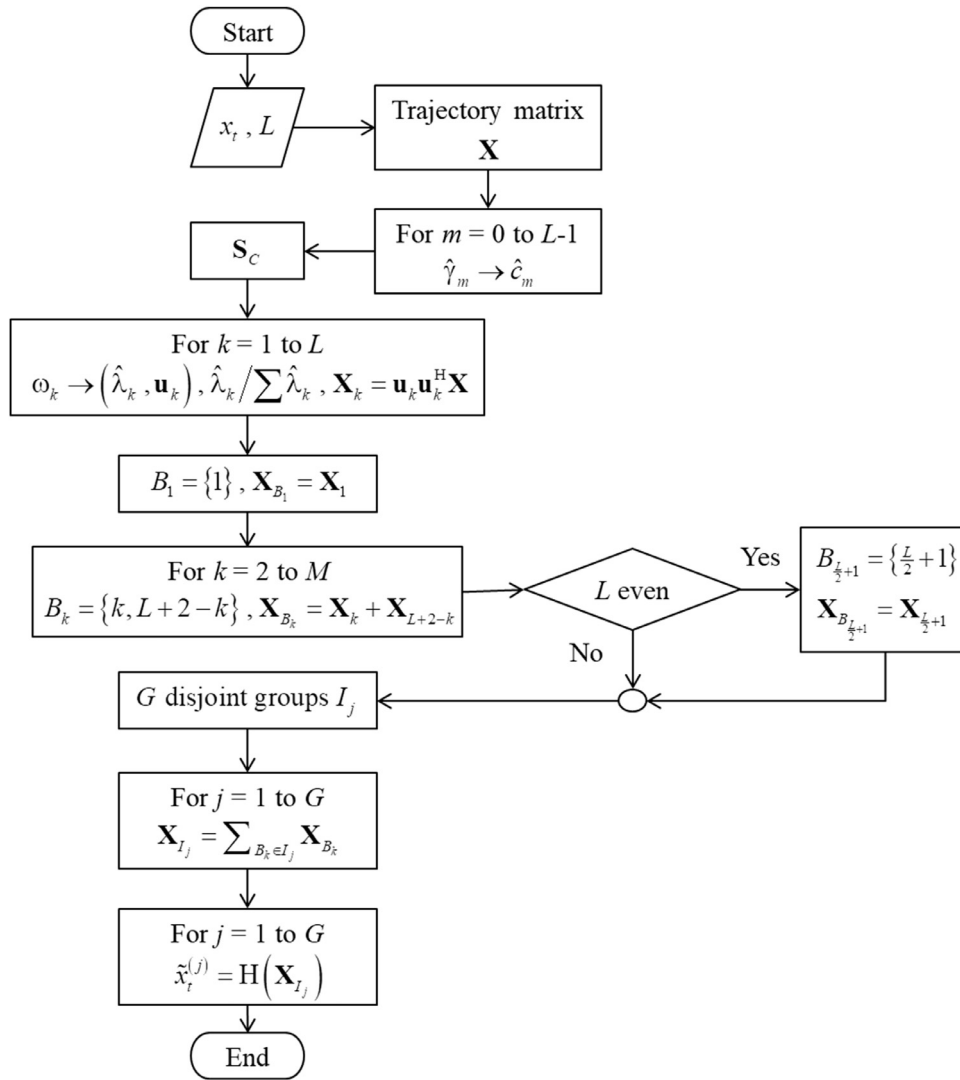
Fig. 1. Flowchart of Circulant SSA algorithm.

the result for the new version of SSA that we have introduced in this paper, CiSSA.

3.3. Nonstationary case

In economics, many time series are nonstationary in the sense that the power spectral density function has discontinuities. This has important consequences in our analysis and we have to show that Circulant SSA can be applied to nonstationary time series. The next theorem, a generalization of the analogous Gray's theorem [53, Theorem 3], provides the theoretical background needed to apply CiSSA to nonstationary time series.

Theorem 2. Let $\mathbf{T}_L(s)$ be a sequence of Toeplitz matrices with $s(w)$ a real, continuous and 2π-periodic, such that $s(w) \geq 0$, where the equality is reached in a finite number of points $H = \{w_i^0, i = 1, \ldots, l\}$. Given a finite δ, consider the disjoint sets

$$\Omega_i = \left\{ w \in \left[w_i^0 - a_i, w_i^0 + a_i \right] | s(w) \leq \frac{1}{\delta} \right\}, a_i \in \mathbb{R}^+, \; i = 1, \ldots, l$$

and let $g(w)$ be a function defined as

$$g(w) = \begin{cases} f(w) = \frac{1}{s(w)} & \text{if } w \notin \bigcup_{i=1}^l \Omega_i \\ h_i(w) & \text{if } w \in \Omega_i \end{cases}$$

where $h_i(w)$ is any real valued bounded function continuous in Ω_i and symmetric around w_i^0. Let $M_{h_i} = \sup h_i < \infty$ and $m_{h_i} = \inf h_i = h_i\left(w_i^0 - a_i\right) = h_i\left(w_i^0 + a_i\right) = \delta$.

Let $\rho_{L,k}, k = 1, \ldots, L$, be the eigenvalues of $(\mathbf{T}_L(s))^{-1}$ sorted in decreasing order and let $F(x)$ be a continuous function in $\left[\frac{1}{M_s}, \max_i M_{h_i}\right]$ with $M_s = \sup s$, then

$$\lim_{L \to \infty} \frac{1}{L} \sum_{k=1}^{L} F(\min(\rho_{L,k}, \max(\widetilde{g}_k, \delta))) = \int_0^1 F(g(w))dw, \qquad (11)$$

where $\widetilde{g}_k$ are the values of $g(\frac{k-1}{L})$ sorted in descending order.

Proof. The proof is given in the appendix □

Notice that while Gray's Theorem [53] approximates the speudo-spectrum by a constant value, Theorem 2 allows for a better separation of the components around the spectrum discontinuities that is specially relevant for low frequencies. In a similar way to [53], the theorem states that the sequence of eigenvalues of the sequence of matrices $(\mathbf{T}_L(s))^{-1}$ are asymptotically equally distributed (in the sense of Weyl) as the eigenvalues of the sequence of matrices $\mathbf{T}_L(g)$ up to a finite value δ as L tends to infinity. Moreover, the matrices $\mathbf{T}_L(g) \sim \mathbf{C}_L(g)$ and, by Szegö's theorem, the eigenvalues of the sequence of matrices $\mathbf{T}_L(g)$ are asymptotically equally

distributed as the eigenvalues of the sequence of matrices $\mathbf{C}_L(g)$ up to a finite value δ as L tends to infinity.

As a result, for a nonstationary series, the union of the estimation of the pseudo-power spectral density in a point of discontinuity with the estimations in the adjoint frequencies through segments is an easy way of building the functions h_i. If all the functions h_i are constant and equal to a particular value δ finite, we have the particular case proved in [53, Theorem 3]. Therefore, the generalization to functions h_i allows a better approximation of the pseudo-power spectral density when we increase the window length. Fig. 1 and the pseudo-code in Algorithm 1 previously introduced for the description of CiSSA are also valid in the nonstationary case.

4. Simulations

In this section we check the performance of our new proposal, Circulant SSA, in finite samples and compare it with the competing SSA algorithms, i.e. Basic SSA and Toeplitz SSA for a linear as well as a nonlinear time series model. Even though SSA is nonparametric and therefore model free, in this section we generate time series following a known model and check the basic statistical properties related to the signal extraction procedure. In particular, we check if the extracted signals are unbiased. These simulations generalize previous exercises [54] by including CiSSA, but also using more complex time series models in a linear and nonlinear framework.

4.1. Linear time series

The first model is a basic structural time series model

$$x_t = T_t + c_t + s_t + e_t \tag{12}$$

where T_t is the trend component, c_t is the cycle, s_t is the seasonal component and e_t is the irregular component. We assume an integrated random walk for the trend [55] given by

$$T_t = T_{t-1} + \beta_{t-1} \tag{13}$$
$$\beta_t = \beta_{t-1} + \eta_t$$

with $\eta_t \sim N(0, \sigma_\eta^2)$. The cyclical and seasonal components are specified according to [56], where the cycle is given by the first component of the bivariate Vector Autoregressive of order 1, VAR(1), model

$$\begin{pmatrix} c_t \\ \tilde{c}_t \end{pmatrix} = \rho_c \begin{pmatrix} \cos(2\pi w_c) & \sin(2\pi w_c) \\ -\sin(2\pi w_c) & \cos(2\pi w_c) \end{pmatrix} \begin{pmatrix} c_{t-1} \\ \tilde{c}_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \tilde{\varepsilon}_t \end{pmatrix} \tag{14}$$

with $\begin{pmatrix} \varepsilon_t \\ \tilde{\varepsilon}_t \end{pmatrix} \sim N(\mathbf{0}, \sigma_\varepsilon^2 I)$ and $\frac{1}{w_c}$ the period, $w_c \in [0, 1]$. And, the seasonal component is given by

$$s_t = \sum_{j=1}^{[s/2]} a_{j,t} \cos(2\pi w_j t) + b_{j,t} \sin(2\pi w_j t) \tag{15}$$

with $w_j = \frac{j}{s}, j = 1, \ldots, [s/2]$ and s the seasonal period, where $[\cdot]$ is the integer part and $a_{j,t}$ and $b_{j,t}$ are two independent random walks with noise variances equal to σ_j^2. Finally, the irregular component is white noise with variance σ_e^2. All the components are independent of each other. We set $\rho_c = 1$, so the trend, cycle and seasonal components may have poles of module 1 and therefore are nonstationary. We consider that the series are monthly with $s = 12$ and cyclical period equal to $\frac{1}{w_c} = 48$ months. The sample size is $T = 193$ and the noise variances of the different components are given by $\sigma_\eta^2 = 0.0006^2$, $\sigma_j^2 = 0.004^2$, $\sigma_\varepsilon^2 = 0.008^2$ and $\sigma_e^2 = 0.06^2$. We choose as window length $L = 48$ because this value of L is multiple of the seasonal period, it is equal to the cyclical period and $T - 1$ is multiple of L [1].

The trend is related to frequency 0, the cycle to frequency 1/48 and the seasonal components to frequencies 1/12, 1/6, 1/3, 1/4, 5/12 and 1/2. Given (7), we can recover the signal associated to a frequency $w = \frac{k-1}{L}$ by using the elementary components associated to eigenvalues k and $k' = L + 2 - k$, the latter by the symmetry of the spectral density. Therefore, the trend is reconstructed with eigentriple 1, the cyclical component with eigentriples 2 and 48, and the seasonal components with eigentriples 5, 9, 13, 17, 21, 25, 29, 33, 37, 41 and 45. For example, for the frequency $w = \frac{1}{12}$, we have that $\frac{k-1}{L} = \frac{1}{12}$, and therefore, we sum the elementary components $k = \frac{48}{12} + 1 = 5$ and $k' = L + 2 - k = 48 + 2 - 5 = 45$.

