UCLA
UCLA Previously Published Works

Title
Data-adaptive wavelets and multi-scale singular-spectrum analysis

Permalink
https://escholarship.org/uc/item/5v1803mp

Journal
Physica D Nonlinear Phenomena, 142(3-4)

ISSN
0167-2789

Authors
Yiou, Pascal
Sornette, Didier
Ghil, Michael

Publication Date
2000-08-01

DOI
10.1016/s0167-2789(00)00045-2

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

arXiv:chao-dyn/9810034v1 29 Oct 1998

Data-Adaptive Wavelets and Multi-Scale SSA

Pascal Yiou[1,2], Didier Sornette[3,4] and Michael Ghil[2]

Institute of Geophysics and Planetary Physics
University of California Los Angeles at Los Angeles
Los Angeles, CA 90095-1567

1. Permanent address : Laboratoire des Sciences du Climat et de l'Environnement, UMR CEA–CNRS, CEA Saclay, 91191 Gif-sur-Yvette, France
2. Also Department of Atmospheric Sciences, UCLA
3. Also Department of Earth and Space Sciences, UCLA
4. Also Laboratoire de Physique de la Matière Condensée, CNRS UMR 6622 and Université de Nice-Sophia Antipolis, 06108 Nice Cedex 2, France

Physica D, submitted October 26, 1998.

Abstract

Using multi-scale ideas from wavelet analysis, we extend singular-spectrum analysis (SSA) to the study of nonstationary time series of length N whose intermittency can give rise to the divergence of their variance. The wavelet transform is a kind of local Fourier transform within a finite moving window whose width W, proportional to the major period of interest, is varied to explore a broad range of such periods. SSA, on the other hand, relies on the construction of the lag-covariance matrix $\mathbf{C}$ on M lagged copies of the time series over a fixed window width W to detect the regular part of the variability in that window in terms of the minimal number of oscillatory components; here $W = M\Delta t$, with Δt the time step. The proposed multi-scale SSA is a local SSA analysis within a moving window of width $M \leq W \leq N$. Multi-scale SSA varies W, while keeping a fixed W/M ratio, and uses the eigenvectors of the corresponding lag-covariance matrix $\mathbf{C}_M$ as a *data-adaptive* wavelets; successive eigenvectors of $\mathbf{C}_M$ correspond approximately to successive derivatives of the first mother wavelet in standard wavelet analysis. Multi-scale SSA thus solves objectively the delicate problem of optimizing the analyzing wavelet in the time-frequency domain, by a suitable localization of the signal's covariance matrix. We present several examples of application to synthetic signals with fractal or power-law behavior which mimic selected features of certain climatic and geophysical time series. A real application is to the Southern Oscillation index (SOI) monthly values for 1933–1996. Our methodology highlights an abrupt periodicity shift in the SOI near 1960. This abrupt shift between 4 and 3 years supports the Devil's staircase scenario for the El Niño/Southern Oscillation phenomenon. [1]

Contents

1 Introduction and motivation **3**

2 Methodology **4**
2.1 Wavelet transform 4
2.2 Singular spectrum analysis (SSA) 5
2.3 Multi-scale SSA (MS-SSA) 6

3 Data sets **8**
3.1 Synthetic data sets 8
3.2 Climatic data set 9

4 Numerical results **9**
4.1 Cantor-set experiments 9
  4.1.1 SSA analysis 9
  4.1.2 Wavelet transforms 10
  4.1.3 MS-SSA analysis 10
4.2 A log-periodic process 12
4.3 Multiplicative noise 13
4.4 Southern Oscillation Index 14

5 Concluding remarks **15**

6 Acknowledgments **16**

7 Appendix: Construction of the synthetic time series **17**
7.1 Approximate triadic Cantor set 17
7.2 The exact triadic Cantor set 17
7.3 Log-periodic time series 18
7.4 The statistically self-similar process P_3 18

8 Figure captions **23**

---

[1] Preliminary results of this study were presented at the XXII General Assembly of the European Geophysical

1 Introduction and motivation

Intermittent time behavior is common in a large variety of systems in the natural and socio-economic sciences. Multiple regimes of behavior and the transitions between them are ubiquitous in climate dynamics, from the alternation of glacials and interglacials to that of atmospheric blocking [1]. River discharge and rainfall similarly exhibit long periods of relative quiescence interspersed with large floods and bursts [2]. The solid earth is characterized by even stronger intermittency that manifests itself by long tails of earthquake-size distribution with strong clustering of seismic activity in time and space [3]. Even mantle convection is now believed to be intensely intermittent, with internal boundary-layer instability leading to "avalanches" that come in a large range of sizes and follow a power-law distribution [4]. The number of contagious-disease incidences is also intermittent, and has recently been documented to be distributed according to power laws [5]. The evolutionary path of species on earth with its "punctuated dynamics" provides a paradigm of intermittent behavior [6]. Economic and financial time series have long been recognized to have nonstationary variances, possibly described by Lévy processes [2].

These are only a few examples among many that exhibit the ubiquity of intermittent behavior, and hence of extreme events. Additional examples range from volcanic eruptions, hurricanes and tornadoes, landslides, avalanches, lightning strikes, and meteorite and asteroid impacts, to the failure of engineering structures, crashes in the stock market, social unrest leading to large-scale strikes and upheaval, economic drawdowns on national and global scales, regional power blackouts, and traffic gridlock.

Many of these problems involve nonstationary time series, while standard methods of time-series analysis assume explicitly or implicitly stationarity. Stationarity in the wide sense is defined by (i) the existence of first (mean) and second (variance) moments of the time series and (ii) their invariance with respect to time translations. The failure of either assumption can lead to spurious spectral-analysis results.

Wide-sense stationarity implies that the covariance operator of a time series exists and is a function of the separation between epochs only, *i.e.*, of "lags." The spectral decomposition of the resulting lag-covariance operator is connected to the power spectrum of the time series through the Wiener-Khinchin theorem [7, 8]. For time series in continuous time and of infinite length, the corresponding eigenfunctions are always complex exponentials or sine-and-cosine pairs, and one time series differs from another solely through the associated eigenvalues.

For a time series of given, finite length — in discrete time — singular-spectrum analysis (SSA) [9, 10] relies on the Karhunen-Loève decomposition of an estimate of the covariance matrix that is based on M lagged copies of the time series. The eigenvectors of this matrix are often called empirical orthogonal functions, or EOFs and differ, in general, from sines and cosines [11]. For the given time series, the EOFs form an optimal basis that is orthonormal at zero lag; they permit one to decompose the signal into its, possibly anharmonic, oscillatory components and aperiodic ones. The data-adaptive character of the SSA basis functions presents an advantage over the standard Fourier basis of the classical Blackman-Tukey [7, 8] method: it allows one to capture a nonlinear oscillation by a single pair of EOFs — that maximizes the variance associated with the given frequency — rather than by a number of sine-and-cosine pairs associated with that basic frequency and its harmonics.

SSA seems to be less sensitive than classical Fourier analysis to deviations of a given time series from wide-sense stationarity assumptions. Trends are captured quite successfully, most often by the leading one or two EOFs of the given time series [11, 12] and simple discontinuities in the local value of an oscillation's frequency can also be detected by SSA [13, 14].

Wavelet analysis, on the other hand, has become a tool of reference for intermittent, complex and self-similar signals, because it works as a mathematical microscope that can focus on a specific part of the signal to extract *local* structures and singularities [15, 16, 17]. The first step in the definition of a wavelet transform is the choice of the analyzing function, often called "mother wavelet." Considerable work has been devoted to finding mother wavelets that provide an orthogonal expansion that is complete but not (highly) redundant. It is also often desirable to concentrate the spectral energy in an optimal band for the problem at hand, while keeping the localization property.

A large number of analyzing wavelet functions have been introduced to satisfy these often conflicting requirements, with their relative advantages and drawbacks. To provide an optimal multi-scale decomposition of a signal, it may be desirable to (A) automatically adjust the shape of the analyzing wavelet to the signal rather than search through the extensive "libraries" of mother wavelets (http://www.mathsoft.com/wavelets.html) and (B) modify this shape in time and scale, especially if the data set's nonstationarity implies different structures as one slides along time or scale. The intrinsic constraint of a unique mother wavelet does not allow for this flexibility.

We introduce here a simple and powerful resolution to the limitations of both types of methodology. The standard forms of the wavelet transform and SSA are recalled in Secs. 2.1 and 2.2 respectively. The SSA method – modified to use varying windows with width W proportional to the order M of the lag-covariance matrix $\mathbf{C}$ – provides data-adaptive wavelet transforms with analyzing functions given by the leading EOFs of the matrix $\mathbf{C}$; it is described in Sec. 2.3. A set of synthetic time series that exhibit intermittent and self-similar properties, to wit, a Cantor set, a log-periodic process, and a multiplicative noise, are introduced in Sec. 3.1; the details of their construction appear in Appendices 7.1-7.4. We also introduce a climatic time series, as a "real-world" test set. The multi-scale SSA (MS-SSA) of Sec. 2.3 is applied to these four test sets in Secs. 4.1 through 4.4, respectively.

MS-SSA provides a new approach for testing self-similarity and its breakdown by comparing the data-adaptive wavelets at different scales, while it retains the usual scaling of wavelet coefficients which is obtained from any wavelet transform [18]. A summary of our results and comparison with complementary approaches to address the requirements (A) and (B) above appear in Sec. 5.

2 Methodology

2.1 Wavelet transform

A wavelet transform requires the choice of an analyzing function ψ, with general admissibility properties [15, 16, 17], and with the more specific property of time and frequency localization, *i.e.*, ψ and its Fourier transform $\hat{\psi}$ must decay rapidly outside a given interval. Functions based on a Gaussian, $f(x) = \exp(-x^2)$, possess the localization property even if they do not verify the admissibility condition that the integral of the mother wavelet vanish. This zero-mean condition ensures that knowledge of all position- and scale-dependent wavelet coefficients allows one to retrieve the initial function through an inversion formula (similar to a two-parameter inverse Fourier transform) [16].

In the sequel, we shall follow [18] and use a Gaussian wavelet and its first derivative for the sake of simplicity. The Gaussian itself does not satisfy the admissibility condition, while its first derivative does. The overall statistical characterization of complex structures depends only weakly on the choice of the mother wavelet [18] and we choose this simple example to illustrate its multi-scale low-pass filter properties and contrast them to the proposed MS-SSA technique.

A ψ-wavelet transform W_ψ in continuous time and frequency is simply a projection of a signal $X(t)$, $-\infty < t < \infty$, onto b-translated and a-dilated versions of ψ:

$$
\begin{aligned}
W_\psi(a,b) &= \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} X(t)\psi\left(\frac{t-b}{a}\right) dt &\quad (1)\\
&= \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} X(t+b)\psi\left(\frac{t}{a}\right) dt. &\quad (2)
\end{aligned}
$$

If most of ψ is concentrated between, $[-1, 1]$ say (up to a rescaling), then Eq. (1) is clearly an analysis of X in the interval $[b-a, b+a]$, where the integral is nonvanishing *a priori*.

Using the successive derivatives $\psi^{(n)}$ of a given wavelet ψ in Eq. (1) is equivalent to a ψ-analysis (up to a normalization factor) of the successive derivatives of the time series X; this is easy to see through an integration by parts. It is also easy to show, for a wide class of smooth analyzing functions ψ, that the number of oscillations in $\psi^{(n)}$ increases with n; this is true, in particular for our Gaussian wavelet of choice.

The original signal (or a filtered version of it) can be reconstructed from the family of wavelet transforms. Hence, for scale values a in an interval I, a reconstructed version X_I of the signal $X(t)$ is:

$$
X_I(t) = A_\psi \int_{a\in I} \int_{b=-\infty}^{\infty} W_\psi(a,b)\psi\left(\frac{t-b}{a}\right) \frac{da\,db}{a^2}; \quad (3)
$$

A_ψ is a normalization factor which only depends on the mother wavelet ψ. This formulation is essentially a bandpass filter of X through I. If I is the positive real line, $I = \mathbf{R}^+$, then $X_I = X$. Formula (3) applies when the mother wavelet is the derivative of the Gaussian, which we use below. As mentioned previously, the Gaussian $\psi(x) = \exp(-x^2/2)$ itself cannot be used in this reconstruction formula, because it does not satisfy $\int \psi(x)dx = 0$. However, the forward transform of Eq. (2) is well-defined and provides nevertheless a useful multi-scale decomposition of the signal.

The counterpart of Eqs. (1) and (3) for discrete time series can be obtained in a straightforward way. Given a discrete signal $X_i = X(t_i)$, $t_i = i\Delta t$, $0 \leq i \leq N$, there are numerous ways of estimating Eq. (1) numerically that avoid the problems of high-frequency noise and aliasing, while minimizing computational cost [17, 19]. Computationally efficient algorithms based on the fast Fourier transform can be used [17], but we will only need here the simplest numerical quadrature formulae to process Eq. (1).

2.2 Singular spectrum analysis (SSA)

SSA has been described theoretically, for time-continuous signals $X(t)$, by Vautard and Ghil [10]. Since it is still less familiar to a general readership than the wavelet transform, we summarize here its form for time-discrete signals $\{X_i = X(i\Delta t),\ 0 \leq i \leq N\}$, which is generally used in practice, as it will be in the present study. Methodological and review papers include [11, 13, 21, 22]. SSA proceeds by finding the eigenvectors, called EOFs, of the lag-covariance matrix $\mathbf{C}$ of the time series $\{X_i\}$ as above. The true lag-covariance matrix $\tilde{\mathbf{C}}$ of the process that generated $\{X_i : 1 \leq i \leq N\}$ is estimated by a lag-covariance matrix $\mathbf{C}_M$; assuming stationarity of X yields a Toeplitz structure

$$\mathbf{C}_{M,i} = \frac{1}{N-i} \sum_{j=1}^{N-i} X_j X_{j+i}, \tag{4}$$

where i $(0 \leq i \leq M-1)$ is the diagonal index of the matrix, and the time series X is normalized to have zero mean.

The eigenvectors of this symmetric matrix are orthonormal and provide the Karhunen-Loève basis for expanding the time series X_i with respect to its M lagged copies. These EOFs ρ_k of $\mathbf{C}_M$ are sorted by decreasing order of the associated eigenvalue λ_k, $0 \leq \lambda_M \leq \cdots \leq \lambda_2 \leq \lambda_1$. The principal component (PC) $\mathbf{a}_k = \{a_{ki} : 0 \leq i \leq N-M\}$ is computed by projecting the time series onto $\rho_k = \{\rho_{kj} : 0 \leq j \leq M\}$:

$$a_{ki} = \sum_{j=1}^{M} X_{i+j-1}\rho_{kj}. \tag{5}$$

For a given k, this equation is formally similar to Eq. (2), in which the integration interval is essentially restricted to the domain over which the wavelet function is numerically nonnegligible. Furthermore, the EOFs ρ_k have an oscillation property which resembles that of the derivatives $\psi^{(n)}$: the higher the order k, the more sign changes ρ_k has [10, 11]. This oscillation property is due to the formal analogy between the Toeplitz structure of the covariance matrix $\mathbf{C}_M$ and that of finite-difference discretizations of Sturm-Liouville problems on an interval $W = M\Delta t$ [20]. In our case, Δt is the sampling interval of the time series; in the discretized Sturm-Liouville problem, it is the mesh size. In both cases, the true operator being approximated is positive definite (or, more precisely, non-negative semi-definite), *i.e.*, its eigenvalues are all positive (possibly one or more, but not all, being zero).

As in wavelet and Fourier analyses, time-dependent reconstructions of the entire signal $\{X_i\}$ or of a suitable filtered version thereof are possible from its PCs and EOFs. Hence, for a given set of eigenelements $\mathcal{K}$, the corresponding reconstructed component (RC) is [11]:

$$r_{\mathcal{K},i} = \frac{1}{M_i} \sum_{k\in\mathcal{K}} \sum_{j=1}^{M} a_{k,i-j}\rho_{kj}, \tag{6}$$

where M_i is a scaling factor equal to M when i is sufficiently far from the end points, and contains corrections when i approaches the interval's endpoints to within a distance $M\Delta t$

2.3 Multi-scale SSA (MS-SSA)

Systematic comparisons between SSA, the wavelet transform and other spectral-analysis methods have been carried out in Ref. [13, 23] (see, in particular, Table 1 in [23]). Further analogies between certain mathematical features of SSA and wavelet analysis were mentioned in [24]. Table 1 here summarizes the most useful mathematical parallels between the two time-series analysis methods; these parallels provide the basis for extending global SSA analysis to a local one. In SSA, the largest scale at which the signal X is analyzed in Eq. (5) is approximately N (the length of the time series), and the largest period is M. As a consequence, the EOFs ρ_k contains information from the whole time series, as in the Fourier transform.

Method	SSA	Wavelet transform
Analyzing function	EOF ρ_k	Mother wavelet ψ
Basic facts	ρ_k eigenvectors of C_M	ψ chosen *a priori*
Decomposition	$\sum_{t'=1}^{M} X(t+t')\rho_k(t')$	$\int X(t)\psi\left(\frac{t-b}{a}\right)dt$
Scale	$W = \alpha M$	a
Epoch	t	b
Average & trend	ρ_1	$\psi^{(0)}$
Derivative	ρ_2	$\psi^{(1)}$

Table 1: Analogy between SSA and the wavelet transform.

In order to define a local SSA, we propose to extend the SSA methodology by using a time-frequency analysis within a running time window whose size W is proportional to the order M of the covariance matrix. Varying M, and thus W in proportion, we obtain a multi-scale representation of the data. We perform *local* SSA on a time series by sliding windows of length $W \leq N$, centered on times $b = W/2, \ldots, N - W/2$. When using this method, we assume that considerable information content resides in the local variance structure and that the time series is locally the sum of a trend, statistically significant variability, and noise.

A priori, we can vary the two scales W and M independently, as long as W is larger than M, $W/M \geq \alpha > 1$, and α is large enough [11]. In the wavelet transform, however, for instance the Mexican hat which corresponds to the second derivative of the Gaussian function, the number of oscillations of the mother wavelet is fixed and independent of the scale (width) of the analyzing wavelet. In this spirit, we fix the ratio $W/M = \alpha$ and rely therewith on the oscillation property of the EOFs described in Sec. 2.2 above to provide a fixed number of zeroes for the data-adaptive "wavelet" ρ_k of each local SSA analysis; $\alpha = 3$ in the calculations presented below and, as we shall see, it will in fact suffice to use $k = 1$ or $k = 2$ on each W-interval. This provides an analysis at a fixed scale W. Sampling a set of W values that follow a geometrical sequence, for instance in powers of 2 or 3, provides a multi-scale analysis very similar to the wavelet transform.

For a given position b and fixed W, we obtain local EOFs, for which PCs or RCs can be computed as described in Sec 2.2. The EOFs are the direct analogs of analyzing wavelet functions, as summarized in Table 1. Similarly to successive analyzing wavelets (for instance the derivatives of the Gaussian function), the number of EOF oscillations increases roughly with order and the zeroes of ρ_{k+1} separate those of ρ_k. This is consistent with the orthogonality of distinct EOFs, with their being even or odd about the mid-point of the window, and with their tending in certain cases to pairs of sines and cosines [10] or of Legendre polynomials [25]. The first EOF thus corresponds approximately to an analyzing wavelet function with a single extremum and no zero, for instance the Gaussian. The second EOF is reminiscent of the first derivative of the Gaussian and so on.

Then for each b and each EOF ρ_k, it is possible to obtain local PCs a_k and RCs r_k. The kth

PC at time b is

$$a_{ki}^{b} = \sum_{j=1}^{M} X_{i+j-1}\rho_{kj}^{b}, \tag{7}$$

and the corresponding RC is

$$r_{ki}^{b} = \frac{1}{M_i}\sum_{j=1}^{M} a_{k,i-j}^{b}\rho_{kj}^{b}, \tag{8}$$

with $b - W/2 \leq i \leq b + W/2$. The crucial difference between this local version and global SSA is that the RCs are obtained here from local lag-covariance matrices. As b varies from $W/2$ to $N - W/2$, this implies that the RCs will be truncated near the edges of the time series.

We thus see that the local SSA method provides simultaneous "wavelet transforms" of the data by a set of analyzing wavelet functions, corresponding to the M different EOFs of the lag-covariance matrix. When W (and thus M) is small, local SSA provides a small-scale analysis of the signal with a few distinct analyzing functions, namely a (sub)set of EOFs indexed by k. This is reasonable as there are not many possible structures at scales that approach the sampling time scale. On the other hand, at large scales, local SSA can also provide the simultaneous analysis by many different analyzing mother wavelet functions, $\{\rho_k : 1 \leq k \leq M\}$, and thus reflect the large possible complexity of the structures that can develop over the entire time series.

The most important property of this local SSA analysis is that the analyzing functions (analogous to the wavelets) are *data adaptive*: they are just the EOFs of the lag-covariance matrix with time lags up to M, within a time window of size $W = \alpha M$. In other words, the shape of these analyzing functions is not imposed *a priori*, like in a wavelet analysis, but explicitly depends on the time series itself. This property is particularly suitable in analyzing time series that exhibit dominant structures which differ along the signal. For instance, an oscillatory behavior could be followed by white or colored noise and then by deterministically intermittent behavior; this could indicate regime transitions that the system which generates the signal underwent while under observation. If so, one would like to have an analyzing wavelet which is adapted to each section of the signal. Our data-adaptive scheme will definitely help follow such regime transitions in time.

In addition, there is information to be gained from the shape of the EOFs that are provided by the data. Wavelet transforms are sometimes used to test for self-similarity: one tests for the existence of a power-law dependence of a wavelet coefficient at a given time as a function of scale. If the power law is observed, one concludes on the existence of a local singularity associated with such a law. This is useful in analyzing multifractal structures characterized by a set of distinct singularities [18].

The approach presented here provides an additional test. Indeed, if a signal is locally self-similar, the shape of the EOFs must also be the same, as we change scale (*i.e.*, M and W). We can thus carry out an analysis at different scales, obtain the data-adaptive analyzing functions and test for their self-similarity. This possibility arises naturally out of our multi-scale SSA (MS-SSA) method and does not exist in any technique that assumes *a priori* the shape of the analyzing function, however carefully selected. We present a few tests of MS-SSA below, and compare the results with those of a standard wavelet analysis.

The interesting properties of MS-SSA come however at a significant computational cost. Indeed, the numerical cost for matrix diagonalization far exceeds that of fast Fourier transforms or simple algorithms for wavelet analysis, even when several wavelet functions are used. In order to circumvent this problem and reduce the computational cost, we carry out the local SSA analysis only for a subset of time steps, $b = \beta\Delta t$. This sampling interval for the analysis must be substantially smaller than the scale parameter W, *i.e.* $\beta \ll W$. Then we use Eq. (8) for $b \neq \beta\Delta t$ to obtain the local RCs where the MS-SSA algorithm is not applied. We have checked that this interpolation procedure is very accurate when compared to exhaustive local SSA analysis at each time step, *i.e.* computing r_{kb}^{b} in Eq. (8) for each b and hence discarding its values at epochs surrounding b. This procedure is based on the relative robustness of the EOFs, which is also used in SSA-based time-series prediction [11, 26, 27].

3 Data sets

3.1 Synthetic data sets

We would like to investigate the behavior of MS-SSA on time series that are self-similar or fractal, as they pose *a priori* the biggest challenge to techniques based on well-behaved finite covariances. As already mentioned, such features are ubiquitous in the geosciences [28].

The first time series $\tilde{P}_1$ that we shall investigate was obtained by using an Iterated Function System (IFS) [29] to approximate a simple triadic set. The procedure that we have used is described in Appendix 7.1 and the resulting time-series is shown in Fig. 1. A closely related time series P_1 is obtained from the characteristics of the exact triadic Cantor set. Its construction is described in Appendix 7.2. In the "perfect" construction scheme of P_1, we anticipate large "resonances" in the lag-covariance matrix that should be smoothed out by the noise present in the construction of $\tilde{P}_1$. Finally, a multiple-rule Cantor set is generated in Appendix 7.2 to yield the time-series $\hat{P}_1$.

Figure 1: A 1000-point realization of the $\tilde{P}_1$ process (top) and its blow-up (bottom).

The second process is a log-periodic signal with a square-root singularity. In the context of critical phenomena, the log-periodicity corresponds to the existence of an imaginary part in the critical exponent, and is associated with a discrete scale invariance [30], *i.e.* to the invariance of the system or of its properties only under magnifications that are integer powers of a fundamental ratio. Initially, complex exponents were proposed as formal solutions of renormalization-group equations in the seventies [31, 32, 33]. In the eighties, they have been shown to emerge in various physical problems that arise in discrete hierarchical systems [34, 35].

Recently, it has been realized that discrete scale invariance and its associated complex exponents and log-periodicity may appear "spontaneously," without the need for a pre-existing hierarchy, as a result of dynamical processes. Such behavior has been found in models and experimental data from irreversible growth processes, rupture in heterogeneous systems, earthquakes, percolation and even in financial crashes [36] (see [30] for a review, and references therein). Another source of time series with log-periodic behavior is provided by Boolean delay equations, whose solutions are Boolean-valued processes in continuous time [37]. These equations have been used to model paleoclimatic variability [40, 38], interdecadal [39] and most recently, seasonal-to-interannual [41] climate change, although not all of these applications exhibit the behavior in question. In the case of Boolean delay equations, the log-periodicity exhibits a period and amplitude that increase as $t \rightarrow t_c = +\infty$ [37] and their treatment by MS-SSA is left for subsequent study.

The log-periodicity we concern ourselves with here is reflected in oscillations of decreasing period and amplitude that are superimposed on a power-law behavior and culminate at a singularity t_c in finite time. At the critical point t_c, the instantaneous frequency of the oscillation becomes infinite, but its amplitude vanishes. Again, a global spectral analysis method would fail to represent this signal in a satisfactory manner since its nonstationarity is characterized by a continuously varying period and a nonlinear trend. The log-periodic time series P_2 is constructed according to the procedure given in Appendix 7.3 and is shown in Fig. 2.

Figure 2: Log-periodic time series P_2.

While the first two processes $(P_1, \tilde{P}_1)$ and P_2 are purely deterministic, the third process we have studied still presents self-similar properties, albeit with a large stochastic component. Its self-similarity stems from its multiplicative-noise nature. The process P_3 has the property of being statistically stable under affine transformations [42, 43]. A 512-point realization of P_{3i}, for $1 \leq i \leq 512$, is shown in Fig. 3 and a histogram is provided to illustrate the power-law behavior and its long tail.

Figure 3: Multiplicative-noise time series P_3.

3.2 Climatic data set

In addition to the synthetic data sets that possess well-understood properties, we chose to analyze a "real-world" climatic time series, the Southern Oscillation Index (SOI). This time series illustrates well another type of behavior often encountered in the geosciences, viz., the superposition of quasi-periodic and aperiodic behavior [27, 44]. SOI is a climatic index connected with the recurring El Niño/Southern Oscillation (ENSO) conditions in the tropical Pacific; it is essentially the monthly mean difference in sea-level pressure between Darwin, Australia, and Tahiti (Fig. 4). An anomalously negative value of this index indicates a warm ENSO event (El Niño), while a highly positive value is associated with a cold event (La Niña). SOI has been used in numerous studies for understanding the dynamics of ENSO [46] and improving its prediction [26, 27].

In the data set we use here, the annual cycle was removed and the time series was normalized by its variance. The time interval considered goes from January 1933 to December 1996, during which very few observations are missing at either station.

Figure 4: SOI variations between 1933 and 1996.

ENSO models have shown that, in the Tropical Pacific, phase locking of the coupled ocean-atmosphere system's self-sustained oscillation to the annual cycle leads to a Devil's staircase [47]. This behavior has been confirmed by studies of sea-surface temperature data sets using both classical (global) SSA and (fixed-analyzing function) wavelet methods. Such studies have been performed on the sea-level pressure record at Darwin only [48], which spans a longer time interval, using wavelet and waveform analysis, and Tropical Pacific sea-surface temperatures with global SSA [14, 45, 46]. Two of these analyses have shown a fairly sudden shift in the characteristic period of ENSO's low-frequency component that occurred in the 1960s, with a longer period before and a shorter one prevailing since [14, 48].

4 Numerical results

4.1 Cantor-set experiments

4.1.1 SSA analysis

Global methods of time-series analysis will clearly miss the multi-scale feature of the IFS construct $\tilde{P}_1(t)$, and essentially see one or more nearly periodic components that arise from a smoothed version of the exact Cantor set $P_1(t)$. A Monte-Carlo SSA [22, 49] of $\tilde{P}_1(t)$ with a lag of $M = 40$ shows two frequencies that appear as pairs of eigenvalues (0.12 and 0.25 cycles per time unit) emerging high above a red-noise-like background; two additional eigenvalues stand slightly above the red-noise background at frequencies 0.006 and 0.06 cycles per time unit (Fig. 5). The pair at frequency 0.25 (period 4) represents the small-scale oscillations in $\tilde{P}_1$; the smallest possible scale at this step of the IFS is 2, but it is not systematically reached throughout the interval of interest, so that the smallest scale which is always present is 4 (see Fig. 1b). The pair at frequency 0.12 cycles per unit (0.12 cpu; period 8.33) accounts for the grouping of two smaller-scale oscillations. Finally, the pair near 0.09 cpu (period 11.1; not shown) appears to be closely associated with the 0.12 cpu pair; hence the latter two pairs should be viewed as a quadruplet. The first eigenvalue has an associated period (156) larger than the lag window $M = 40$, so that it has no statistical significance; visually, it corresponds to roughly half the length of the largest-scale alternation between quiescence and oscillations (1/6 of the entire interval).

Figure 5: Monte-Carlo SSA of $\tilde{P}_1$.

SSA enhances one characteristic scale (which depends on M) and smears out the faster ones. On the other hand, the RCs of the low-order EOFs exhibit mixtures of fast and slow oscillations. This is so because all the oscillations have nearly the same variance by construction (Fig. 6): the amplitude of small-scale oscillations is exactly the same as the one of large-scale oscillations, and so the variance only depends on the distribution of the oscillations over the interval. The fast oscillations can thus be viewed as spurious because they do not provide information on the way the data set was constructed. Other global methods (Maximum Entropy: [13, 22, 23, 50]; Multi-Taper: [13, 22, 23, 51]) do not provide more relevant information on $\tilde{P}_1(t)$ either (not shown).

Figure 6: Global RCs of $\tilde{P}_1$.

4.1.2 Wavelet transforms

We next performed wavelet transforms of $\tilde{P}_1$ with the Gaussian analyzing function $\psi(x) = \exp(-x^2/2)$ and its derivative ψ'. Geometric scale increments $a = 3^n$, $n = 0, \ldots, 6$, were used. The choice of an increment of 3 is suggested by the triadic structure of the Cantor set; other choices (e.g., $a = 2^k$) clearly cannot provide the same optimal scale decomposition (see below). More generally, when the data set contains a discrete scale invariance with respect to dilations, the use of the preferred scaling ratio — such as the ratio 3 in the Cantor set studied here — is optimal [18, 28]. The wavelet analyses of $\tilde{P}_1$ are shown in Fig. 7. The triadic and self-similar structure of the Cantor set is clarified, as each scale picks up one step of the construction process.

Figure 7: Wavelet transforms of $\tilde{P}_1$ with scale increments of 3.

A wavelet analysis with ψ' as the mother wavelet (right column of Fig. 7), rather than ψ (left column), reveals further aspects of the process's singularities. The triadic construction of $\tilde{P}_1$ is also apparent in this decomposition, albeit not as clearly as with ψ. The wavelet transforms of $\tilde{P}_1$ in Fig. 7 do not differ significantly from those of its "exact" counterpart P_1 (not shown).