If the procedure for signal extraction works well, the simulated component y_t (y_t can be the trend, cycle or seasonal component) could be written as

$$y_t = \hat{y}_t + u_t$$

where u_t is the noise and $\hat{y}_t$ is the extracted signal. Then, in the regression

$$y_t = a + b\hat{y}_t + u_t \tag{16}$$

$a = 0$ (unbiasedness) and $b = 1$ (the scale is not changed). Notice that y_t and $\hat{y}_t$ should be cointegrated. We simulate 10,000 times the model and perform signal extraction with Circulant SSA. Table 1 shows the percentiles of the empirical distribution of the estimated coefficients of the regression in (16).

Table 1 shows that the median of the estimated intercept is almost zero for the three estimated components (cycle, seasonal component and trend). The median for the scale parameter b is almost one for the three components, but looking at the values for different quantiles, the empirical distribution for the estimated b associated to the cycle indicates a larger dispersion.

The estimated residuals from Eq. (12) are given by $\hat{e}_t = x_t - \hat{T}_t - \hat{c}_t - \hat{s}_t$, and should be white noise, where $\hat{T}_t, \hat{c}_t$, and $\hat{s}_t$ are the estimates of the trend, cycle and seasonal component respectively. In order to check this, we fit an AR(1) to $\hat{e}_t$. Table 2 shows the quantiles of the empirical distributions of the mean, standard error and autoregressive coefficient of the residuals of the 10,000 replications. The median of the mean and autoregressive coefficient are close to zero. The median of the standard deviation is 0.0529 (the value used for the simulations was 0.06).

The results from the simulations seem very good. In order to compare Circulant SSA with alternative algorithms as Basic and Toeplitz SSA we also simulate the linear model given by (12) and extract the trend, cycle and seasonal components for 10,000 simulations. Basic and Toeplitz SSA require first to calculate the principal components and then to identify the frequency they represent with some procedure as stated in the first paragraph of this section. However, given that we are using simulated time series and we know beforehand the frequencies that might be more informative, we proceed in a heuristic way. According to model (12), we know that the informative frequencies are $\Omega = \{0, 1/48, 1/12, 1/6, 1/4, 1/3, 5/12, 1/2\}$ and the window length $L = 48$ coincides with the cycle periodicity and is multiple of the seasonal periodicity of a monthly time series. Also each eigenvector generates a linear subspace associated to a frequency. In this way, we calculate the periodogram for each eigenvector and obtain the frequency associated with the maximum. If that frequency belongs to the set Ω, the associated component to that eigenvector is assigned to the trend, cycle or seasonal component and, on the contrary it is assigned to the residual $\hat{e}_t$.

In analogous way to Circulant SSA, we perform regressions as in (16) between simulated an estimated components and check $a = 0$ and $b = 1$. Table 1 shows the quantiles of the 10,000 estimated values for a and b. Results are very similar for the three versions of SSA and it can be accepted that the estimated values are close to $a = 0$ and $b = 1$. These simulations allow to conclude that, at least

Table 1

Statistics related to the goodness of fit of the extracted signals for the different methods. Simulations for the linear model, N=10000. Columns show the quantiles of the empirical distribution of the estimated coefficients of the regression of the generated components over the estimated ones.

Statistic	Component	Quantiles				
		5	25	50	75	95
Circulant SSA						
$\hat{a}$	Trend	-0.0613	-0.0209	-0.0006	0.0194	0.0600
	Cycle	-0.0109	-0.0043	0.0000	0.0045	0.0108
	Seasonal	-0.0015	-0.0006	0.0000	0.0006	0.0015
$\hat{b}$	Trend	0.9748	0.9951	1.0032	1.0143	1.0651
	Cycle	0.8481	0.9569	1.0029	1.0476	1.1340
	Seasonal	0.9451	0.9819	1.0049	1.0277	1.0630
Basic SSA						
$\hat{a}$	Trend	-0.0610	-0.0206	-0.0006	0.0191	0.0598
	Cycle	-0.0165	-0.0066	0.0001	0.0065	0.0167
	Seasonal	-0.0033	-0.0010	0.0000	0.0010	0.0033
$\hat{b}$	Trend	0.9881	1.0063	1.0153	1.0326	1.1292
	Cycle	0.7891	0.9618	1.0177	1.0794	1.2793
	Seasonal	0.9471	0.9911	1.0166	1.0431	1.0867
Toeplitz SSA						
$\hat{a}$	Trend	-0.0588	-0.0203	-0.0007	0.0186	0.0566
	Cycle	-0.0178	-0.0061	0.0001	0.0062	0.0170
	Seasonal	-0.0017	-0.0007	0.0000	0.0007	0.0018
$\hat{b}$	Trend	0.9820	1.0003	1.0088	1.0264	1.1415
	Cycle	0.7852	0.9863	1.0537	1.1310	1.2754
	Seasonal	0.9554	0.9982	1.0273	1.0605	1.1207

Table 2

Statistics related to the residual term $\hat{e}_t$ in Circulant SSA: Average, standard deviation and autoregressive coefficient of AR(1). Simulations for the linear model, N=10000.

Statistic	Quantiles				
	5	25	50	75	95
Average	-0.0033	-0.0012	0.0000	0.0011	0.0033
Standard deviation	0.0478	0.0508	0.0529	0.0551	0.0581
AR(1) coefficient	-0.1693	-0.0870	-0.0313	0.0285	0.1075

for the proposed linear model, empirically, the three versions of SSA are equivalent. However, some differences can be found in the estimation of the cycle, where the distributions of the estimates of a and b show less dispersion around 0 and 1 with CiSSA.

4.2. Nonlinear time series

For the case of nonlinear time series, we borrow the model from [56] for UK travellers given by

$$x_t = T_t + c_t + \exp(a_0 + a_1 T_t)\gamma_t + e_t$$

where T_t is the trend, c_t is the cycle and γs_t is the seasonal component specified as in (13), (14) and (15), respectively. The parameters a_0 and a_1 are unknown fixed coefficients. Coefficient a_0 scales the seasonal component. The sign of the coefficient a_1 determines whether the seasonal variation increases or decreases when a positive change in the trend occurs. The overall time varying amplitude of the seasonal component is determined by the combination $a_0 + a_1 \mu_t$.

As for the linear case, we simulate the model 10,000 times for series of length $T = 193$ observations. We set a_0 and a_1 such that for each replication $0.5 \leq \exp(a_0 + a_1 \mu_t) \leq 1.5$, with $a_1 > 0$. We apply Circulant SSA with a window length $L = 48$. Table 3 shows the quantiles of the empirical distribution of the estimated coefficients of the regression in (16) and again we can see that the values of a and b estimated are located around 0 and 1 respectively with low dispersion.

In order to check that the estimated residuals are white noise, we fit an AR(1) to $\hat{e}_t$ as in the linear case. Table 4 shows the quantiles of the empirical distribution of the mean, standard error and autoregressive coefficient of the residuals of the 10,000 replications. The median of the mean and autoregressive coefficient are close to zero. The median of the standard deviation is 0.053 (the value used for the simulations was 0.06).

As in the linear case, the results from the simulations seem very good. To compare Circulant SSA with alternative algorithms as Basic and Toeplitz SSA, we repeat the simulations described in the previous section and apply the same steps to obtain their trend, cycle and seasonal components. Again, we perform regressions as in (16) between simulated an estimated components and check $a = 0$ and $b = 1$. Table 3 shows the quantiles of the 10,000 estimated values for a and b. The same conclusions as in the linear case apply: it can be accepted that the estimated values are close to $a = 0$ and $b = 1$; empirically, the three versions of SSA are equivalent for the proposed linear model; and some differences can be found in the cycle estimations, where the distribution of the estimates of a and b show less dispersion around 0 and 1 with CiSSA.

5. Applications

5.1. Industrial production

We consider monthly series of Industrial Production (IP), index 2010=100, of six countries: France, Germany, Italy, UK, Japan and US. Industrial Production is widely followed since it is pointed out in the definition of a recession by the National Bureau of Economic Research (NBER), as one of the four monthly indicators series to check in the analysis of the business cycle. The sample covers from January 1970 to September 2014, so the sample size is $T = 537$. The data source is the IMF database. As it can be seen in Fig. 2, these indicators show different trend, seasonality and cyclical behavior, and our goal is to extract these components and discuss about the results.

The first step is to establish the window length. Due to the monthly periodicity and seasonality, we select a window length

Table 3

Statistics related to the goodness of fit of the extracted signals for the different methods. Simulations for the nonlinear model, N=10000. Columns show the quantiles of the empirical distribution of the estimated coefficients of the regression of the generated components over the estimated ones.

Statistic	Component	Quantiles				
		5	25	50	75	95
Circulant SSA						
$\hat{a}$	Trend	-0.0603	-0.0199	0.0004	0.0202	0.0609
	Cycle	-0.0111	-0.0045	-0.0001	0.0043	0.0112
	Seasonal	-0.0015	-0.0006	0.0000	0.0006	0.0015
$\hat{b}$	Trend	0.9742	0.9951	1.0037	1.0154	1.0682
	Cycle	0.8442	0.9567	1.0029	1.0475	1.1353
	Seasonal	0.9241	0.9779	1.0072	1.0335	1.0720
Basic SSA						
$\hat{a}$	Trend	-0.0602	-0.0198	0.0005	0.0199	0.0605
	Cycle	-0.0167	-0.0065	0.0000	0.0066	0.0163
	Seasonal	-0.0035	-0.0010	0.0000	0.0009	0.0030
$\hat{b}$	Trend	0.9880	1.0064	1.0158	1.0337	1.1284
	Cycle	0.7626	0.9588	1.0158	1.0763	1.2660
	Seasonal	0.9269	0.9888	1.0236	1.0561	1.1084
Toeplitz SSA						
$\hat{a}$	Trend	-0.0581	-0.0191	0.0002	0.0195	0.0602
	Cycle	-0.0176	-0.0063	-0.0001	0.0064	0.0185
	Seasonal	-0.0019	-0.0007	-0.0001	0.0006	0.0016
$\hat{b}$	Trend	0.9814	1.0004	1.0093	1.0284	1.1424
	Cycle	0.7609	0.9812	1.0513	1.1279	1.2767
	Seasonal	0.9351	0.9977	1.0315	1.0667	1.1316

Table 4

Statistics related to the residual term $\hat{e}_t$ in Circulant SSA: Average, standard deviation and autoregressive coefficient of AR(1). Simulations for the nonlinear model, N=10000.

Statistic	Quantiles				
	5	25	50	75	95
Average	-0.0034	-0.0011	0.0000	0.0012	0.0033
Standard deviation	0.0476	0.0508	0.0531	0.0554	0.0590
AR(1) coefficient	-0.1727	-0.0899	-0.0339	0.0250	0.1066

Table 5

Contribution of the different signals to IP in the six countries in percentage.

Component	Country					
	France	Germany	Italy	Japan	UK	USA
Trend	52.1	77.3	42.7	79.0	72.0	87.9
Cycle	9.5	12.6	7.8	13.8	11.1	10.3
Seasonal	35.6	6.7	47.3	5.1	13.5	0.3
Irregular	2.8	3.4	2.2	2.1	3.4	1.5

multiple of 12. Assuming that the period of the cycle in these series goes from 1 year and a half to 8 years, we choose a window length multiple of 8 × 12=96 months. From the two available options, 96 and 192 months, we select the second one since it is larger.