4.1.3 MS-SSA analysis

- *RC analysis.* We performed MS-SSA analyses of $\tilde{P}_1$ with scales W varying with geometric increments, $W = 2 \cdot 3^n$, and a ratio $W/M = 3$. We kept the first two EOFs for each scale. The first RCs of $\tilde{P}_1$ reveal very well the triadic structure of the Cantor set (left column of Fig. 8), as in the Gaussian-wavelet analysis.

Figure 8: RCs 1 and 2 of $\tilde{P}_1$ for scale increments of 3.

Unlike the wavelet analysis, MS-SSA restricted to EOF-1 reproduces almost exactly the steps taken to generate $\tilde{P}_1$, *i.e.*, the divisions do not overlap, even though parts of the reconstruction are lost due to side effects. Thus, in the case of $\tilde{P}_1$, the triadic structure seems to be more faithfully captured by this method. A slight drawback, on the other hand, is that SSA exhibits a mild "Gibbs effect," *i.e.*, an overshoot at certain discontinuities; it does not guarantee, therewith, the positiveness of the signal reconstruction which is ensured by a wavelet analysis with a positive function, like the Gaussian (but not its derivative: see right column of Fig. 7).

The RCs corresponding to EOF-2 (right column of Fig. 8) provide an analysis of the Cantor set's "derivative," as with the ψ'-wavelet transform. Hence, EOF-2 is analogous to ψ' (see Table 1 and ref.[25]), and the resulting RCs yield an analysis of the derivative of $\tilde{P}_1$ (not shown).

- *EOF analysis.* As anticipated in Section 2, we expect the EOFs $\rho_k^{(W)}$ to be similar for fixed order k and on different scales W, due to the choice of the scaling ratio 3 being equal to that used in constructing the time series. Figure 9 shows EOFs 1 through 4 of $\tilde{P}_1$, for scales W in a fixed ratio of 3 between 9 and 243 (a scale of 3 implies $M = 1$, for which the analysis is trivial), and at a location $b = 333$, situated at one third of the time series. The EOFs are scaled by $3^{n/2}$, so that their amplitudes match. We notice a near-convergence in the shape of EOFs 1–3 as the scale increases; this indicates that the time series is self-similar with scale increments of 3. EOF 4, on the other hand, needs larger scales (and more data) in order to converge.

Figure 9: EOFs $1-4$ of $\tilde{P}_1$ at $b=333$, with a scaling of $3^{n/2}$.

In contrast, we show in Fig. 10 the same EOFs 1-4 of $\tilde{P}_1$, for scales W in a fixed ratio of 2, between 16 and 256, at the same position $b=333$ along the time series. There is practically no convergence with W for fixed k, reflecting the mismatch between the scale factor 2 of the multiscale analysis and the preferred scale factor 3 of the time series. This suggests a novel method for detecting the existence of a preferred scaling factor λ in a time series by optimizing the scale ratio l in MS-SSA: only when $l \approx \lambda^m$, where m is an integer, will the convergence of the EOFs $\rho_k^{(W)} \to \rho_k^{(\infty)}$ be good as $W \to \infty$.

Figure 10: EOFs $1-4$ of $\tilde{P}_1$ at $b=333$, with a scaling of $2^{n/2}$.

This phenomenon reflects a "scale-locking" between the analyzing (multi)scale factor and the intrinsic scale ratio(s) present in the time series. This provides a scaling test that conventional wavelet analysis cannot offer, due to the *a priori* choice of ψ. In MS-SSA, the functional-shape constraint is relaxed and new information on the signal can thus be obtained.

• *Resonance.* An interesting phenomenon appears when applying MS-SSA to the "perfect" Cantor set $P_1(t)$. The local lag-covariance matrices seem to exhibit resonances between scales so that a given "slow" scale contains information on faster scales and hence exhibits oscillations that blur the triadic decomposition (not shown). This was not seen in the MS-SSA of the "approximate" time series $\tilde{P}_1(t)$, which contains noise in both time and scale that smoothes out such resonances and leads to a clear reconstruction of the triadic structure. The presence of a small amount of noise often increases the robustness, and hence ease of identification, of a phenomenon's main features. Well-known instances are the randomization method in general probability theory [52] and the addition of process or observation noise in sequential estimation theory [53].

• *Change of geometrical structure along the scale axis.* MS-SSA analysis of the multiple-rule Cantor set $\hat{P}_1$ (Fig. 11) shows that the method can capture the difference in structure beyond a threshold scale, even though the transition may be smoothed out by the two construction procedures. The RCs do not give as clearcut a triadic reconstruction of this Cantor set as for the single-rule set $\tilde{P}_1$ in Fig. 8. But, in spite of the relatively small data set we used (only 729 points), we can still detect in Fig. 11 the difference between the use of rule (14) for the larger scales and rule (16) for the smaller ones at the passage between the scales $W=9$ and $W=27$.

Figure 11: RCs 1 and 2 of $\hat{P}_1$ for scale increments of 3.

• *Reconstruction process by summing over scales for a single EOF.* By construction, there is a finite number M of EOFs for each scale W in MS-SSA. The EOFs form an orthonormal basis and hence the local decomposition of the time series occurs on a finite number M of modes, at each scale. Thus, for a given scale W, the sum of the corresponding RCs reproduces the original time series in each window. In wavelet language, this corresponds to reconstructing a signal by summing, at a fixed scale a (see Table 1), over a finite number of wavelet transforms; each of these transforms uses a different mother wavelet − for instance, the successive derivatives $\psi^{(n)}, 0 \leq n \leq M-1$, of the same mother wavelet $\psi = \psi^{(0)}$ − such that the finite set of mother wavelets form a (not necessarily orthogonal, but still nondegenerate) basis. This, however, does not correspond to the usual reconstruction formula (3).

In *continuous* wavelet analysis, the reconstruction is usually performed by summing over the scales for a fixed mother wavelet, according to Eq. (3), often called a "resolution of the identity (operator)" on admissible functions [16]. For discrete wavelet transforms, there does not exist, in general, a "resolution of identity" formula analogous to (3); the expansion of a function over a discrete wavelet basis is not orthogonal in general and special iterative methods must be developed to derive an inversion formula (see Chapter 3 in [16]).

Likewise, for our MS-SSA method, the expansion over discrete scales of a function with a fixed EOF order is not orthogonal in general. The MS-SSA approach is thus more closely related to the theory of "frames" [16]. It is easy to show from the definition of a frame that the family of EOFs ρ_k^W of given order k for discrete scales $W_n = a^n$ with n integer constitutes such a frame. This yields an algorithm to reconstruct the signal from the discrete set of wavelet transforms, in our case the RCs of same order k at different scales, for special choices of the scale W and translation parameter b [16] :

$$X_i = \sum_m \sum_{j=1}^{M} \hat{\rho}_{k,j}^{W_m} a_{k,i-j}^{W_m} \; ; \tag{9}$$

here the sum over m corresponds to the different scales, and $a_{k,i-j}^{W_m}$ is given by

$$a_{k,i-j}^{W_m} = \sum_{l=1}^{M} \rho_{k,l}^{W_m} \; X_{i-j+l-1} \; , \tag{10}$$

while $\hat{\rho}_{k,j}^{W_m}$ is a vector related to the EOF $\rho_{k,j}^{W_m}$ through an iterative algorithm [16].

The inversion can thus be formulated in principle, but in practice it is quite laborious. Emphasizing the practical aspect, we propose an approximate inversion formula similar to (9) but with $\hat{\rho}_{k,j}^{W_m}$ replaced by $\rho_{k,j}^{W_m}$:

$$X_i = \sum_m \sum_{j=1}^{M} \frac{1}{M} \rho_{k,j}^{W_m} a_{k,i-j}^{W_m} \; . \tag{11}$$

This formula is tested in Fig. 12. The top panel shows the time series P_1. The second panel shows the RC using EOF-2 ($k = 2$) for the single scale $W = 9$. This corresponds to formula (11) with a single value of $k = 2$, while $a = 3$ and thus $W_2 = 9$. The third, fourth and fifth panels show the sum (11) with two, three and four terms; these terms correspond to $W_2 = 9$, $W_3 = 27$, $W_4 = 81$ and $W_5 = 243$, while the order of the EOF $\rho_k^{W_m}$ is always $k = 2$.

Figure 12: Reconstruction of P_1 using Eq. (11) for scale increments of 3.

4.2 A log-periodic process

In this subsection, we focus on the detection and quantification of self-similar structures in a log-periodic process. In the case of Eq. (17), we expect discrete self-similarity near $t = t_c$, with a preferred scaling of $\lambda = 2$. Indeed, in Fig. 2 the distance between successive maxima of the oscillations decreases geometrically with a scaling of about 2, and thus converges to zero as t approaches t_c. In the case of P_2, a straightforward application of MS-SSA (with EOFs 1 and 2) and of the wavelet transform (with ψ and ψ') give extremely similar results; they both act as low-pass filters with geometrically increasing width (not shown) and hence do not provide a satisfactory reconstruction of the signal.

We analyzed therefore $P_2(t)$ by MS-SSA with scaling increments of 2 but placing ourselves near $t_c = 100$ in the sense that the local windows always end at t_c, and their midpoint moves to the left. The first four EOFs, shown in Fig. 13, clearly converge rapidly as the size of the window — and thus the scale analyzed — increases. If we choose another ratio, say 3, to scale the windows no such convergence is observed (not shown), as was the case in comparing Fig. 9 with Fig. 10 for the Cantor set. This sensitivity of MS-SSA's EOFs to the correct scaling factor suggests a new way to characterize log-periodicity.

Figure 13: EOFs 1–4 of P_2 near t_c, with 2^n scaling.

EOFs 1 and 3 capture the local mean of the signal and scale the period of its oscillation, respectively, having no zero and two zeroes (see Fig. 13). RC-2 describes best the local fast oscillations near the critical point $t = t_c$, while RC-1 at different scales captures well the overall behavior of the time series in Fig. 2 (see Fig. 14).

Figure 14: RCs 1 and 2 of P_2 for scale increments of 2.

We look next at what happens near $t = 0$, *i.e.*, we use windows starting at the beginning of the time series. The convergence of EOFs 2–4 is still excellent as W increases (not shown). Only EOF-1, which is sensitive to the overall behavior of the time series, exhibits slower convergence with W, due to the lack of exact self-similarity of $P_2(t)$ near the origin. This robust, albeit imperfect, convergence of its EOFs for log-periodic time series makes the MS-SSA approach potentially useful for the analysis of real log-periodic data, in the presence of noise and truncated singularities. The application of this technique to real data on earthquakes [54, 55] and financial [36] time series will be investigated separately.

4.3 Multiplicative noise

As expected, classical global methods which assume the existence of finite variance do not give a particularly good insight into the process $P_3(t)$. Monte-Carlo SSA yields a periodicity that is statistically significant above 95%; the actual period, though, lies between 3 and 4 cycles and fluctuates from realization to realization (not shown). The time series cannot be otherwise distinguished from red noise; this is obviously wrong and can be explained by the lack of stability in the global variance estimate.

The global RCs (Fig. 15) provide a sensible scale-dependent reconstruction for the slower scales: the characteristic scale halves from RC-1 to RCs 2–3, on to RC-4 and, again to RCs 5–8, but the reconstructions become noisier as the order increases further (not shown). The total reconstruction of $P_3(t)$ from the eight leading RCs (Fig. 15e) is an excellent coarse-scale representation of $P_3(t)$, as depicted in Fig. 3.

Figure 15: Global RCs of $P_3(t)$.

Next, wavelet and MS-SSA analyses were both performed with geometric scale increments of 2, while the latter used a ratio $W/M = 3$ as in Secs. 4.1 and 4.2. A wavelet analysis using the Gaussian mother wavelet (Fig. 16, left column) progressively smoothes the details of the time series as the scale parameter a increases. The analysis reveals the natural self-similarity of P_3, as the analysis at scale $a = 32$ can be normalized to match the first half of the analysis at scale $a = 8$ (Fig. 16). In turn, the analysis at scale $a = 8$ can be normalized to fit the analysis at scale $a = 2$ in the same way. Thus, we can infer that scale increments of about 4 are sufficient for the analysis of P_3.

Figure 16: Wavelet analysis of P_3.

MS-SSA using EOF-1 (Fig. 17, left column) behaves in a similar way, with a progressive smoothing of the time series and the same scaling result. RC-2 captures the oscillating variability (right column of Fig. 17) around EOF-1, as does ψ' with respect to ψ (right column of Fig. 16). In this case, the parallel between the behavior of wavelet analysis and MS-SSA is striking. As we saw in Fig. 15e, the reconstruction of P_3 can also be completed with a finite sum of EOFs, for a fixed scale of analysis, whether global of not.

In contrast to the results found in Fig. 9 for the Cantor set $\tilde{P}_1$, while the EOFs of the same order have a similar shape across scales, they cannot be exactly superimposed (not shown). This simply reflects the fact that the time series P_3 is not exactly self-similar with a preferred scaling ratio of 2. In fact, no choice of a single scaling ratio provides a good convergence of the EOFs for fixed k: P_3 is not self-similar but rather exhibits a more complex scaling structure, that is approximately multifractal [43]. The shape of the EOFs (not shown) is in fact quite similar to that of the Gaussian

Figure 17: RCs 1 and 2 of P_3 with scale increments of 2.

wavelets. The data-adaptive virtues of the EOFs are illustrated in Fig. 18, where the shape of the local EOF-1 follows closely the variations of P_3.

Figure 18: The evolution in time of the local EOF-1 for P_3.

On this data set, the two methods under consideration perform in a rather similar way. The advantage of MS-SSA is that it fully uses the information contained in the time series in order to compute analyzing functions, and that the reconstruction is intrinsically complete, given a small set of EOFs, which is not the case for wavelets.

4.4 Southern Oscillation Index

We performed MS-SSA on the monthly SOI data for the years 1933-1996 (see Section 3.2). The parameters were, as before, $W/M = 3$ and geometric scale increments of 2. In addition to the previous analyses, we also computed an "instantaneous" frequency for each local EOF. This was simply done by least-square fitting a sine wave to each local EOF of interest. The instantaneous frequency can also be obtained from a complex wavelet transform [56], by using information on the phase of the transform.

Our analysis does not reveal any evidence of self-similarity or fractality, as obtained for the previous synthetic data sets. Instead, we find a preferred scale of variability between 3 and 5 years (not shown), which corresponds to ENSO's low-frequency mode [45, 46]? The first two local EOFs are consistently paired and in phase quadrature, which shows that the nonlinear oscillation associated with this mode is robust and persists throughout the sixty odd years being examined.

The computation of the instantaneous frequency allows us to detect an abrupt frequency shift of the ENSO mode near 1960 (Fig. 19). The characteristic periodicity goes from 57 months (between 1943 and 1961) to 39 months (between 1963 and 1980). A comparable decrease in period was observed in the early 1960s by Moron et al. [14] in Tropical Pacific sea-surface temperatures, by using multi-channel global SSA, and by Wang and Wang [48] in a sea-level pressure record at Darwin, using a wavelet transform.

Figure 19: Instantaneous SOI frequencies, based on EOF-1 of MS-SSA at scale $W = 128$.

Our MS-SSA approach improves on these previous results in two respects. First, compared to both the global SSA analysis [14] and the wavelet and waveform (matching pursuit) approach [48], the combination of our local SSA in a sliding window with the local sine-wave fit gives a much sharper signature of the increase in instantaneous frequency. Second, we can check this regime change at different scales by performing the same analysis separately at each scale. In contrast, the wavelet and waveform approaches need all the scales together in order to be able to detect the frequency change. We thus have at our disposal an additional degree of freedom to test for the reliability and robustness of the signal.

We find that, for $W = 64$ and 32, the transition is not as sharp as with $W = 128$ (the value used in Fig. 19), but is still identifiable. This degradation is expected since the maximum period $M = W/3$ that can in principle be retrieved at these values of W is less that the ENSO period of interest. This means in practice that only scales of the order of $W = 128$ or larger will capture the ENSO period in question and the change in it. At scales that are even larger, however, the analysis becomes limited in its reliability by the finite length of the time series. Hence the value $W = 128$ provides a good compromise between the two conflicting constraints of resolution and statistical significance [10, 11, 22, 23].

5 Concluding remarks

We have presented a multi-scale extension of the orthogonal component decomposition of a time series known as singular spectrum analysis (SSA). This multi-scale extension, which we called MS-SSA, has been performed by varying the width W of the analyzing window, by analogy with wavelet analysis. The successive analyses used a fixed ratio – equal to 2 or 3 in the numerical examples – between the width W and the order M of the local covariance matrix.

Our multi-scale SSA (MS-SSA), or data-adaptive wavelet analysis, extends various other approaches in the same spirit. Mallat et al. [57] describe a hierarchical binary tree of basis functions for locally stationary time series. Lilly and Park [58] select their Slepian wavelets [59] at a given scale by optimizing the spectral energy in a given frequency band [51] and vary the scale. Mann and Park [60] have applied this approach to a multivariate analysis of interannual temperature variations. Wavelet packets [61, 62] can be defined as smooth versions of the Walsh system used in binary signal processing. While they extend wavelet analysis and take a step in the direction of satisfying our requirements (A) and (B) above, the wavelet-packet approach does not allow the same degree of flexibility and data-adaptivity as provided by our method.

Coifman and Saito [63] use the Karhunen-Loève expansion that underlies SSA to construct a dictionary of orthonormal bases that has a tree structure. Their starting point is thus very similar to ours; Coifman and Saito, however, use additional criteria to choose the orthonormal bases from the overdetermined dictionary in order to perform a multi-scale analysis. Our Karhunen-Loève decomposition is performed in moving windows of size W that maintain a fixed ratio to the order M of the decomposition. As a consequence, our orthogonal basis is unique, minimal and offers a very intuitive interpretation of the decomposition.

The empirical orthogonal functions (EOFs) play the same role in MS-SSA as the "mother wavelets." In contrast to standard wavelet transforms, our basis functions are data-adaptive, *i.e.* they are determined from a diagonalization of the covariance matrix C_M of the data within the observing window. The data-adaptive character of the MS-SSA basis allows us to test further for self-similar properties of a time series by comparing the EOFs at different scales.

In MS-SSA's orthogonal decomposition, the sum of a finite set of coefficients weighted by the corresponding "wavelets," *i.e.*, the reconstructed components (RCs) *at the same scale*, reconstructs exactly the initial signal. This is in contrast to the standard wavelet decomposition in which one needs to sum over all scales or sum over an infinite set of wavelets (for instance corresponding to the successive derivatives of a Gaussian) at a given scale. We also propose an approximate "resolution of identity" inversion formula to reconstruct the signal from a nonorthogonal sum over scales of RCs that correspond to a fixed EOF order. This formula is the counterpart of the usual wavelet inversion formula and arises from the theory of "frames" [16].

We tested MS-SSA on a few irregular time series for which conventional single-scale spectral analysis techniques would fail to reveal their properties: Cantor sets and intermittently amplifying multiplicative noise. In many respects, wavelets and MS-SSA behave in similar ways: MS-SSA decomposition onto the first and second EOFs is very similar to a wavelet analysis that uses the Gaussian mother wavelet and its first derivative. But in the case of the approximate Cantor set $\tilde{P}_1$, the MS-SSA decomposition gives a clearer insight on the way the set was generated. In addition, the shape of the EOFs also allowed us to check precisely for the self-similarity of that set.

The latter property is probably the most interesting one to investigate in the future, as the change in shape of the EOFs as a function of scales provide a new method to test for discrete scale invariance. Indeed, the detection of the preferred scaling ratio in a discretely scale-invariant system may be obtained by finding the ratio of the multi-scale analysis that minimizes the difference between the data-adaptive EOF "wavelets" obtained at different scales. This test appears to be more robust than previously used methods [30] and has been demonstrated successfully on a time series that presents the scaling structure of the triadic Cantor set.

Finally, we applied the MS-SSA methodology to a real climatic time series, given by the monthly Southern Oscillation index (SOI) for 1933-1996. The SOI is a broadly used indicator for the El Niño/Southern Oscillation (ENSO). The sharp transition in ENSO's low-frequency mode from a period of roughly 5 years to roughly 3 years in the early 1960s refines earlier results obtained using global SSA [14] and standard wavelet [48] analysis. It supports the Devil's staircase theory of ENSO [47] in the following sense: a sudden shift in frequency can be reconciled most easily with the idea that an interdecadal change in mean thermocline depth in the Tropical Pacific [27, 67] causes the coupled ocean-atmosphere system there to “jump” from one broad step of the staircase to another. We hope that similar physical inferences will be facilitated in other areas of the geosciences to which MS-SSA might be applied.

6 Acknowledgments

It is a pleasure to thank A. Fournier, M. Mann, J.-F. Muzy and N. Saito for stimulating conversations and references. This work was supported by NSF grant ATM95-23787 (MG and PY), a NATO travel grant (PY) and research funds of UCLA’s Institute of Geophysics and Planetary Physics (IGPP) during PY’s leave from LSCE–CEA, and NSF grants EAR96-15357 and EAR97-06488 (DS). This is contribution No. XXXX of LSCE and contribution No. YYYY of IGPP at UCLA.

7 Appendix: Construction of the synthetic time series

7.1 Approximate triadic Cantor set

The time series $\tilde{P}_1$ was obtained by using an Iterated Function System (IFS) [29] to approximate a simple Cantor set. An IFS is a contracting stochastic map constructed to converge towards a self-similar attractor of zero measure. The example we study consists in a time series whose characteristics mimic the triadic Cantor set. To generate this time series, we use the following one-dimensional IFS:

$$x_{n+1} = \begin{cases} \frac{1}{3}x_n & \text{with probability } \frac{1}{2}, \\ \frac{1}{3}x_n + \frac{2}{3} & \text{with probability } \frac{1}{2}. \end{cases} \tag{12}$$

Starting from an arbitrary point $x_0 \in [0,1]$, it is straightforward to verify that this stochastic IFS map converges to a triadic Cantor set on $[0,1]$, characterized by the fractal dimension of $\log 2/\log 3$ [29]. We iterated Eq. (12) 10000 times and discarded the first 1000 values of the set so generated in order to obtain a set $\{x_i\}$ that has converged close to the Cantor set target.

We sorted these values, $x_i \in [0,1]$, into increasing order, $0 \le x_0 \le ... \le x_i \le x_{i+1} \le ... \le x_N \le 1$. A time series $\tilde{P}_1$ with, say $N = 1000$ equidistant points, is generated by interpolating the set $\{x_i\}$ every $\Delta t = 1/N$ and scaling it to N:

$$\tilde{P}_{1j} = \begin{cases} 1 & \text{if } \exists x_i \in [j/N, (j+1)/N], \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

with $0 \le j \le N$. As a result of the way this time series is generated with the stochastic IFS, it contains residual noise with respect to a perfect Cantor set both in time and scale, because of the interpolation process (Fig. 1). This means that the smallest scales are not always of unit length, due to incomplete interval division.

7.2 The exact triadic Cantor set

The second time series P_1 was obtained recursively by iterating the binary map:

$$1 \mapsto 101, \tag{14}$$

$$0 \mapsto 000\ , \tag{15}$$

on the initial list $s_0 = (1)$. At the L-th iteration, the list s_L contains 3^L bits which have an exact Cantor set structure. We generated the time series P_1 with $L = 6$ iterations, so that $P_{1j} = s_L(j)$, $1 \le j \le 3^L$. A "noisy" Cantor time series could be obtained by assigning a probability of occurence between rule (14) and $1 \mapsto 111$. This procedure emulates the incomplete division scheme towards smaller scales that we observe in $\tilde{P}_1$.

A large variety of Cantor sets can be obtained by modifying the rules (14, 15) in a straightforward manner. For instance, we generated a multi-rule Cantor set by iterating rules (14, 15) three times, then used the rules

$$\begin{aligned} 1 &\mapsto 110, \\ 0 &\mapsto 000. \end{aligned} \tag{16}$$

three more times, in order to obtain a set $\hat{P}_1$ of 729 ($= 3^6$) points.

The fractal dimension of a set generated after iterating infinitely many arbitrary permutations of rules (16) and (14) is still $\log 2/\log 3$, but the geometry of the resulting set obviously depends on the sequence in which the rules were applied. This is a clear instance where the fractal dimension is not enough to characterize the geometric structure [2, 64, 65]. Other rules and combinations of rules can be used to obtain other fractal dimensions and, for a given dimension, distinct geometric structures of the resulting set.

7.3 Log-periodic time series

The functional form of the log-periodic process P_2 is simply:

$$P_2(t) = A + B(t - t_c)^{\alpha} \; \{1 + C \cos\left[2\pi \log(t - t_c)/\log(\lambda)\right]\} \,. \tag{17}$$

Here, we take $A = 2$, $B = -1$, $C = 0.2$, $\alpha = 1/2$ and $\lambda = 2$ (see Fig. 2). The critical time is $t_c = 100$; we regularly sample 1000 time steps between 1 and t_c with a time interval $\Delta t = 0.1$, and define the P_2 time series by $P_{2i} = P_2(i\Delta t)$, $1 \leq i \leq 1000$. Note that in our case $\alpha = 1/2$ and thus $\lim_{t\to t_c} P_2(t) = A$; hence P_2 is indeed continuous on the $[0, t_c]$ interval, although it is not differentiable at $t = t_c$.

7.4 The statistically self-similar process P_3

The process P_3 is defined by

$$P_{3,i+1} = a_i P_{3i} + b_i, \tag{18}$$

where a_i is a uniformly distributed random variable that can take values larger than 1 — here, $a_i \in [0.48, 1.48]$ — and b_i is uniformly distributed over $[0, 1]$. In order to ensure that P_3 does not grow to infinity but is "globally" stationary, we enforce the technical condition that the average growth rate $\langle \ln a_i \rangle$ be negative. For the parameters taken above, it follows that $\langle \ln a_i \rangle = -0.06747$.

The apparently innocuous self-affine map (18) leads to time series with surprisingly rich properties [42, 43]. The probability density function of P has an asymptotic tail for large P of the form of a power law $p(P) \sim P^{-(1+\mu)}$, where the exponent μ is the smallest real solution of $\langle [a_i]^{\mu} \rangle = 1$. Complex solutions μ of this probabilistic equation also exist that have equal or larger real parts; they correspond to higher-order corrections to the leading power-law behavior of $p(P)$ with log-periodic modulations. In practice, however, these modulations are not always observable due to their averaging by the large fluctuations in the a_i and b_i values [66].

In the present numerical example, the smallest real exponent is $\mu \approx 1.5$. Since $\mu > 1$, the mean $\langle P_{3i} \rangle$ is well-defined and finite. However, since $\mu < 2$, the variance $\langle [P_{3i} - \langle P_{3i} \rangle]^2 \rangle$ is infinite. In practice, this means that, for a given realization and a fixed time length T, the estimation of the variance by integration over $[t_0, t_0 + T]$ is very unstable as a function of the position t_0 of the time window and its length T. Furthermore, the variance diverges approximately as $T^{2-\mu}$ with the length T of the time series.

References

[1] M. Ghil and S. Childress. *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory and Climate Dynamics*. Springer-Verlag, New York, 1987.

[2] B. B. Mandelbrot. *The Fractal Geometry of Nature*. Freeman, San Francisco, 2nd edition, 1982.

[3] Y. Y. Kagan. Observational evidence for earthquakes as a nonlinear dynamic process. *Physica D*, 77:160–192, 1994.

[4] W. R. Peltier. Phase transition modulated mixing in the mantle of the earth. *Phil. Trans. Roy. Soc. London Series A*, 354:1425–1443, 1996.

[5] C. J. Rhodes and R. M. Anderson. Power laws governing epidemics in isolated populations. *Nature*, 381:600–602, 1996.

[6] S. J. Gould. *Full House*. Harmony Books, New York, 1996.

[7] E. J. Hannan. *Time Series Analysis*. Methuen, New York, 1960.

[8] D. B. Percival and A. T. Walden. *Spectral Analysis for Physical Applications*. Cambridge University Press, Cambridge, UK, 1993.

[9] D. S. Broomhead and G. P. King. Extracting qualitative dynamics from experimental data. *Physica D*, 20:217–236, 1986.

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

[11] R. Vautard, P. Yiou, and M. Ghil. Singular spectrum analysis: A toolkit for short noisy chaotic signals. *Physica D*, 58:95–126, 1992.

[12] M. Ghil and R. Vautard. Interdecadal oscillations and the warming trend in global temperature time series. *Nature* 350:324–327, 1991.

[13] P. Yiou, M. F. Loutre, and E. Baert. Spectral analysis of climate data. *Surveys Geophys.*, 17:619–663, 1996.

[14] V. Moron, R. Vautard, and M. Ghil. Trends, interdecadal and interannual oscillations in global sea-surface temperatures. *Clim. Dyn.*, 14:545–569, 1998.

[15] Y. Meyer. *Ondelettes et Opérateurs I: Ondelettes*. Hermann, Paris, 1989.

[16] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, PA, 1992.

[17] Y. Meyer. *Wavelets: Algorithms and Applications*. SIAM, Philadelphia, PA, 1993.

[18] A. Arneodo, F. Argoul, J. F. Muzy, and M. Tabard. Beyond classical multifractal analysis using wavelets: Uncovering a multiplicative process hidden in the geometrical complexity of diffusion limited aggregates. *Fractals*, 1:629–649, 1993.

[19] M. Holschneider. *Wavelets: An Analysis Tool*. Clarendon Press, New York, 1995.

[20] R. Courant and D. Hilbert. *Methods of Mathematical Physics*. New York, Interscience Publishers, Vol. I, 1953.

[21] M. D. Dettinger, M. Ghil, C. M. Strong, W. Weibel, and P. Yiou. Software expedites singular-spectrum analysis of noisy time series. *Eos Trans. AGU*, 76:12, 20, 21, 1995.

[22] M. Ghil and P. Yiou. Spectral methods: What they can and cannot do for climatic time series. In *Decadal Climate Variability: Dynamics and Predictability*, D. Anderson and J. Willebrand, editors, Elsevier, Amsterdam, pp. 446-482, 1996.

[23] M. Ghil and C. Taricco. Advanced spectral analysis methods. In *Past and Present Variability of the Solar-Terrestrial System: Measurement, Data Analysis and Theoretical Models*, G. Cini Castagnoli and A. Provenzale (eds.), Societa Italiana di Fisica, Bologna, & IOS Press, Amsterdam, pp. 137–159, 1997.

[24] P. Yiou. Dynamique du Paléoclimat: Des Données et des Modèles. Ph.D. thesis, Université Pierre et Marie Curie, Paris 6, 1994.

[25] J.F. Gibson, J.D. Farmer, M. Casdagli and S. Eubank, An analytic approach to practical space reconstruction. *Physica D*, 57:1–30, 1992.

[26] C.L. Keppenne and M. Ghil. Adaptive filtering and prediction of the Southern Oscillation index. *J. Geophys. Res.*, 97:20449–20454, 1992.

[27] M. Ghil and N. Jiang. Recent forecast skill for the El Nio/Southern Oscillation. *Geophys. Res. Lett.*, 25:171–174, 1998.

[28] B. Dubrulle, F. Graner, and D. Sornette, editors. *Scale Invariance and Beyond*. EDP Sciences and Springer, Berlin, 1997.

[29] M. F. Barnsley. *Fractals Everywhere*. Academic Press, Boston, 2nd edition, 1993.

[30] D. Sornette. Discrete scale invariance and complex dimensions. *Physics Reports*, 297:239–270, 1998.

[31] G. Jona-Lasinio. The renormalization group: a probabilistic view. *Nuovo Cimento*, 26B:99, 1975.

[32] M. Nauenberg. Scaling representations for critical phenomena. *J.Phys. A*, 8:925, 1975.