According to (7) for $k = 1$, we have $w_1 = \frac{k-1}{L} = 0$ and it will be associated to the trend. In the same way, for $k = 2$, we have $w_2 = 1/192$, that corresponds to 192 months or 16 years that are beyond cyclical movements between 1.5 and 8 years. Therefore, given (7) and the symmetry of the power spectral density, the trend is reconstructed with the eigentriples 1, 2 and $L + 2 - k = 192$ with the elementary groups by frequencies from B_1 and B_2 respectively. In an analogous way, assuming that the business cycle goes from 1.5 to 8 years, this component is associated to frequencies $w_k = 1/96, 1/64, 1/48, 5/192, 1/32, 7/192, 1/24, 3/64, 5/96$ and the cycle signal is reconstructed with the eigentriples 3 to 11 and 183 to 191, with the elementary groups by frequencies from B_3 to B_{11}. Finally, the seasonal component is associated to the frequencies $w_k = 1/12, 1/6, 1/4, 1/3, 5/12, 1/2$ and reconstructed in a similar way with the eigentriples 17, 33, 49, 65, 81, 97, 113, 129, 145, 161 and 177 and with the elementary groups by frequencies $B_{17}, B_{33}, B_{49}, B_{65}, B_{81}$, and B_{97}.

Table 5 shows the contributions of the signals to the original IP variations in percentage. First, we highlight that the contribution of the irregular component (those oscillations not explained by the trend, cyclical or seasonal components) is smaller than 3.5% in all the countries. Main contributions come from the trend and seasonality, that account for more than 84% in all the countries. As expected, the contribution of the seasonal component is almost negligible in US, and quite small in Japan and Germany, while it is very relevant in Italy and France. Finally, the cycle contributes in a range between 7.8% in Italy to 13.8% in Japan.

Fig. 2 shows the estimated trends for every country. The trend is a smooth component that has shown a decreasing evolution since the last decade for France, Italy and UK as a consequence of the last economic crisis. On the contrary, in Germany and US, the trend shows an upward evolution in all the sample period.

Fig. 3 shows the cyclical component where the shaded areas correspond to recessions as dated by the OECD[3]. We can see that the extracted cycle reflects quite well the business cycle for all countries.

5.1.1. Separability of the estimated components with CiSSA

One desirable property of the signal extraction method is that the resulting components should be orthogonal. However, in practice, they usually exhibit cross-correlation. Residual seasonality in seasonal adjusted time series is another concern in any signal extraction method from very early times [57,58], and it is still a matter of interest nowadays. Findley et al. [59] point out that "The most fundamental seasonal adjustment deficiency is detectable seasonality after adjustment". This is also a concern for policy makers [60].

---

[3] https://www.oecd.org/sdd/leading-indicators/oecdcompositeleadingindicatorsreferenceturningpointsandcomponentseries.htm

France

Germany

Italy

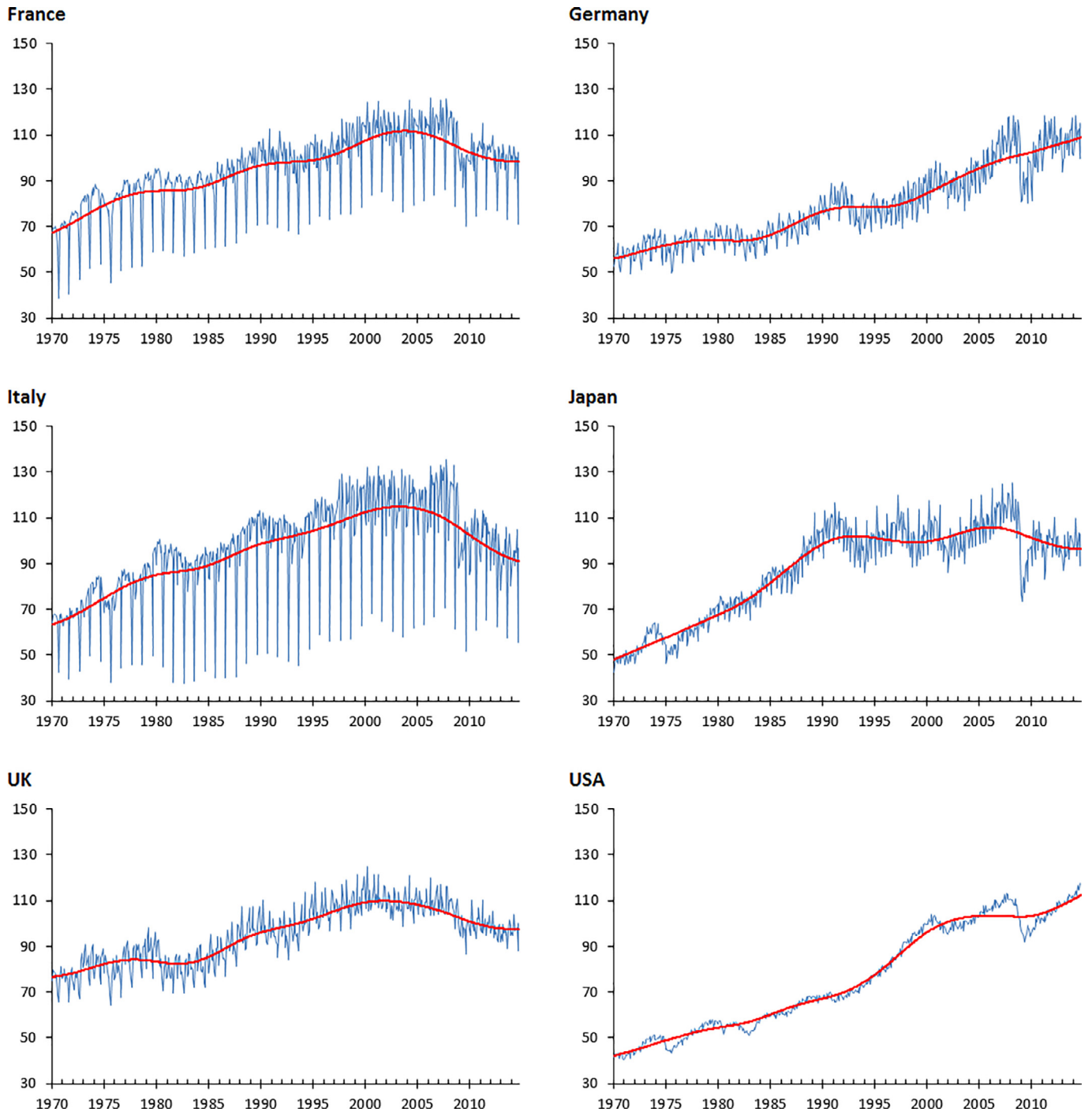**Japan**

UK

USA

Fig. 2. Original IP and trend for the different countries.

Separability of the elementary series as well as those grouped by frequencies is an assumption of SSA and should also be a characteristic of the estimated components. This characteristic is important since many signal extraction procedures assume zero correlation between their underlying components, whereas the estimated signals can be quite correlated. The SSA decomposition can be successful only if the resulting additive components of the series are quite separable from each other [39].

For a fixed window length L, given two series $\{x_t^{(1)}\}$ and $\{x_t^{(2)}\}$ extracted from the series $\{x_t\}$, we say that they are weakly separable if both their column as well as row spaces are orthogonal, that is $\mathbf{X}^{(1)}(\mathbf{X}^{(2)})' = \mathbf{0}_{L \times L}$ and $(\mathbf{X}^{(1)})'\mathbf{X}^{(2)} = \mathbf{0}_{N \times N}$. Furthermore, we

say that two series $\{x_t^{(1)}\}$ and $\{x_t^{(2)}\}$ are strongly separable if they are weakly separable and the two sets of singular values of the trajectory matrices $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are disjoint. When the trajectory matrix of the original time series has no multiple singular values or, equivalently, each elementary reconstructed series belongs to a different harmonic, strong separability is guaranteed according to the previous definition.

Usually, separability is measured in terms of **w**-correlation [1,39] that it is given by

$$\rho_{12}^w = \frac{\left\langle \mathbf{X}^{(1)}, \mathbf{X}^{(2)} \right\rangle_w}{\left\| \mathbf{X}^{(1)} \right\|_w \left\| \mathbf{X}^{(2)} \right\|_w},$$

France

Germany
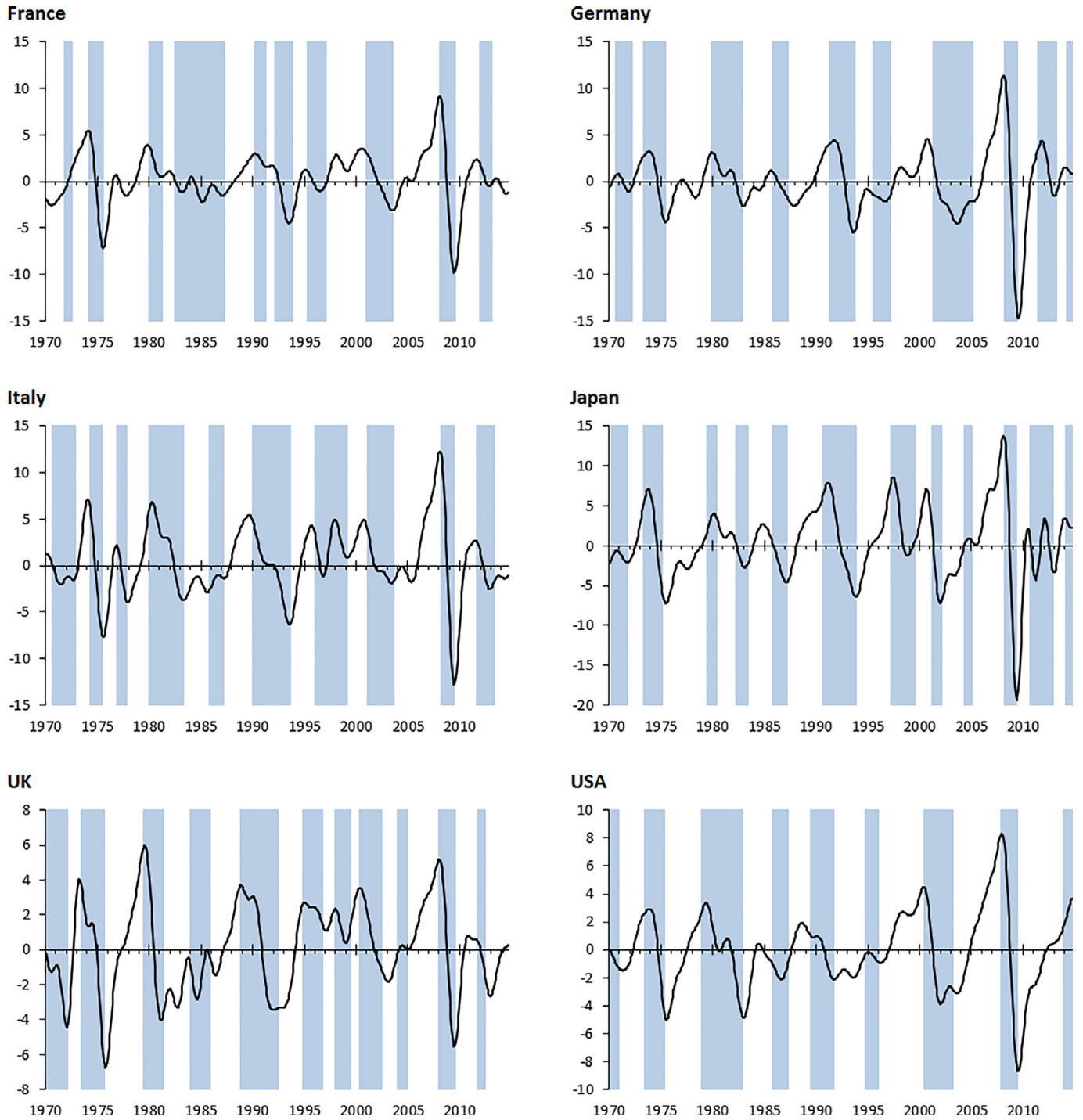

Italy

Japan

UK

USA

Fig. 3. Estimated IP cycles and OECD announced recessions (shadowed areas).