[33] Th. Niemeijer and J.M.J. van Leeuwen. *In Phase Transitions and Critical Phenomena*, Vol.6, C. Domb and M.S. Green, eds. (Academic Press, London, 1976), p.425.

[34] B. Derrida, L. De Seze and C. Itzykson. Fractal structure of zeros in hierarchical models. *J. Stat. Phys.*, 33:559-569, 1983; B. Derrida, C. Itzykson and J.M. Luck. Oscillatory critical amplitudes in hierarchical models. *Commun.Math.Phys.*, 94:115-132, 1984.

[35] D. Bessis, J.-D. Fournier, G. Servizi, G. Tourchetti and S. Vaienti. Mellin transforms of correlation integrals and generalized dimension of strange sets. *Phys. Rev. A*, 36:920–928, 1987.

[36] D. Sornette and A. Johansen. Large financial crashes. *Physica A*, 245:411–422, 1997.

[37] D. Dee, and M. Ghil. Boolean difference equations, I: Formulation and dynamic behavior. *SIAM J. Appl. Math.*, 44:111–126, l984; M. Ghil and A. Mullhaupt. Boolean delay equations II: Periodic and aperiodic solutions. *J. Stat. Phys.*, 41:125–173, 1985.

[38] M. Ghil, A. Mullhaupt and P. Pestiaux, Deep water formation and Quaternary glaciations, *Clim. Dyn.*, 2:1–10, 1987.

[39] D.G. Wright, T. F. Stocker, and L. A. Mysak. A note on Quaternary climate modelling using Boolean delay equations. *Clim. Dyn.*, 4:263–267, 1990.

[40] M.S. Darby and L.A. Mysak, A Boolean delay equation model of an interdecadal arctic climate cycle, *Clim. Dyn.*, 8:241–246, 1993.

[41] A. Saunders, A., A Boolean Delay Equation Model of ENSO Variability, M. S. Thesis, UCLA (1998).

[42] D. Sornette and R. Cont. Convergent multiplicative processes repelled from zero: power laws and truncated power laws. *J. Phys. I France*, 7: 431–444, 1997.

[43] D. Sornette. Linear stochastic dynamics with nonlinear fractal properties. *Physica A*, 250:295–314, 1998.

[44] M. Ghil, M. Kimoto, and J. D. Neelin. Nonlinear dynamics and predictability in the atmospheric sciences. *Rev. Geophys.*, 36 (Supplement U. S. National Report IUGG, 1987–91):46–55, 1991.

[45] E.M. Rasmusson, X. Wang, and C.F. Ropelewski. The biennial component of ENSO variability. *J. Mar. Syst.*, 1: 71–96, 1990.

[46] N. Jiang, D. Neelin, and M. Ghil. Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific. *Clim. Dyn.*, 12:101–112, 1995.

[47] P. Chang, B. Wang, T. Li, and L. Ji. Interactions between the seasonal cycle and the Southern Oscillation: Frequency entrainment and chaos in an intermediate coupled ocean-atmosphere model. *Geophys. Res. Lett.*, 21:2817–2820, 1994; F.-F. Jin, J. D. Neelin, and M. Ghil. El Niño on the Devil's staircase: Annual subharmonic steps to chaos. *Science*, 264:70–72, 1994; E. Tziperman, L. Stone, M. Cane and H. Jarosh. El Niño chaos: Overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator. *Science*, 264:72–74, 1994.

[48] B. Wang and Y. Wang. Temporal structure of the Southern Oscillation as revealed by waveform and wavelet analysis. *J. Clim.*, 9:1586–1598, 1996.

[49] M. R. Allen and L. A. Smith. Monte Carlo SSA: Detecting irregular oscillations in the presence of coloured noise. *J. Clim.*, 9:3373–3404, 1996.

[50] D. G. Childers, editor. *Modern Spectrum Analysis.* IEEE Press, New York, 1978.

[51] D. J. Thomson, Spectrum estimation and harmonic analysis. *IEEE Proc.*, 70:1055–1096, 1982.

[52] W. Feller. *An Introduction to Probability Theory and its Applications.* John Wiley and Sons, New York, 1971.

[53] Ghil, M. and P. Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. *Adv. Geophys.*, 33:141–266, 1991; Miller, R. N., M. Ghil and F. Gauthiez. Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.*, 51:1037–1056, 1994.

[54] D. Sornette and C.G.Sammis. Complex critical exponents from renormalization group theory of earthquakes: Implications for earthquake predictions. *J.Phys.I France*, 5:607–619, 1995.

[55] A. Johansen, D. Sornette, H. Wakita, U. Tsunogai, W.I. Newman and H. Saleur. Discrete scaling in earthquake precursory phenomena: evidence in the Kobe earthquake, Japan. *J.Phys.I France*, 6:1391–1402, 1996.

[56] M. Farge. Wavelet transforms and their applications to turbulence. *Ann. Rev. Fluid Mech.*, 24:395–457, 1992.

[57] S. Mallat, G. Papanicolaou, and Z. Zhang. Adaptive covariance estimation of locally stationary processes. *Annal. of Statistics*, 26:1–47, 1998.

[58] J. M. Lilly and J. Park. Multiwavelet spectral and polarization analyses of seismic records. *Geophys. J. Int.*, 122:1001–1021, 1995.

[59] S. Slepian. Prolate spheroidal wave functions, Fourier analysis and uncertainty-V: The discrete case. *Bell. Syst. Tech. J.*, 1371–1430, 1978.

[60] J. Park and M. E. Mann. Interannual temperature events and shifts in global temperature: A "multiwavelet" correlation approach. Preprint, 1997.

[61] M. V. Wickerhauser. Lectures on Wavelet Packet Algorithms. Technical Report, Dept. Mathematics, Washington U., St. Louis, MO, 1991.

[62] R. E. Learned and A. S. Willsky. A wavelet packet approach to transient signal classification. *Appl. Comput. Harmonic Anal.*, 2:265–278, 1995.

[63] R. R. Coifman and N. Saito. The local Karhunen-Loève bases. *Proc. IEEE Intl. Symp. Time-Frequency and Time-Scale Analysis*, pp. 129–132, 1996.

[64] R. Blumenfeld and B.B. Mandelbrot. Lévy dusts, Mittag-Leffler statistics, mass fractal lacunarity, and perceived dimension. *Phys. Rev. E*, 56:112–118, 1997.

[65] F.J. Solis and L. Tao. Lacunarity of random fractals. *Phys. Lett. A*, 228:351–356, 1997.

[66] P. Jögi, D. Sornette and M. Blank. Fine structure and complex exponents in power law distributions from random maps. *Phys. Rev. E*, 57:120–134, 1998.

[67] T.R. Knutson, S. Manabe, and D. Gu. Simulated ENSO in a global coupled ocean-atmosphere model: Multidecadal amplitude modulation and CO_2 sensitivity. *J. Clim.*, 10:138-161, 1997; D. Gu, and S. G. H. Philander. Secular changes of annual and interannual variability in the tropics during the past century. *J. Clim.*, 8:64-876, 1995.

8 Figure captions

Figure 1 : Iterated funtion system (IFS) simulation of a classical triadic Cantor set. (a) 1000-point interpolation to generate the $\tilde{P}_1$ time series; (b) blow-up of $\tilde{P}_1$ between 1 and 333 in order to visualize the (incomplete) triadic structure of the time series.

Figure 2 : Log-periodic time series P_2, with critical point $t_c = 100$, regularly sampled with 1000 points; see text and Eq. (17) for parameter values.

Figure 3 : Multiplicative noise process P_3: (a) individual realization over 512 equidistant points; (b) histogram of the variations of P_3, with 100 bins — the ordinate represents the range of values (same as in panel a) and the abscissa is a frequency count.

Figure 4 : Variations of the Southern Oscillation Index (SOI) between January 1933 and December 1996 (from the Climate Research Unit Data at the University of East Anglia, U.K.). Time on the abscissa in years and SOI on the ordinate centered to have mean zero and normalized by its standard deviation.

Figure 5 : "Global" Monte-Carlo SSA of $\tilde{P}_1$, with window width (*i.e.* number of lags) $M = 40$. The vertical bars are 95% confidence intervals for 100 red-noise realizations with the same variance and exponential-decorrelation time as $\tilde{P}_1(t)$. The diamonds represent the singular values of the signal; those that stand out above the error bars are statistically significant. The frequencies associated with the latter are obtained by least-square fitting the corresponding EOF to a sine function.

Figure 6 : Reconstructed components (RCs) of $\tilde{P}_1$ with global SSA. (a–c) Reconstructions using groups of eigenelements, as indicated in the legend of each panel; (d) the sum of RCs 1–6. The SSA window is $M = 40$, as in Fig. 5.

Figure 7 : Wavelet transforms of $\tilde{P}_1$, with scale increments of 3. The left panels use the Gaussian analyzing function $\psi(x)$ (shown as "Psi" in the legend) and the right ones use its derivative $\psi'(x)$ (shown as "dPsi"); the scales $a = 3, 9, 27$ and 81 (for $n =$ 1–4, see text) are also indicated from top to bottom on the left.

Figure 8 : Multi-scale SSA (MS-SSA) of $\tilde{P}_1$, with scale increments of 3. The left panels show the reconstruction of $\tilde{P}_1$ using the local RCs associated with EOF-1, the right panels that using the RCs obtained by projection onto the local EOF-2; the window widths $W = 9, 27, 81$ and 243 are shown from top to bottom on the left (like the scale a in Fig. 7).

Figure 9 : Normalized EOFs 1–4 (panels a–d) of $\tilde{P}_1$ at $b = 333$. The EOFs at scale 3^n were multiplied by a factor of $3^{(n-1)/2}$; the abscissa in each panel is linearly normalized to $[0, 1]$. The scales are $W = 9$, 27, 81 and 243, respectively in thick continuous, thin continuous, dotted and dashed lines (see also legend in panel a).

Figure 10 : Normalized EOFs 1–4 of $\tilde{P}_1$ at $b = 333$. The EOFs at scale 2^n are multiplied by a factor of $2^{(n-1)/2}$; abscissa normalized as in Fig. 9. The scales are $W = 16$, 32, 64, 128 and 256, respectively in thick continuous, thin continuous, dotted, dashed and dot-dashed lines (see also legend in panel b).

Figure 11 : MS-SSA of $\hat{P}_1$, with scale increments of 3. The left panels contain the RCs associated with EOF-1, the right panels those associated with EOF-2; the same window widths W as in Fig. 8 (shown on the left).

Figure 12 : Reconstruction of P_1 using formula (11) for scales that are powers of 3. The top panel shows the time series P_1 derived from the exact triadic Cantor set. The remaining panels show the reconstructions of P_1 using the local RCs based on EOF-2. The second panel uses RC-2 for the single scale $N = 9$; this corresponds to Eq. (11) with a single value of $m = 2$, $a = 3$ and thus $W_2 = 9$. The third panel shows the sum (11) with two terms, corresponding to $W_2 = 9$ and $W_3 = 27$; the fourth panel shows the sum (11) with three terms, corresponding

corresponding to $W_2 = 9$, $W_3 = 27$, $W_4 = 81$ and $W_5 = 243$ and still the same order $k = 2$ of the EOF.

Figure 13 : Normalized EOFs 1–4 (panels a-d) of P_2 at $t = t_c - W/2$, so that each window ends at t_c. The EOFs at scale 2^n are multiplied by a factor of $2^{(n-1)/2}$, and the abscissa is normalized to $[0, 1]$. The scales are $W = 16$, 32, 64, 128 and 256, respectively in thick continuous, thin continuous, dotted, dashed and dash-dotted lines (see panel a).

Figure 14 : MS-SSA of P_2 with RC-1 (left column) and RC-2 (right column). The scales W vary from 16 to 256, going from the top to the bottom (indicated along the left edge of the figure).

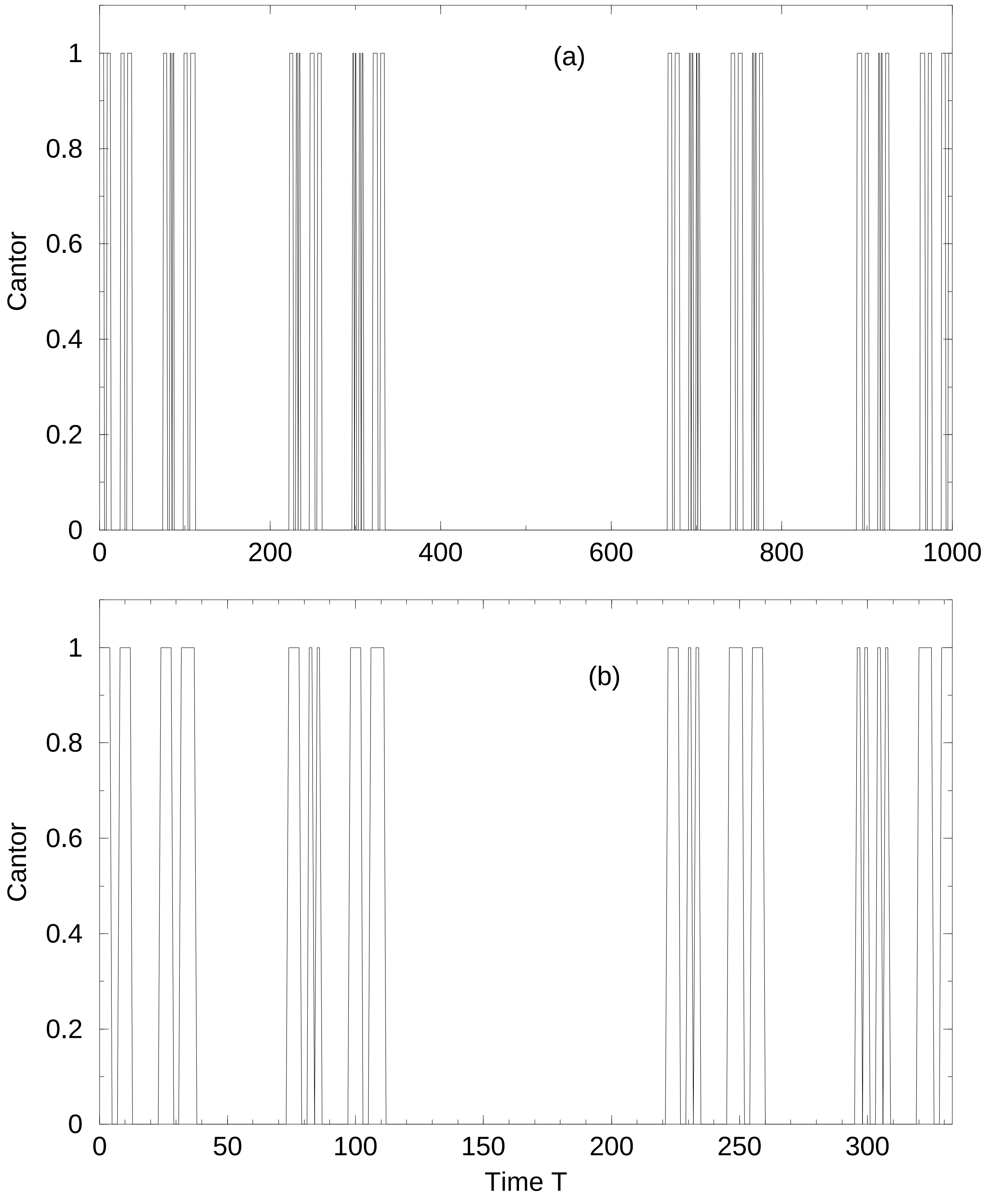
Figure 15 : RCs of P_3 using global SSA. Panels (a–d) represent reconstructions that use groups of RCs as indicated to the left of each panel; panel e contains the sum of RCs 1-8.