where $\left\langle \mathbf{x}^{(1)}, \mathbf{x}^{(2)} \right\rangle_w = (\mathbf{x}^{(1)})'\mathbf{W}\mathbf{x}^{(2)}$ is the so called w-inner product, $\left\| \mathbf{x}^{(1)} \right\|_w = \sqrt{\left\langle \mathbf{x}^{(1)}, \mathbf{x}^{(1)} \right\rangle_w}$ and $\mathbf{W} = \text{diag}(1, 2, \cdots, \underbrace{L, \cdots, L}_{T-2(L-1) \text{ times}}, \cdots, 2, 1)$. Note that the window length L enters the definition of **w**-correlation. We are interested on producing components with **w**-correlation (ideally) zero because, in this case, we can conclude that the component series are **w**-orthogonal, i. e. $\left\langle \mathbf{x}^{(1)}, \mathbf{x}^{(2)} \right\rangle_w = 0$ and separable [39].

To show that Circulant SSA produces components that are strongly separable, first notice that the real eigenvectors $\sqrt{2}R_{\mathbf{u}_k}$

and $\sqrt{2}I_{\mathbf{u}_k}$ (linked to eigenvalues λ_k and λ_{L+2-k}, respectively, $\lambda_k = \lambda_{L+2-k}$) are orthogonal and have information associated only to frequency $\frac{k-1}{L}$. Those are the only eigenvectors that have information related to this frequency. As eigenvectors can be considered filters [23,24], these pair of eigenvectors extract elementary series linked to the same frequency without mixing harmonics of other frequencies. As a result, the two elementary series, when reconstructed in step 4, have spectral correlation close to 1 between them and close to zero with the remaining ones. Taking into account the pairs of reconstructed series per frequency, any grouping of the reconstructed series results in disjoint sets from the point of view of
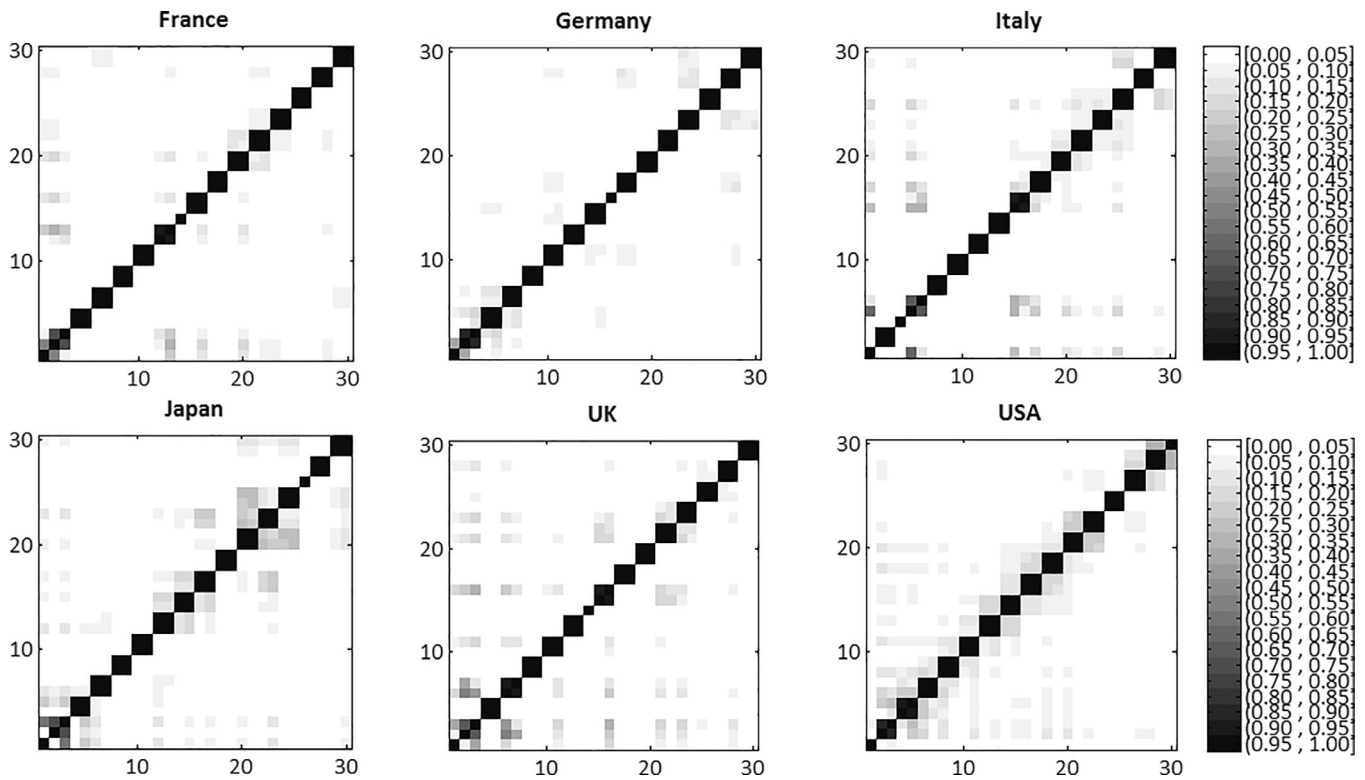
Fig. 4. w-correlation matrix for the IP elementary reconstructed series for the 30 greatest eigenvalues.

the frequency. Then, Circulant SSA produces components that are approximately strongly separable.

To quickly check how separable the components are, Fig. 4 plots the matrix of the absolute values of the **w**-correlations for all the IP components, coloring in white the absence of **w**-correlation, in black **w**-correlations in absolute value equal to 1 and in a scale of grey colors the remaining intermediate values. It can be seen that, as expected, Circulant SSA produces components that are strongly separable.

Furthermore, seasonal adjusted time series for Industrial Production are largely followed by real time analysts, and one desirable property is that they have no remaining seasonality. To check the quality of seasonal adjustment by Circulant SSA, we have applied the *combined test for seasonality* [61] used in X12-ARIMA. We found that there were no signs of any remaining seasonality in any of the seasonal adjusted time series for the different countries[4].

5.2. Speech processing

To further illustrate CiSSA, we consider a segment of voiced speech from the file handel.mat available in Matlab Central. The segment takes the first 2.08 seconds that reproduces the word "Alleluia" when it appears for the first time. The segment length consists of 17,000 observations and a sample frequency of 8192 Hz and it is represented by the blue line in Fig. 5.

We consider $L = 8192$, that is the frequency sample, a value below $T/2$ that will pick the existing oscillations[5]. Fig. 6 shows the estimated power spectral density (PSD) in dB. The black line separates positive PSD values. These values correspond to normalized

[4] Results are available from the authors upon request.

[5] Results are robust to other values of L, like for instance $L = 1024$ or $L = 4096$ (chosen to be a multiple or a fraction of the frequency sample) and are available from the authors upon request.

frequencies around 0.07, 0.13 and 0.14 and values of $k = 574$, 1065 and 1148 respectively.

Summing up the reconstructed components with positive PSD around these values we obtain the red line in Fig. 5 that captures the first vowel "a", /æ/, that is high-pitched and persistent. They represent 26.5% of the original voice signal variability. To deeper understand this signal we calculate the amplitude of the reconstructed components as the result of a low-pass Butterworth filter of order 4 and cutoff normalized frequency 0.01 to the module of their Hilbert Transform. Fig. 7 confirms the high-pitched and the persistence of the vowel /æ/ during the first second of the recording and the slowly fading in the next second.

Finally, the red line in Fig. 6 indicates PSD values over percentile 95% and the green line over percentile 90%. The sum of the reconstructed components over percentile 95% accounts for 69.7% of total variability and allows to recognize the full word "Alleluia". Finally, taking as reference percentile 90%, the corresponding sum of the reconstructed components accounts for 84% of the total variability of the original signal and reproduces the word with clarity.
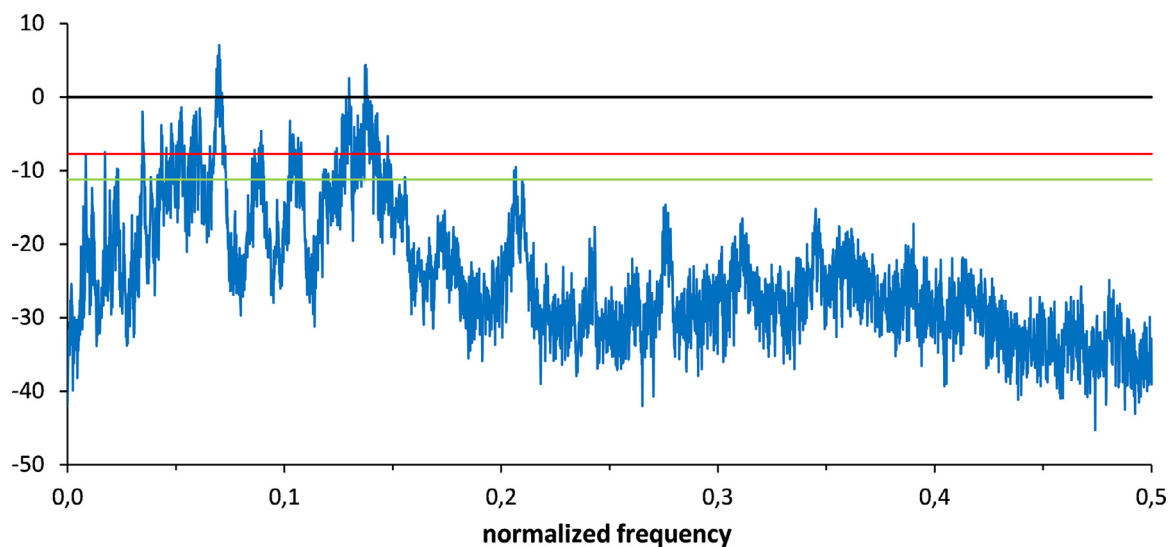
6. Comparison with other signal extraction procedures

In this section we compare CiSSA with other state-of-the art techniques for signal extraction applied in different research fields like voice recognition [30,62], medicine [31], finance [63], or art and logo design [64] among others. The signals in these applications can be nonlinear and nonstationary in the form of amplitude and frequency changing in time. Such signals are modelled by multicomponent AM-FM decomposition methods like the Hilbert Huang Transform (HHT) [29,65], the Iterative Hilbert Transform (IHT) [30,62] or the Hilbert Vibration Decomposition (HVD) [31,66]. For this type of signals, alternative versions of SSA are acknowledged as useful tools to represent a slowly changing amplitude [2]. The simulations in Section 4.2 already illustrate that CiSSA also performs well with changing amplitude signals. In this
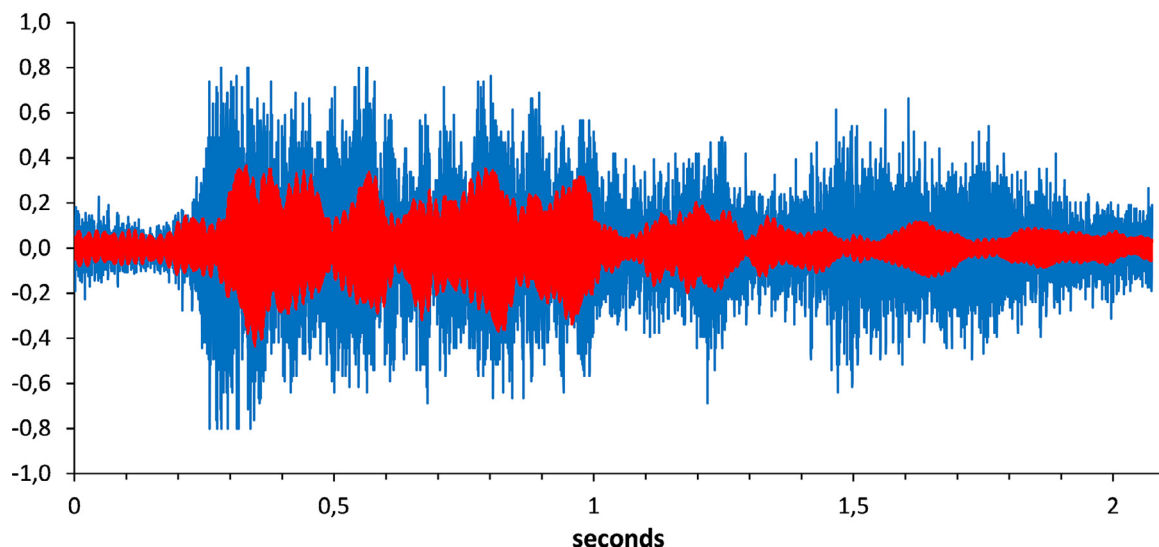
Fig. 5. Original voice segment for the word Alleluia (blue line) and sum of reconstructed CiSSA components (red line) with positive PSD. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. Estimated CiSSA power spectral density in dB ($L = 8192$) for the word Alleluia.
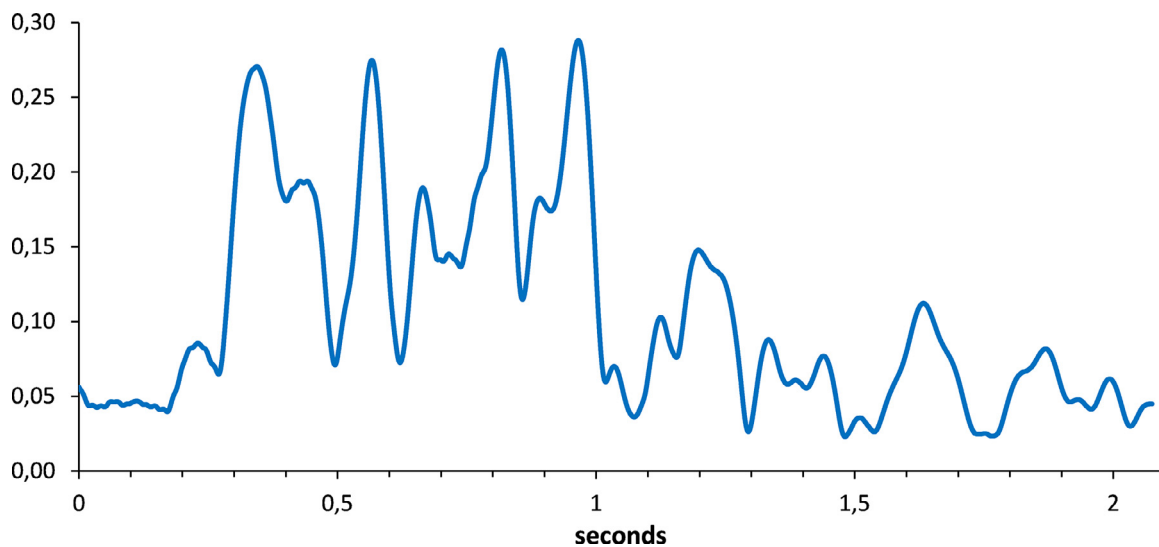
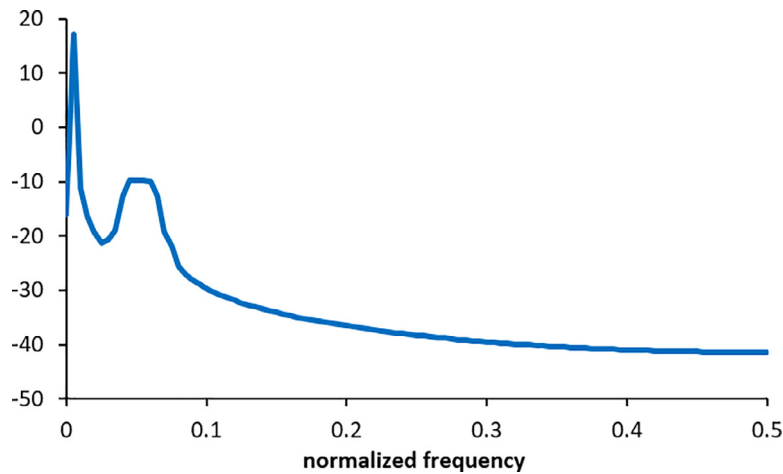**Fig. 7.** Amplitude of the sum of reconstructed components with positive PSD for the word Alleluia.

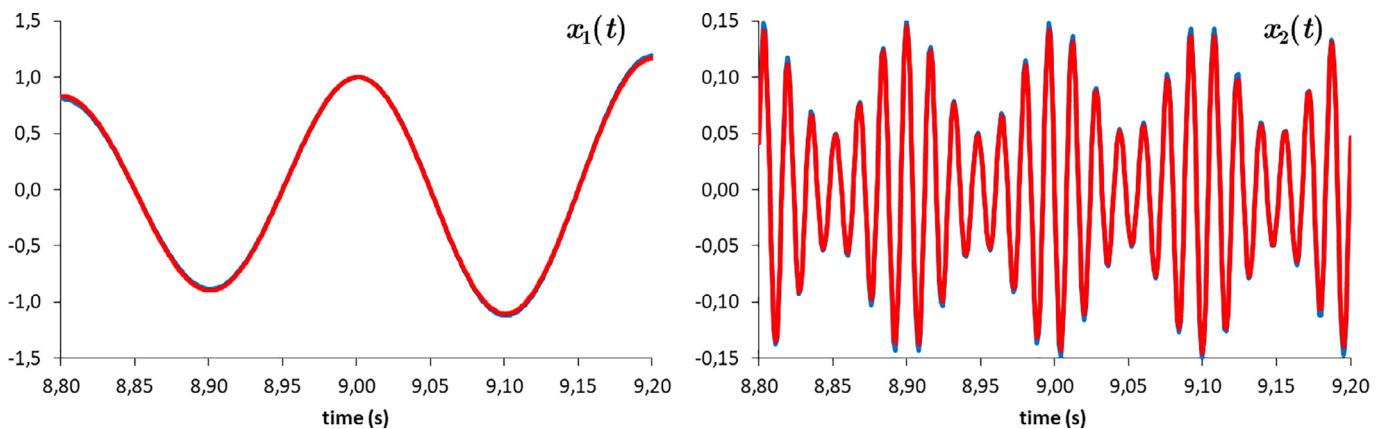**Fig. 8.** Spectral density estimation of the synthetic signal $x(t)$.

Fig. 9. Synthetic components $x_1(t)$ and $x_2(t)$ (in blue) and the corresponding CiSSA estimation (in red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

section we show that CiSSA is also suitable to capture FM signals and illustrates its functioning in this context by understanding its performance with the application to a synthetic signal previously used in the literature where it has been shown that AM-FM techniques work well [31]. Additionally, we also compare the behaviour of CiSSA and the above mentioned methods extracting the signals of a real economic time series that is characterized by evolving trend, strong seasonality and business cycle oscillations.

We consider a synthetic example taken from [31] that results from the sum of a simple AM signal and another one that is both amplitude and frequency modulated. Let $x(t) = x_1(t) + x_2(t)$ where $x_1(t) = a_1(t)\cos(w_1 t)$ and $x_2(t) = a_2(t)\cos(w_{2,0} t + w_{2,1} \frac{t^2}{2T})$ with $a_1(t) = 1 + 0.2\sin(w_{A,1} t)$ and $a_2(t) = 0.1 + 0.05\cos(w_{A,2} t)$. See that $x_2(t)$ shows linearly increasing frequency in time as $w_2(t) = w_{2,0} + w_{2,1} \frac{t}{2T}$. In this particular example, the authors choose the following values $f_1 = 5Hz$, $f_{2,0} = 40Hz$, $f_{2,1} = 25Hz$, $f_{A,1} = 1Hz$ and $f_{A,2} = 10Hz$, being $w_. = 2\pi f_.$. We apply CiSSA choosing as window length $L = 200$.

Fig. 8 shows the estimation of the power spectral density of $x(t)$ measured in dB. We clearly see a peak at the normalized frequency 0.005 that corresponds to $x_1(t)$. Regarding the modulated signal described by $x_2(t)$, it appears in the spectrum as a constant gain around a central value. Therefore, CiSSA is able to capture the variability of the frequency by adding several components of adjacent frequencies.

Fig. 9 shows the generated components and their estimations by CiSSA. The first component $x_1(t)$ corresponds to k=2 and its

frequency $w_2 = \frac{k-1}{L}$. The second signal $x_2(t)$ appears as the sum of the reconstructed components associated to the plateau (k=6 to k=17 and their frequencies). This range of adjacent frequencies should be selected as one of the G disjoint groups determined in line 24 of the pseudo-code presented in Algorithm 1. As it can be seen, CiSSA is able to characterize signals with both constant and varying frequencies.