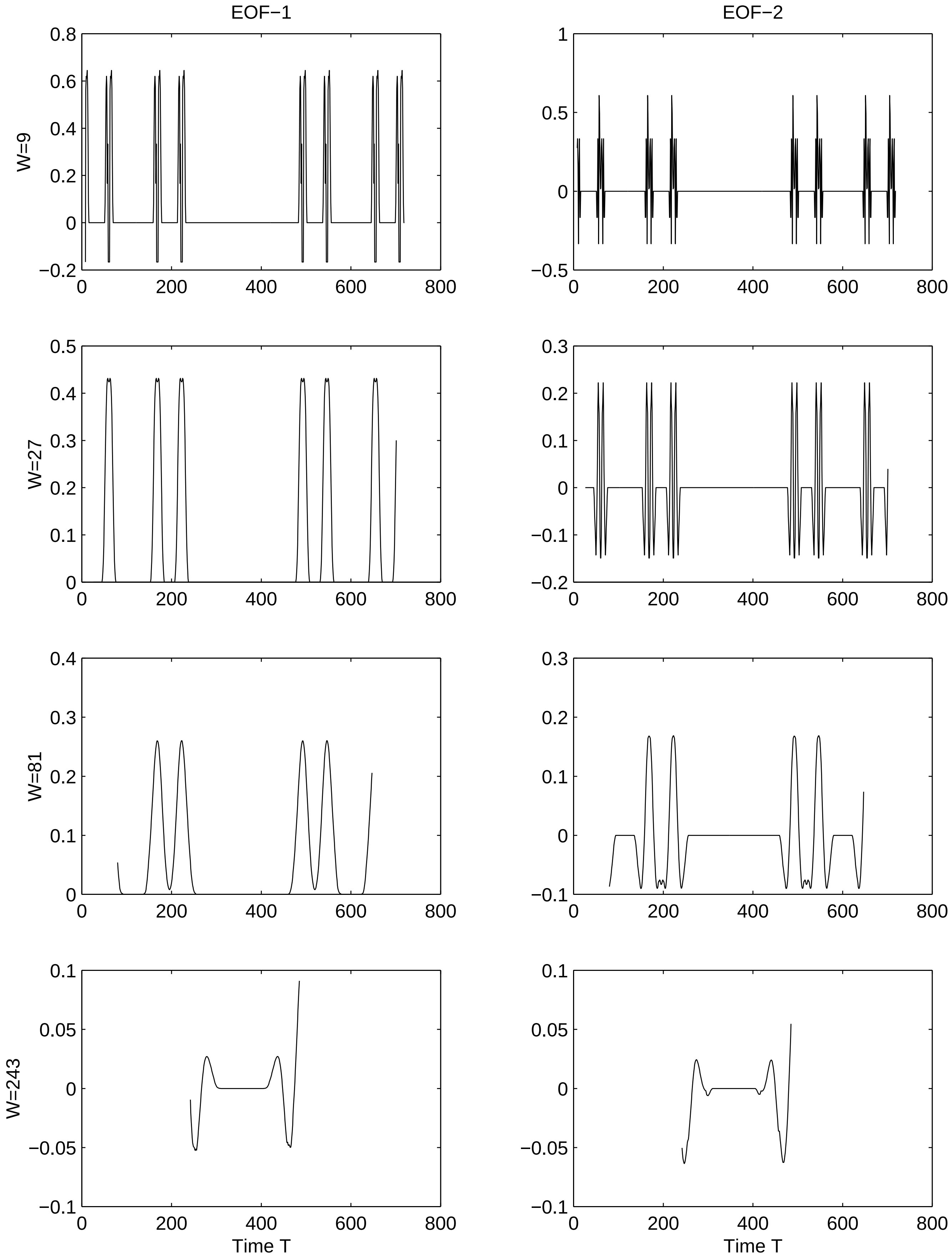
Figure 16 : Wavelet analysis of P_3 with a Gaussian wavelet and its derivative. The scales vary from $a = 2$ to 32 in powers of 2; same layout as for Fig. 7.

Figure 17 : MS-SSA of P_3 with RC-1 and RC-2. The scales W vary from 4 to 64, in powers of 2; same layout as in Fig. 8.

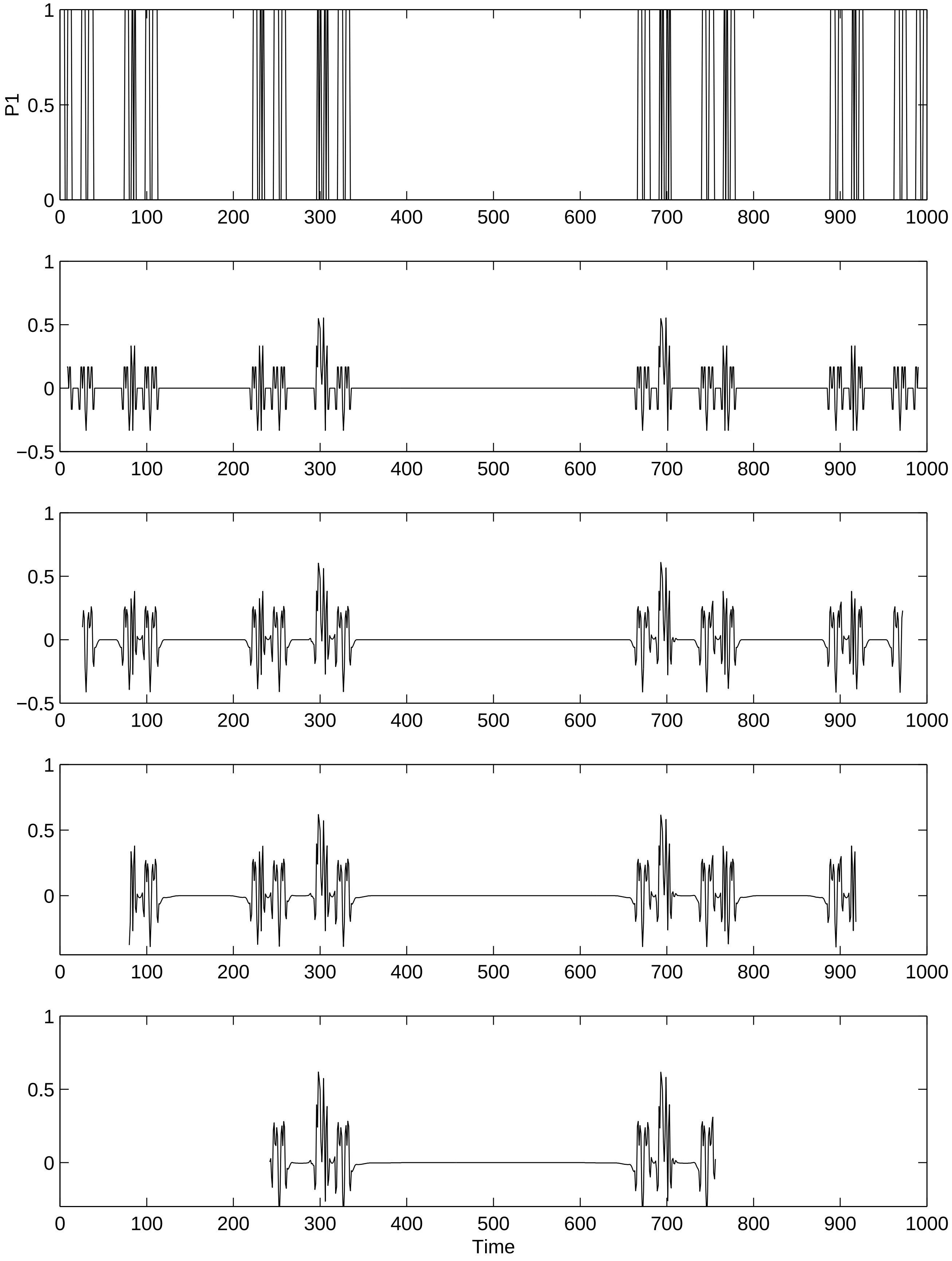
Figure 18 : Evolution in time of the local EOF-1's for P_3, at $W = 32$ ($M = W/3 = 10$). The axes are the time and the lag.

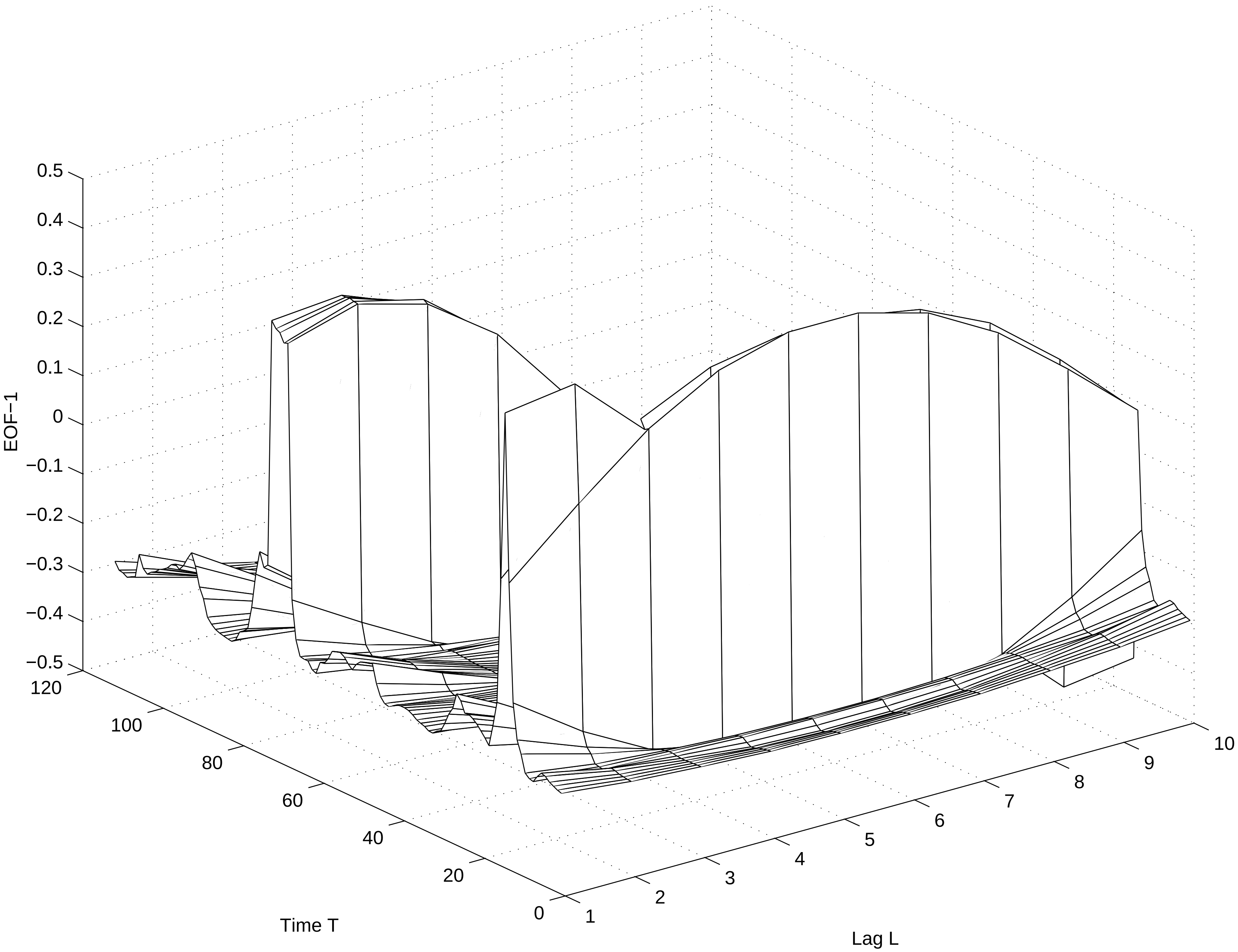
Figure 19 : Instantaneous frequencies of the local EOFs 1 (solid) and 2 (dotted) for the SOI time series in Fig. 4; the MS-SSA scale is $W = 128$ and the number of lags is $M = W/3 = 42$.


(a)
Cantor
0
0.2
0.4
0.6
0.8
1
0
200
400
600
800
1000
(b)
Cantor
0
0.2
0.4
0.6
0.8
1
0
50
100
150
200
250
300
Time T

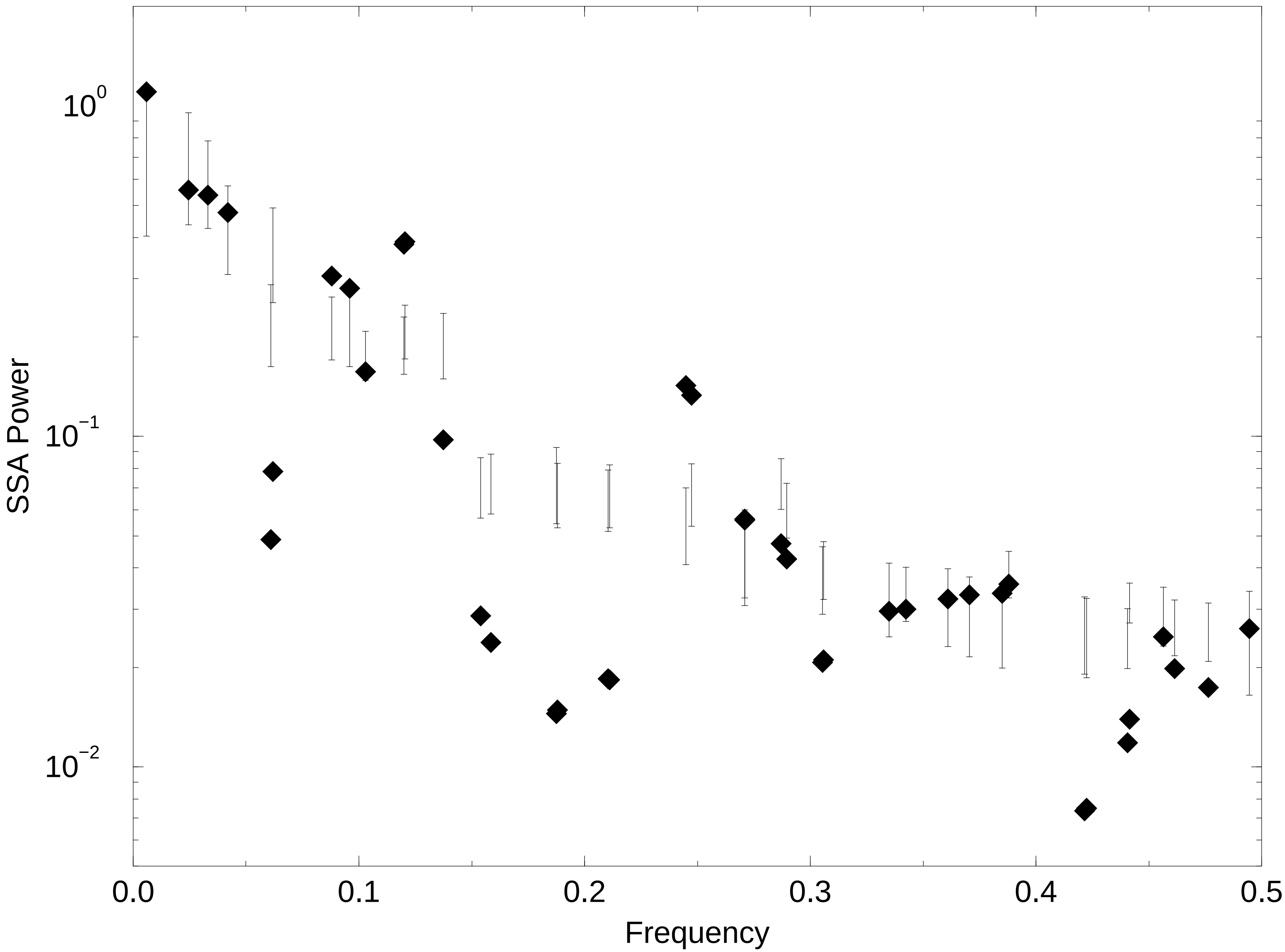


EOF−1
EOF−2
W=9
W=27
W=81
W=243
Time T
Time T
0.8
0.6
0.4
0.2
0
−0.2
1
0.5
−0.5
0.5
0.4
0.3
0.1
−0.1
0.05
−0.05
0
200
400
600
800