To further discuss and compare the characteristics of the alternative methods, we apply CiSSA, HHT, IHT and HVD to a real data set. HHT works in two steps: first, it decomposes the signal into a small number of intrinsic mode functions (IMF) by means of the Empirical Mode Decomposition (EMD); and, second, it applies the Hilbert transform to the IMFs estimated in the first step to obtain instantaneous frequencies as a function of time. The first component extracted by EMD is highly oscillating and the last component is referred to a constant mean or a trend if there is one, just opposite to SSA since the trend typically corresponds to the leading components of the decomposition. IHT iteratively applies the Hilbert transform to a filtered version of the amplitude envelopes. In a second step, it obtains the instantaneous frequencies by linear regression over time intervals applied to the extracted phases. The number of iterations is the number of estimated components. HVD is also an iterative algorithm. It estimates one component in each iteration by computing the instantaneous frequency and amplitude of the current residual signal. It obtains the instantaneous frequency by applying a low-pass filter to the argument of the Hilbert transform and the instantaneous amplitude by coher-
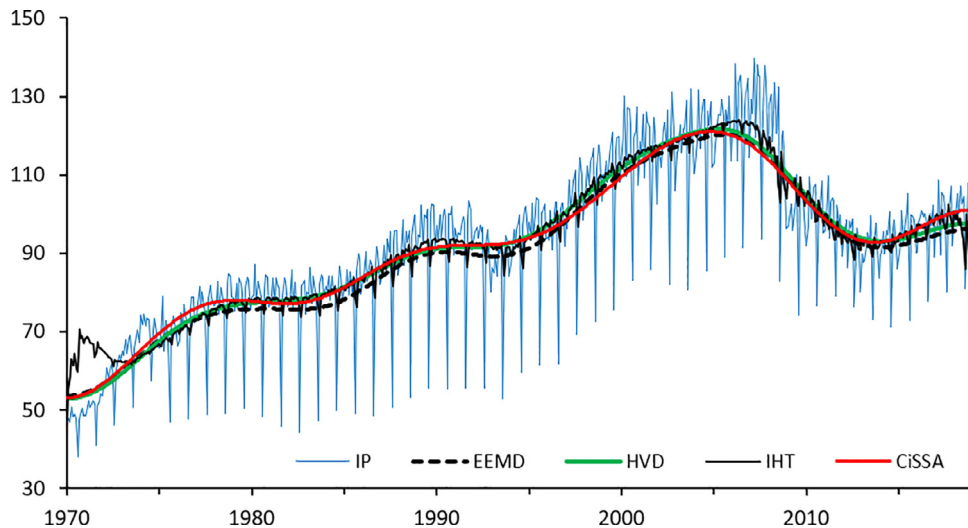
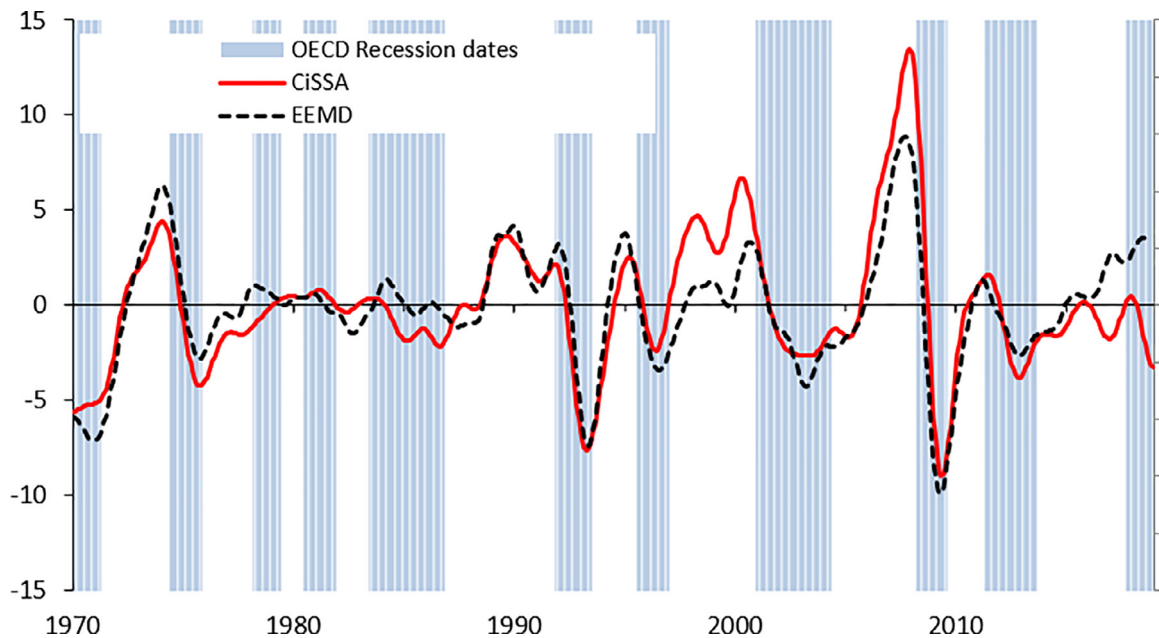**Fig. 10.** Spanish IP and Trends estimated by different methods.

Fig. 11. Spanish IP business cycle estimated by CiSSA and EEMD.

ent demodulation. Biagetti et al. [31] enhance previous HVD versions eliminating the distortions at the beginning and the end of the sample by mirrored convolution.

We consider the monthly Spanish Production Index (IP), 2010=100. The sample covers from January 1970 to December 2018, and the data source is the International Monetary Fund database. As it can be seen in Fig. 10, it is a time series with evolving trend, strong seasonality and business cycle oscillations (between 1.5 and 8 years). Business cycle fluctuations trigger different economic measures from policy makers depending on the state of the economy: expansion or recession. They are characterized by long and slow expansions and short and deep recessions. In this sense, it seems that the analysis made through AM-FM methods can be a good alternative to pick up this component.

We apply CiSSA with a window length $L = 192$. Regarding HHT we use Ensemble EMD (EEMD) [67], implemented in Matlab [68], that solves the mixing-mode problems present in the original technique. Similar results are also obtained with the solutions pro-

posed in [69][6]. For IHT we have used a low pass digital Butterworth filter and the Hilbert transform iteratively. Finally, for HVD we have used the matlab code in [70].

Fig. 10 shows the estimated trend. We can see that all the methods are able to capture the trend in a very similar way, although the trend extracted by IHT shows some remaining seasonality and instability at the beginning and the end of the sample.

IHT and HVD extract components with mixed seasonal and business cycle oscillations. Regarding CiSSA and EEMD, Fig. 11 shows the corresponding estimations of the business cycle as well as recession dates (shadowed areas in blue) estimated by the OECD[7]. See that both signals look quite similar and are able to identify the well-known recessions of the first oil-crisis, the one due to the industrial re-conversion, the recession at the beginning

---

[6] Results are available from the authors upon request.

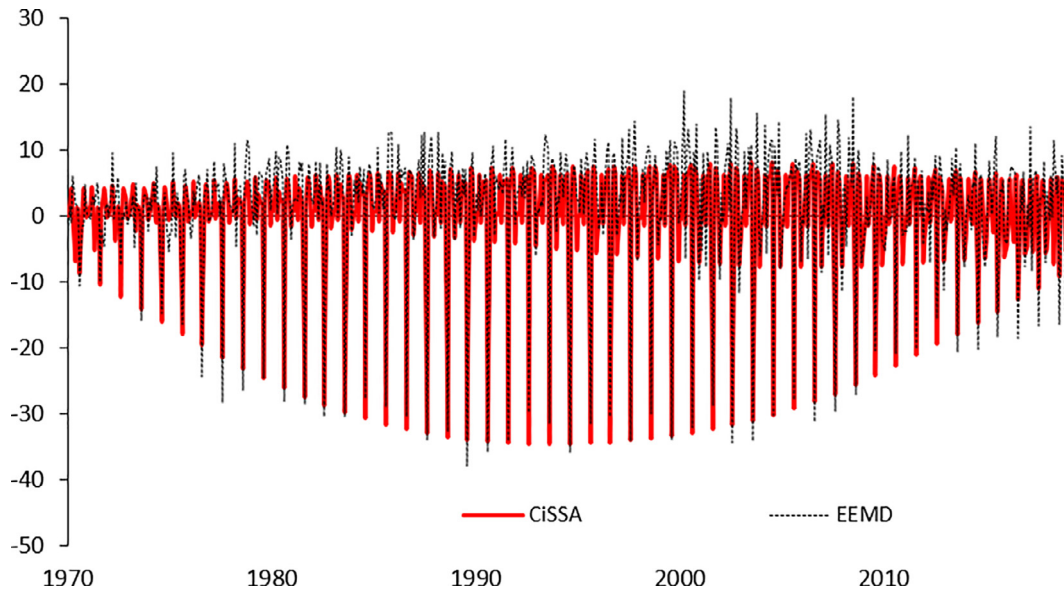[7] https://fred.stlouisfed.org/series/ESPREC

Fig. 12. Spanish IP seasonal component estimated by CiSSA and EEMD.

of the 90s, the one at the beginning of the 21st century and the last Great Recession and subsequent sovereign debt crisis. However, at the end of the sample CiSSA and EEMD are providing a different assessment of the business cycle momentum. While CiSSA, in line with the OECD estimation recession dates, is showing a deceleration of the business cycle, EEMD shows an increasing path. The reason for this difference is the discrepancy in the estimation of the trend at the end of the sample where EEMD could be underestimating the trend given that it is below the local average of the last observations. Fig. 12 shows a very similar estimation of the seasonality by CiSSA and EEMD, though the signal estimated by EEMD seems to be incorporating also short-run noise.

7. Conclusions

In this paper we propose CiSSA, Circulant SSA, an automated procedure that allows to extract the signal associated to any given frequency specified beforehand. This is different to previous versions of SSA that, after extracting the principal components of the trajectory matrix, they need to identify their frequency of oscillation and group them in order to form the desired signals.

CiSSA relies on the eigenstructure of a circulant matrix related to the second moments of the time series. Circulant matrices have closed form solutions for their eigenvalues and eigenvectors. Additionally, we can use them to evaluate the power spectral density at specific frequencies. We prove that CiSSA is asymptotically equivalent to Basic and Toeplitz SSA.

We also extend the algorithm of Circulant SSA to the nonstationary case providing a generalization of Gray's theorem.

The properties of Circulant SSA have been checked through a set of simulations for linear and nonlinear time series models as well as through empirical applications in economics and speech processing. Regarding the application to Industrial Production, we find that CiSSA does a good job extracting the business cycle and estimates cycles that match the dating proposed by the OECD. Within speech processing CiSSA is able to identify the main features of the word under study and to reproduce it.

Finally, we have illustrated the possibility of using CiSSA in the context of time-varying amplitude and frequency and compared its results with other state of the art models commonly used for AM-FM signals.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Theorems and Proofs

The proof of Theorem 1 relies on a set of lemmas and propositions that need to be proven before. Proposition 1 shows the asymptotic equivalence between the Toeplitz matrices of sample and population second moments, $\mathbf{S}_T \sim \mathbf{\Gamma}_L(f)$. Proposition 2 shows that the sequence of matrices $\mathbf{S}_B$ are also asymptotically equivalent to the Toeplitz matrix of population second moments $\mathbf{\Gamma}_L(f)$. We also need two auxilliary lemmas regarding probability convergence of sample and population second moments.