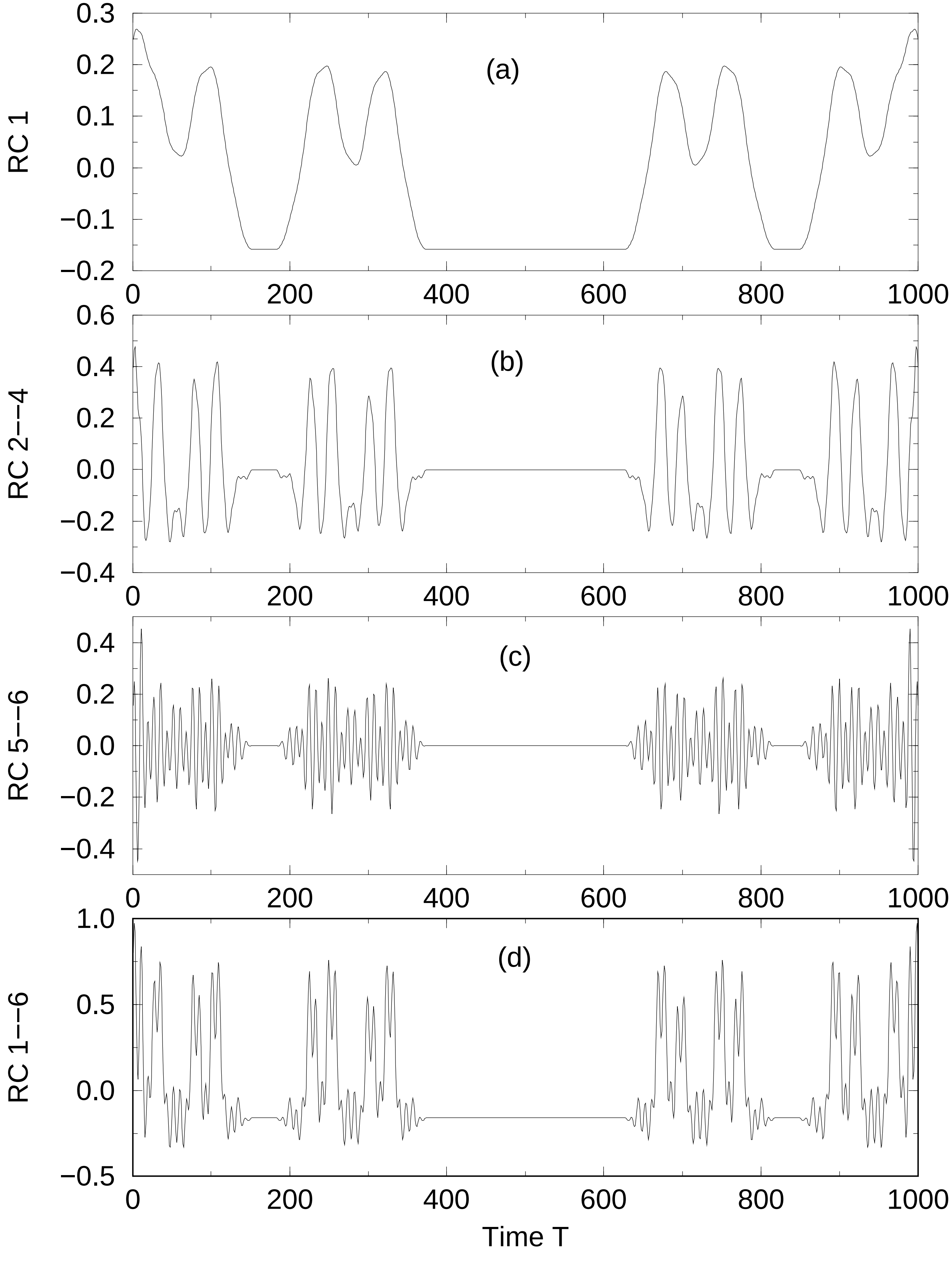

P1
0
0.5
1
−0.5
0
100
200
300
400
500
600
700
800
900
1000
Time

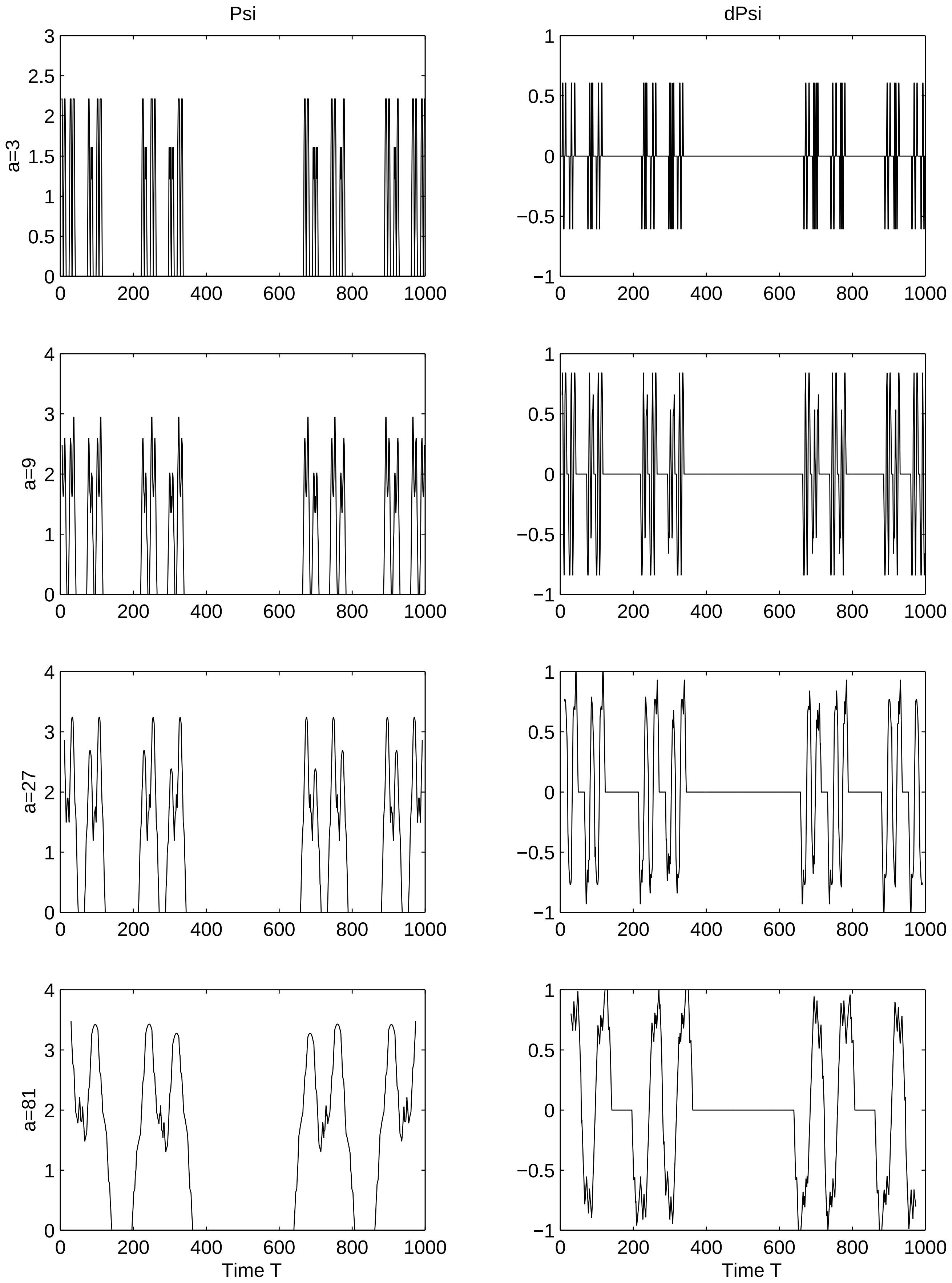


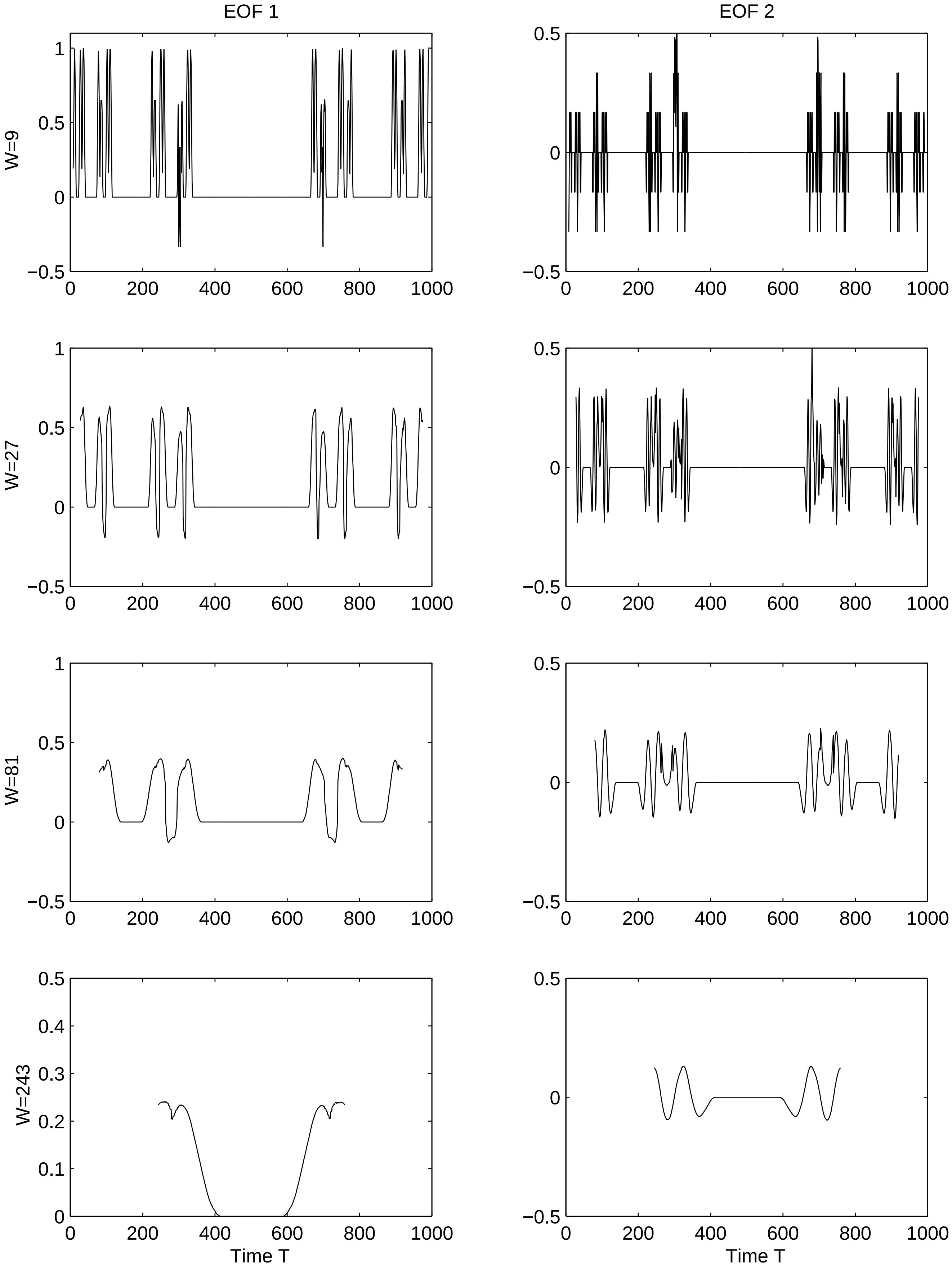
EOF−1
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
120
100
80
60
40
20
0
Time T
1
2
3
4
5
6
7
8
9
10
Lag L



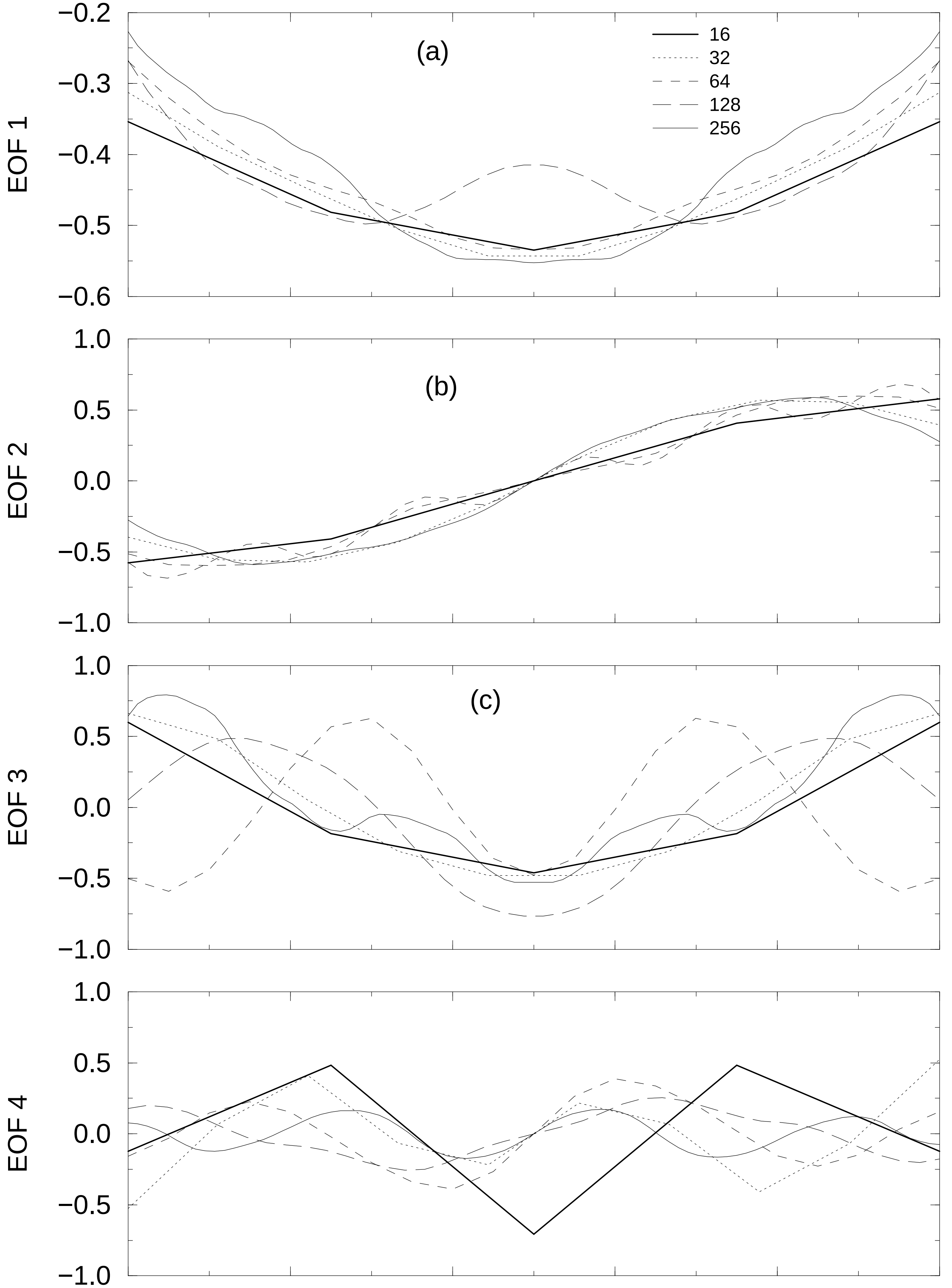
10^0
10^{-1}
10^{-2}
SSA Power
0.0
0.1
0.2
0.3
0.4
0.5
Frequency