Lemma 3. *For a stationary time series, the sequence* $S_L = \sum_{m=0}^{L-1} (\hat{\gamma}_m - \gamma_m)^2$ *converges in probability to 0 when $L \longrightarrow \infty$.*

Proof. The sum S_L can be decomposed as

$$S_L = \sum_{m=0}^{L-1} \left(\hat{\gamma}_m - \gamma_m \right)^2 = \sum_{m=0}^{L-1} \gamma_m^2 + \sum_{m=0}^{L-1} \hat{\gamma}_m^2 - 2 \sum_{m=0}^{L-1} \gamma_m \hat{\gamma}_m .$$

The first term in the previous equation is finite when $L \longrightarrow \infty$ by Parseval's Theorenm, that is $\sum_{m=0}^{\infty} \gamma_m^2 = K$. Preserving $L < T/2$, L is a monotonically increasing sequence as a function of T so $L \longrightarrow \infty$ when $T \longrightarrow \infty$. Thus if $L \longrightarrow \infty$ means that $T \longrightarrow \infty$ and, therefore, the sum of infinite addends of the second term converges in probability to K given that $\hat{\gamma}_m \longrightarrow \gamma_m$ in probability when $T \longrightarrow \infty$. And, because of the same reasoning, the third term converges in probability to $2K$ when $L \longrightarrow \infty$. As a consequence, the sum S_L converges in probability to $K + K - 2K = 0$ as $L \longrightarrow \infty$. □

Proposition 4. *Let $\{\mathbf{S}_T\}$ and $\{\mathbf{\Gamma}_L(f)\}$ be two sequences of matrices defined in function of the window length defined by (2) and (3) respectively. Then, $\mathbf{S}_T \sim \mathbf{\Gamma}_L(f)$.*

Proof. We know that the eigenvalues of the Toeplitz matrix $\mathbf{\Gamma}_L(f)$ are bounded [47]. The matrices $\mathbf{S}_T$ are Toeplitz and symmetric, therefore their real eigenvalues are also bounded. We must proof

that $\lim_{L \to \infty} \frac{1}{L}\|\mathbf{S}_T - \mathbf{\Gamma}_L(f)\|_F = 0$. We can write

$$0 \le \frac{1}{L}\|\mathbf{S}_T - \mathbf{\Gamma}_L(f)\|_F^2 = \frac{1}{L}\sum_{m=1-L}^{L-1}(L-m)(\widehat{\gamma}_m - \gamma_m)^2 \le$$

$$\le 2\sum_{m=0}^{L-1}\frac{L-m}{L}(\widehat{\gamma}_m - \gamma_m)^2 \le$$

$$\le 2\sum_{m=0}^{L-1}(\widehat{\gamma}_m - \gamma_m)^2.$$

By the *Squeeze Theorem* and the previous Lemma, we obtain that $\lim_{L \to \infty} \frac{1}{L}\|\mathbf{S}_T - \mathbf{\Gamma}_L(f)\|_F = 0$ and therefore it is proved that $\mathbf{S}_T \sim \mathbf{\Gamma}_L(f)$. □

In Basic SSA, it is possible to substitute the matrix $\mathbf{S} = \mathbf{XX}'$ by $\mathbf{S}_B = \mathbf{XX}'/N$ for stationary time series [39]. Matrices $\mathbf{S}$ and $\mathbf{S}_B$, with dimension $L \times L$, have the same eigenvalues and the eigenvectors of $\mathbf{S}_B$ are those of $\mathbf{S}$ multiplied by $1/N$. The elements of matrix $\mathbf{S}_B$ are given by $\widetilde{s}_{ij} = \frac{1}{N}\sum_{t=1}^{i+N-1} x_t x_{t+j-i}$ and, under stationarity, it holds that $\widetilde{s}_{ij}$ converges to $\gamma_{|t-j|}$ as $N \to \infty$, it is, when $T \to \infty$. From matrix $\mathbf{S}_B$ we obtain a sequence of symmetric matrices $\{\mathbf{S}_B\}$ as a function on the window lenght L. To relate this sequence $\{\mathbf{S}_B\}$ of matrices symmetric with the sequence of Toeplitz symmetric matrices $\{\mathbf{\Gamma}_L(f)\}$ we must proof the following Lemma.

Lemma 5. *Under stationarity, the sequence $S_L = \sum_{m=0}^{L-1} \max_{\substack{1\le i, j \le L \\ |i-j|=m}} \{(\widetilde{s}_{ij} - \gamma_m)^2\}$ converges in probability to 0 when $L \to \infty$.*

Proof. The sum S_L verifies that

$$0 \le S_L \le \sum_{m=0}^{L-1}\gamma_m^2 + \sum_{m=0}^{L-1}\max_{\substack{1\le i, j \le L \\ |i-j|=m}}\{\widetilde{s}_{ij}^2\} - 2\sum_{m=0}^{L-1}\min_{\substack{1\le i, j \le L \\ |i-j|=m}}\{\gamma_m\widetilde{s}_{ij}\}.$$

By Parseval's Theorem, the first term on the right is finite as $L \to \infty$, it is quadratic summable, $\sum_{m=0}^{L-1}\gamma_m^2 = K$. We know that $N = T - L + 1$. Given that $L < T/2$, $N > T/2 + 1$ and, further L and N are monotonically increasing sequences as functions of T, so $L, N \to \infty$, as $T \to \infty$. Therefore, if $L \to \infty$ means that $T \to \infty$ and the sum of infinite addends of the second term converges in probability to K because $\widetilde{s}_{ij} \to \gamma_{|i-j|}$, for all i, j, when $N \to \infty$, that is, when $T \to \infty$. And, following the same reasoning, the third term converges in probability to $2K$ when $L \to \infty$. Therefore the right term of the inequality converges to 0 in probability. Finally by the Squeze Theorem, S_L converges in probability to 0. □

Proposition 6. *Let $\{\mathbf{S}_B\}$ and $\{\mathbf{\Gamma}_L(f)\}$ be the sequences of matrices defined as a function of the window length L. Then, $\mathbf{S}_B \sim \mathbf{\Gamma}_L(f)$.*

Proof. The eigenvalues of the Toeplitz matrix $\mathbf{\Gamma}_L(f)$ are bounded [47]. The symmetric matrices $\mathbf{S}_B$ converge to Toeplitz matrix in probability. Then, their eigenvalues are bounded in probability. Now we must proof that $\lim_{L \to \infty} \frac{1}{\sqrt{L}}\|\mathbf{S}_B - \mathbf{\Gamma}_L(f)\|_F = 0$. We can write,

$$0 \le \frac{1}{L}\|\mathbf{S}_B - \mathbf{\Gamma}_L(f)\|_F^2 = \frac{1}{L}\sum_{i=1}^{L}\sum_{j=1}^{L}\left(\widetilde{s}_{ij} - \gamma_{|i-j|}\right)^2 \le$$

$$\le 2\sum_{m=0}^{L-1}\frac{L-m}{L}\max_{\substack{1\le i, j \le L \\ |i-j|=m}}\{(\widetilde{s}_{ij} - \gamma_m)^2\} \le$$

$$\le 2\sum_{m=0}^{L-1}\max_{\substack{1\le i, j \le L \\ |i-j|=m}}\{(\widetilde{s}_{ij} - \gamma_m)^2\}.$$

Therefore, by the Squeeze Theorem and previous Lemma, it holds that $\lim_{L \to \infty} \frac{1}{L}\|\mathbf{S}_B - \mathbf{\Gamma}_L(f)\|_F^2 = 0$ and $\mathbf{S}_B \sim \mathbf{\Gamma}_L(f)$.

Proof of Theorem 1: We have that $\mathbf{S}_T \sim \mathbf{\Gamma}_L(f)$ and $\mathbf{S}_B \sim \mathbf{\Gamma}_L(f)$ by Propositions 4 and 6 respectively, and that together with the transitive property lead to $\mathbf{S}_B \sim \mathbf{S}_T$. Given that by construction $\mathbf{S}_T \sim \mathbf{S}_C$ [50] and, again, by transitive property we have that $\mathbf{S}_B \sim \mathbf{S}_C$.

Proof of Theorem 2: As defined, the function $g(w)$ is real, continuous and 2π-periodic. Its image is $\left[\frac{1}{M_s}, \max_i M_{h_i}\right]$ being different from zero in the whole interval. Then, by the properties of the inverse of Toeplitz matrices $\left(\mathbf{T}_L(g^{-1})\right)^{-1} \sim \mathbf{T}_L(g)$. Moreover, if $F(x)$ is continuos in $\left[\frac{1}{M_s}, \max_i M_{h_i}\right]$, then $F(\frac{1}{x})$ is continuos in $\left[\frac{1}{\max_i M_{h_i}}, M_s\right]$. Since the assumption of $g(w)$ being a Wiener's class function relaxes to a continuous and 2π-periodic function [71], Szegö's theorem leads to (11).

CRediT authorship contribution statement

Juan Bógalo: Conceptualization, Methodology, Software, Validation, Resources, Writing - review & editing, Visualization. **Pilar Poncela:** Conceptualization, Software, Validation, Resources, Writing - original draft, Writing - review & editing, Project administration, Funding acquisition. **Eva Senra:** Conceptualization, Validation, Resources, Writing - original draft, Writing - review & editing, Visualization, Project administration, Funding acquisition.

References

[1] N. Golyandina, A. Zhigljavsky, Singular spectrum analysis for time series, Springer, 2013.
[2] N. Golyandina, Particularities and Commonalities of Singular Spectrum Analysis as a Method of Time Series Analysis and Signal Processing, in: Wiley Interdisciplinary Reviews: Computational Statistics, 2020, p. e1487.
[3] M.A.R. Khan, D.S. Poskitt, Forecasting stochastic processes using singular spectrum analysis: aspects of the theory and application, Int. J. Forecast. 33 (1) (2017) 199–213.
[4] R. Mahmoudvand, P.C. Rodrigues, Missing value imputation in time series using singular spectrum analysis, International Journal of Energy and Statistics 4 (1) (2016) 1650005.
[5] H. Haghbin, S.M. Najibi, R. Mahmoudvand, M. Maadooliat, Functional singular spectrum analysis, 2019, arXiv:1906.05232.
[6] N. Golyandina, A. Korobeynikov, Basic singular spectrum analysis and forecasting with R, Computational Statistics & Data Analysis 71 (2014) 934–954.
[7] S.M.M. Safi, M. Pooyan, A.M. Nasrabadi, Improving the performance of the SSVEP-based BCI system using optimized singular spectrum analysis (OSSA), Biomed. Signal Process. Control 46 (2018) 46–58.
[8] A. Yurova, L.P. Bobylev, Y. Zhu, R. Davy, A.Y. Korzhikov, Atmospheric heat advection in the kara sea region under main synoptic processes, Int. J. Climatol. 39 (2019) 361–374.
[9] U. Kumar, V.K. Jain, Time series models (grey-Markov, grey model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in india, Energy 35 (4) (2010) 1709–1716.
[10] E. Bozzo, R. Carniel, D. Fasino, Relationship between singular spectrum analysis and fourier analysis: theory and application to the monitoring of volcanic activity, Computers and Mathematics with Applications 60 (2010) 812–820.
[11] H. Hassani, D. Thomakos, A review on singular spectrum analysis for economic and financial time series, Stat. Interface 3 (3) (2010) 377=−397.
[12] H. Hassani, S. Heravi, G. Brown, D. Ayoubkhani, Forecasting before, during, and after recession with singular spectrum analysis, J. Appl. Stat. 40 (10) (2013) 2290–2302.
[13] E.S. Silva, H. Hassani, On the use of singular spectrum analysis for forecasting u.s. trade before, during and after the 2008 recession, International Economics 141 (2015) 34–49.
[14] H. Hassani, A.S. Soofi, A. Zhigljavsky, Predicting inflation dynamics with singular spectrum analysis, Journal of the Royal Statistical Society: Series A (Statistics in Society) 176 (3) (2013) 743–760.
[15] H. Hassani, S. Heravi, A. Zhigljavsky, Forecasting UK industrial production with multivariate singular spectrum analysis, J. Forecast. 32 (5) (2013) 395–408.
[16] M. de Carvalho, P.C. Rodrigues, A. Rua, Tracking the US business cycle with a singular spectrum analysis, Econ. Lett. 114 (1) (2012) 32–35.

[17] M. de Carvalho, A. Rua, Real-time nowcasting the US output gap: singular spectrum analysis at work, Int. J. Forecast. 33 (1) (2017) 185–198.

[18] L. Sella, G. Vivaldo, A. Groth, M. Ghil, Economic cycles and their synchronization: a comparison of cyclic modes in three european countries, Journal of Business Cycle Research 12 (1) (2016) 25–48.

[19] J. Arteche, J. García-Enríquez, Singular spectrum analysis for signal extraction in stochastic volatility models, Econometrics and Statistics 1 (2017) 85–98.

[20] S. Lahmiri, Minute-ahead stock price forecasting based on singular spectrum analysis and support vector regression, Appl. Math. Comput. 320 (2018) 444–451.

[21] T. Alexandrov, N. Golyandina, Automatic extraction and forecast of time series cyclic components within the framework of SSA, Proceedings of the Fifth Workshop on Simulation (2005) 45–50.

[22] R. Vautard, P. Yiou, M. Ghil, Singular-spectrum analysis: a toolkit for short, noisy chaotic signal, Physica D 58 (1992) 95–126.

[23] K. Kume, Interpretation of singular spectrum analysis as complete eigenfilter decomposition, Adv. Adapt. Data Anal. 4 (4) (2013).

[24] A.M. Tomé, D. Malafaia, A.R. Teixeira, E.W. Lang, On the use of singular spectrum analysis, 2018, arXiv:1807.10679.

[25] L. Eldén, E. Sjöström, Fast computation of the principal singular vectors of Toeplitz matrices arising in exponential data modelling, Signal Processing 50 (1–2) (1996) 151–164.

[26] A. Korobeynikov, Computation and space-efficient implementation of SSA, Stat. Interface 3 (2009) 357–368.

[27] O. Das, J.S. Abel, J.O. Smith III, Fast MUSIC - an Efficient Implementation of the MUSIC Algorithm for Frequency Estimation of Approximately Periodic Signals, in: Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18), Aveiro, Portugal, September 4–8, 2018, 2018.

[28] H.G. Ma, Q.B. Jiang, Z.Q. Liu, G. Liu, Z.Y. Ma, A novel blind source separation method for single-channel signal, Signal Processing 90 (12) (2010) 3232–3241.

[29] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, E.H. Shih, Q. Zheng, C.C. Tung, H.H. Liu, The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis, in: Proceedings of the Royal Society London, volume A454, 1998, pp. 903–995.

[30] F. Gianfelici, G. Biagetti, P. Crippa, C. Turchetti, Multicomponent AM–FM representations: an asymptotically exact approach, IEEE Trans. Audio Speech Lang. Process. 15 (3) (2007) 823–837.

[31] G. Biagetti, P. Crippa, A. Curzi, S. Orcioni, C. Turchetti, Analysis of the EMG signal during cyclic movements using multicomponent AM-FM decomposition, IEEE J. Biomed. Health Inform. 19 (5) (2015) 1672–1681.

[32] D. Broomhead, G. King, Extracting qualitative dynamics from experimental data, Physica D 20 (1986) 217–236.

[33] D. Broomhead, G. King, On the Qualitative Analysis of Experimental Dynamical Systems, in: A. Hilger (Ed.), Nonlinear Phenomena and Chaos, Bristol, 1986, pp. 113–144.

[34] K. Fraedrich, Estimating the dimension of weather and climate attractors, J. Atmos. Sci. 43 (5) (1986) 419–432.

[35] R. Vautard, M. Ghil, Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, Physica D 35 (1989) 395–424.

[36] D. Danilov, A. Zhigljavsky, Principal components of time series: the "Caterpillar" method, Saint Petersburg Press, Saint Petersburg, 1997. (in Russian)

[37] V.F. Pisarenko, The retrieval of harmonics from a covariance function, Geophys. J. Int. 33 (3) (1973) 347–366.

[38] M.D. Ortigueira, M.A. Lagunas, Eigendecomposition versus singular value decomposition in adaptive array signal processing, Signal Processing 25 (1) (1991) 35–49.

[39] N. Golyandina, V. Nekrutkin, A. Zhigljavsky, Analysis of time series structure: SSA and related techniques, Chapman &Hall/CRC, 2001.

[40] M. Ghil, K. Mo, Intraseasonal oscillations in the global atmosphere – part i and part II, J. Atmos. Sci. 48 (5) (1991) 752–790.

[41] T. Alexandrov, N. Golyandina, The automatic extraction of time series trend and periodical components with the help of the caterpillar-SSA approach, Exponenta Pro 3–4 (2004) 54–61.

[42] F.J. Alonso, D.R. Salgado, Analysis of the structure of vibration signals for tool wear detection, Mech. Syst. Signal Process. 22 (3) (2008) 735–748.

[43] M. Bilancia, F. Campobasso, Airborne Particulate Matter and Adverse Health Events: Robust Estimation of Timescale Effects, in: Classification as a Tool for Research, Springer, Berlin Heidelberg, 2010, pp. 481–489.

[44] M.S. Solary, Finding eigenvalues for heptadiagonal symmetric Toeplitz matrices, J. Math. Anal. Appl. 402 (2013) 719–730.

[45] P. Lancaster, Theory of matrices, Academic Press, NY, 1969.

[46] R.M. Gray, On the asymptotic eigenvalue distribution of Toeplitz matrices, IEEE Transactions on Information Theory 18 (6) (1972) 725–730.

[47] P. Tilli, Singular values and eigenvalues of non-hermitian block Toeplitz matrices, Linear Algebra Appl. 272 (1–3) (1998) 59–89.

[48] U. Grenander, G. Szegö, Toeplitz forms and their applications, University of California Press, Berkeley and Los Angeles, 1958.

[49] W.F. Trench, Absolute equal distribution of the spectra of hermitian matrices, Linear Algebra Appl. 366 (2003) 417–431.

[50] J. Pearl, On coding and filtering stationary signals by discrete fourier transform, IEEE Trans. on Info. Theory IT-:19 (1973) 229–232.

[51] M. Allen, L. Smith, Monte carlo SSA: detecting irregular oscillations in the presence of colored noise, J. Clim. 9 (1996) 3373–3404.

[52] M. Ghil, R.M. Allen, M.D. Dettinger, K. Ide, D. Kondrashov, M.E. Mann, A. Robertson, A. Saunders, Y. Tian, F. Varadi, P. Yiou, Advanced spectral methods for climatic time series, Rev. Geophys. 40 (1) (2002) 1–41.

[53] R.M. Gray, On unbounded Toeplitz matrices and nonstationary time series with an application to information theory, Information and Control 24 (1974) 181–196.

[54] N. Golyandina, Statistical approach to detection of signals by monte carlo singular spectrum analysis: Multiple testing, arXiv:1903.01485 (2019).

[55] P.C. Young, Recursive estimation and time series analysis: An introduction, Springer, Verlag, Berlin, 1984.

[56] J. Durbin, S.J. Koopman, Time series analysis by state space methods, 2nd, Oxford University Press, 2012.

[57] J.P. Burman, Seasonal adjustment by signal extraction, J. R. Stat. Soc. Ser. A (1980) 321–337.

[58] E.B. Dagum, Modelling, forecasting and seasonally adjusting economic time series with the x-11 ARIMA method, Journal of the Royal Statistical Society. Series D (The Statistician) 27 (3/4) (1978) 203–216.

[59] D.F. Findley, D.P. Lytras, T.S. Mc Elroy, Detecting seasonality in seasonally adjusted monthly time series, Statistics (Ber) (2017) 3.

[60] B.R. Moulton, B.D. Cowan, Residual seasonality in GDP and GDI: findings and next steps, Survey of Current Business 96 (7) (2016) 1–6.

[61] J. Lothian, The Identification and Treatment of Moving Seasonality in the X-11-ARIMA Seasonal Adjustment Method, Business Finance Division. Statistics Canada, 1978. Research Paper

[62] F. Gianfelici, G. Biagetti, P. Crippa, C. Turchetti, AM–FM Decomposition of speech signals: an asymptotically exact approach based on the iterated Hilbert transform, IEEE Workshop on Statistical Signal Processing Proceedings (2005) 333–337.

[63] N.E. Huang, M.L. Wu, W. Qu, S.R. Long, S.S. Shen, Applications of Hilbert-Huang transform to non-stationary financial time series analysis, Appl. Stoch. Models Bus. Ind. 19 (3) (2003) 245–268.

[64] S.C. Pei, K.W. Chang, The mystery curve: a signal processing point of view [lecture notes], IEEE Signal Process. Mag. 34 (6) (2017) 158–163.

[65] M. Feldman, Analytical basics of the EMD: two harmonics decomposition, Mech. Syst. Signal Process. 23 (7) (2009) 2059–2071.

[66] M. Feldman, Time-varying vibration decomposition and analysis based on the Hilbert transform, J. Sound Vib. 295 (3–5) (2006) 518–530.

[67] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (01) (2009) 1–41.

[68] M.E. Torres, M.A. Colominas, G. Schlotthauer, P. Flandrin, A complete ensemble empirical mode decomposition with adaptive noise, 2011, IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP-11, 4144–4147.

[69] R.T. Rato, M.D. Ortigueira, A.G. Batista, On the HHT, its problems, and some solutions, Mech. Syst. Signal Process. 22 (6) (2008) 1374–1394.

[70] M. Feldman, Hilbert transform application in mechanical vibration, John Wiley and Sons, 2011.

[71] J. Gutiérrez-Gutiérrez, P.M. Crespo, Block Toeplitz matrices: asymptotic results and applications, Foundations and Trends in Communications and Information Theory 8 (3) (2012) 179–257.