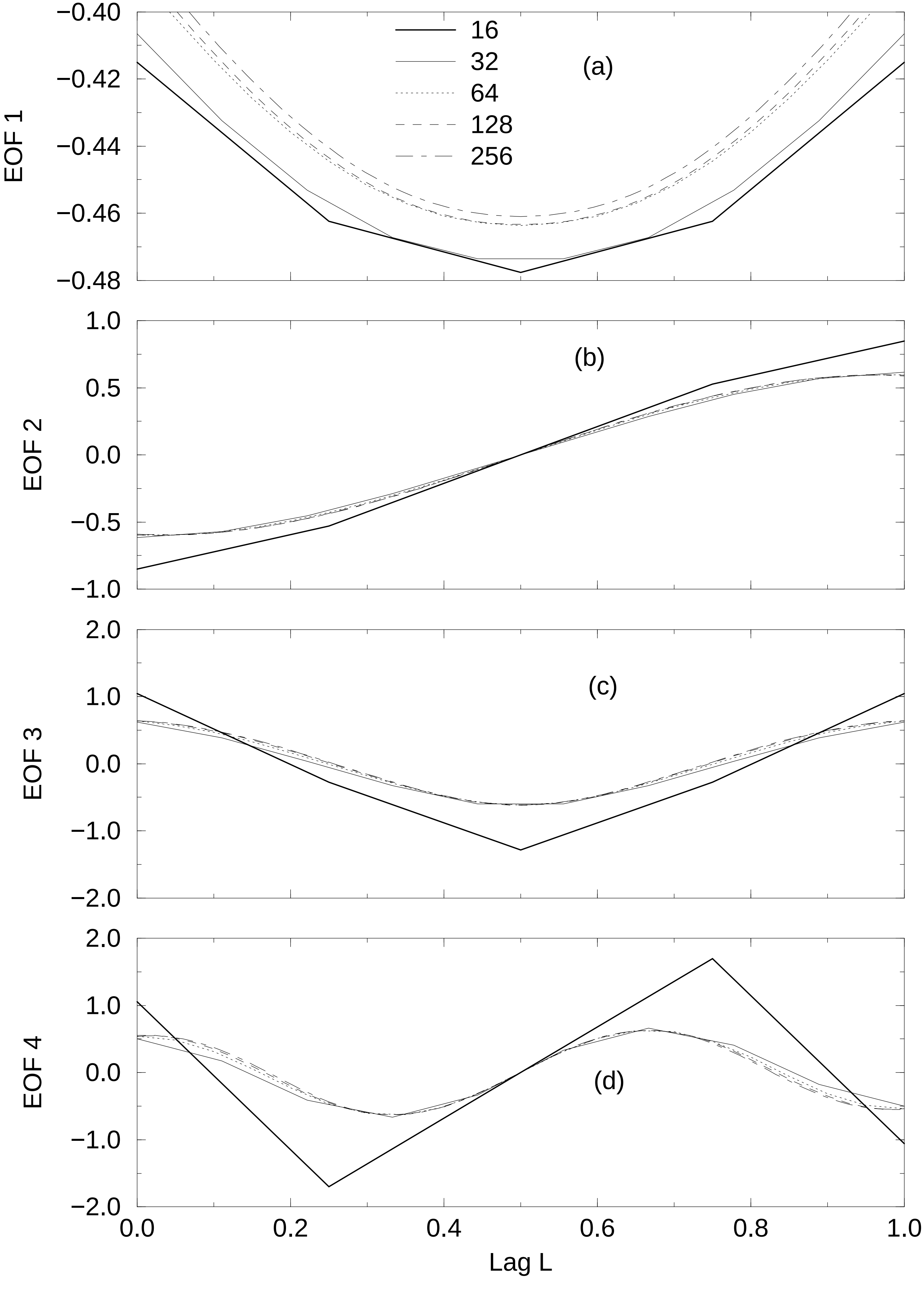

(a)
RC 1
0.3
0.2
0.1
0.0
−0.1
−0.2
(b)
RC 2−−4
0.6
0.4
0.2
0.0
−0.2
−0.4
(c)
RC 5−−6
0.4
0.2
0.0
−0.2
−0.4
(d)
RC 1−−6
1.0
0.5
0.0
−0.5
0
200
400
600
800
1000
Time T

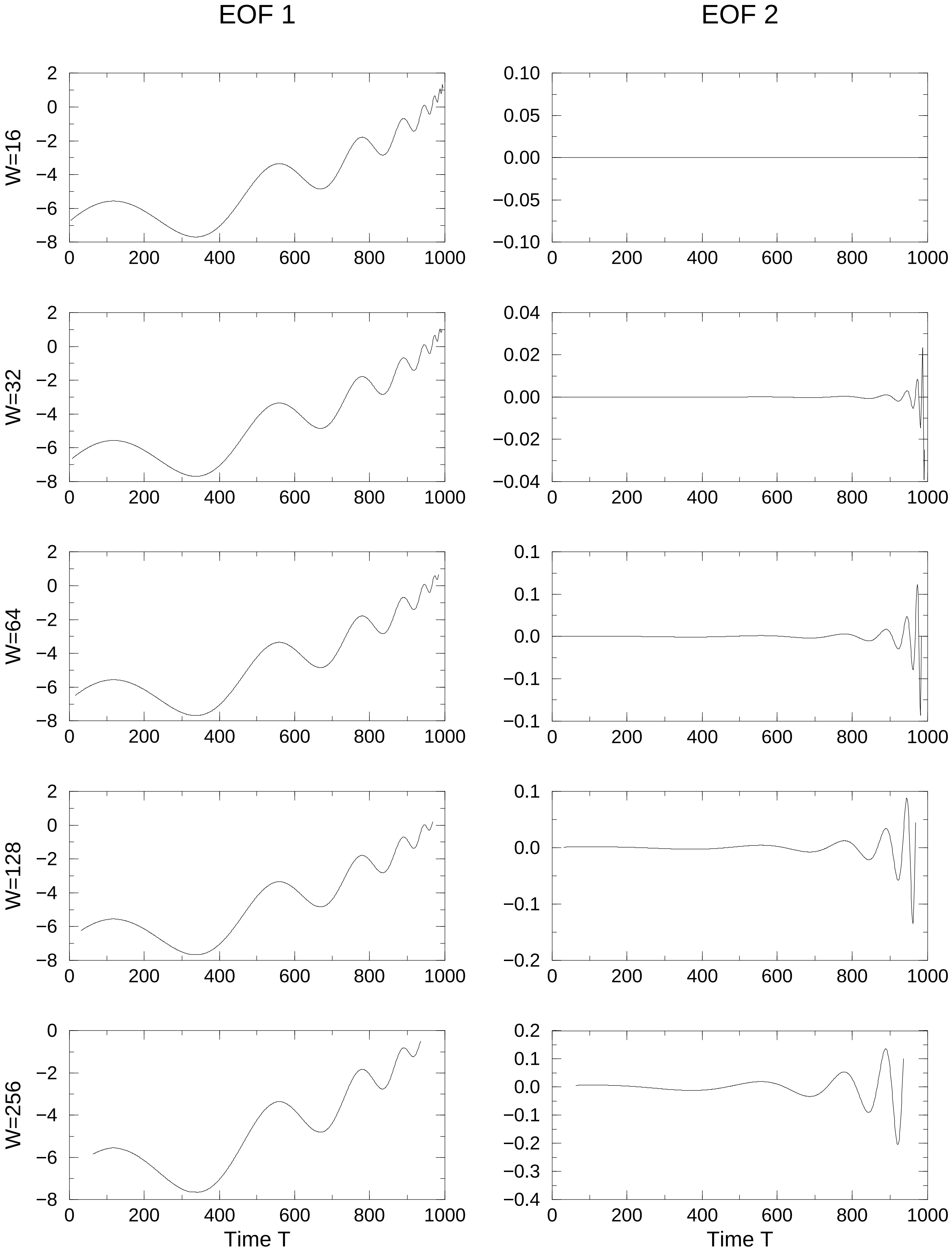

Psi
dPsi
a=3
a=9
a=27
a=81
Time T
Time T
0
200
400
600
800
1000
0
0.5
1
1.5
2
2.5
3
4
−1
−0.5
0
0.5
1

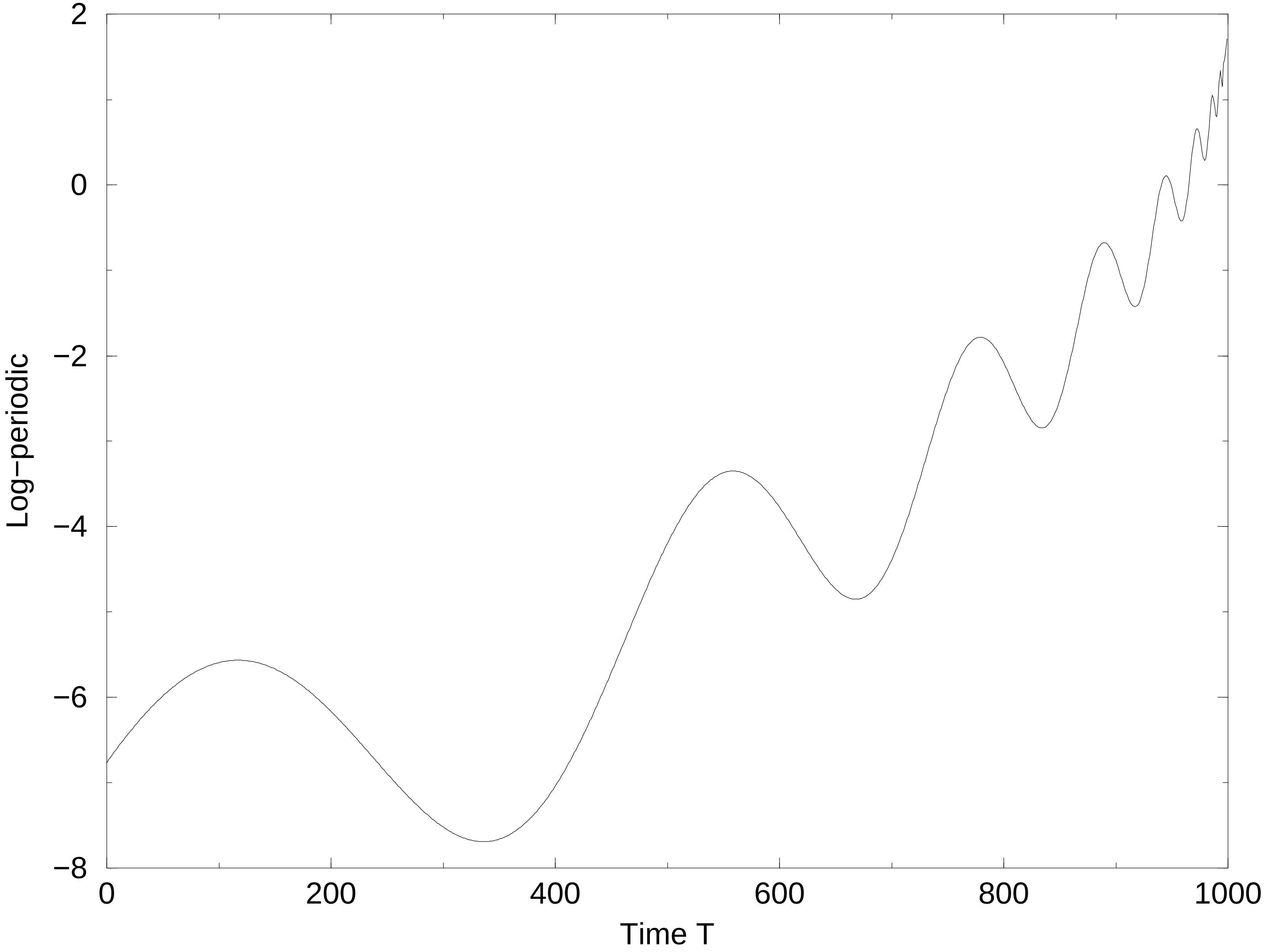


EOF 1
EOF 2
W=9
W=27
W=81
W=243
1
0.5
0
−0.5
0.4
0.3
0.2
0.1
0
200
400
600
800
1000
Time T
Time T

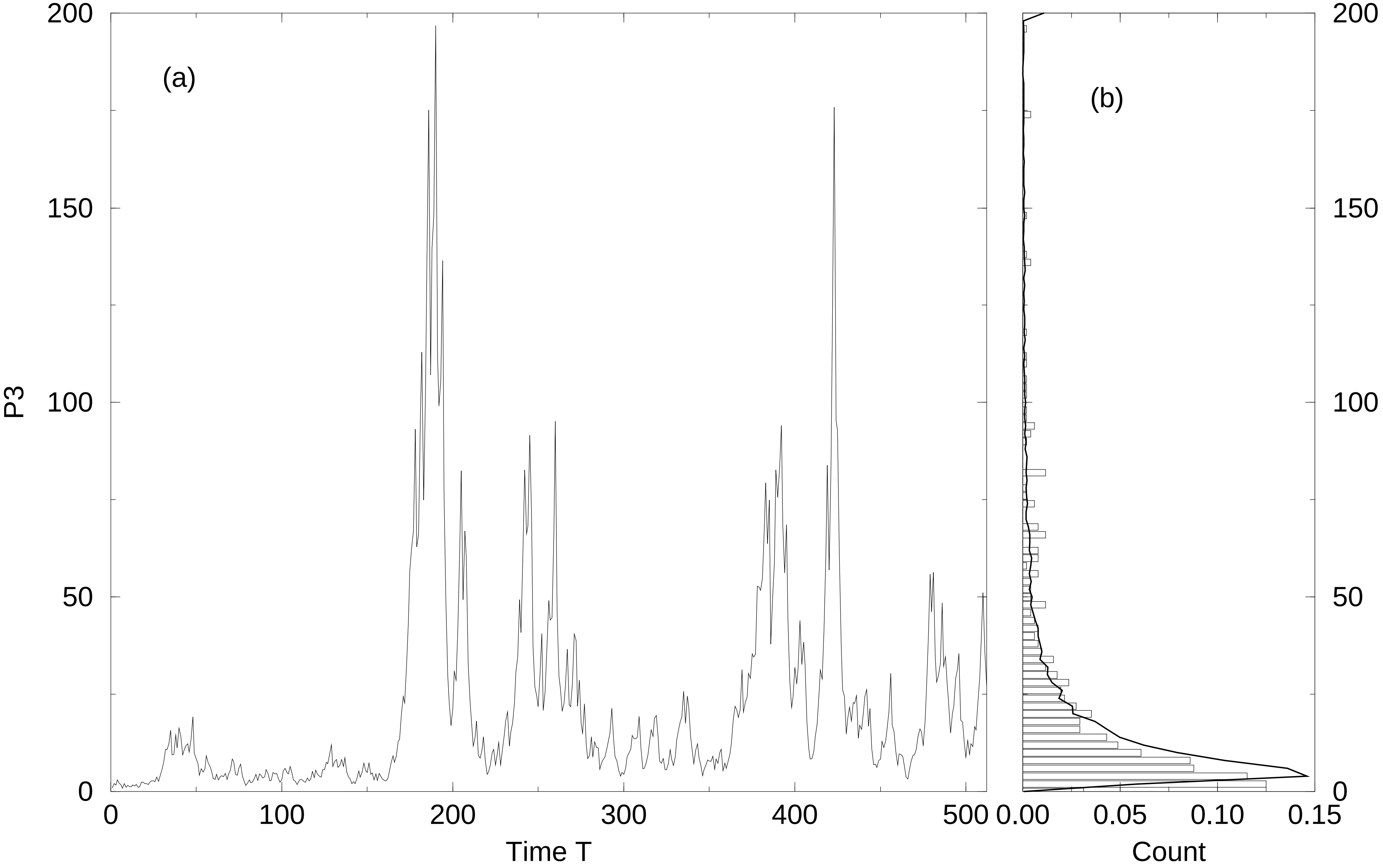

(a)
(b)
(c)
16
32
64
128
256
EOF 1
EOF 2
EOF 3
EOF 4
−0.2
−0.3
−0.4
−0.5
−0.6
1.0
0.5
0.0
−0.5
−1.0
0.0
0.2
0.4
0.6
0.8
1.0
Lag L

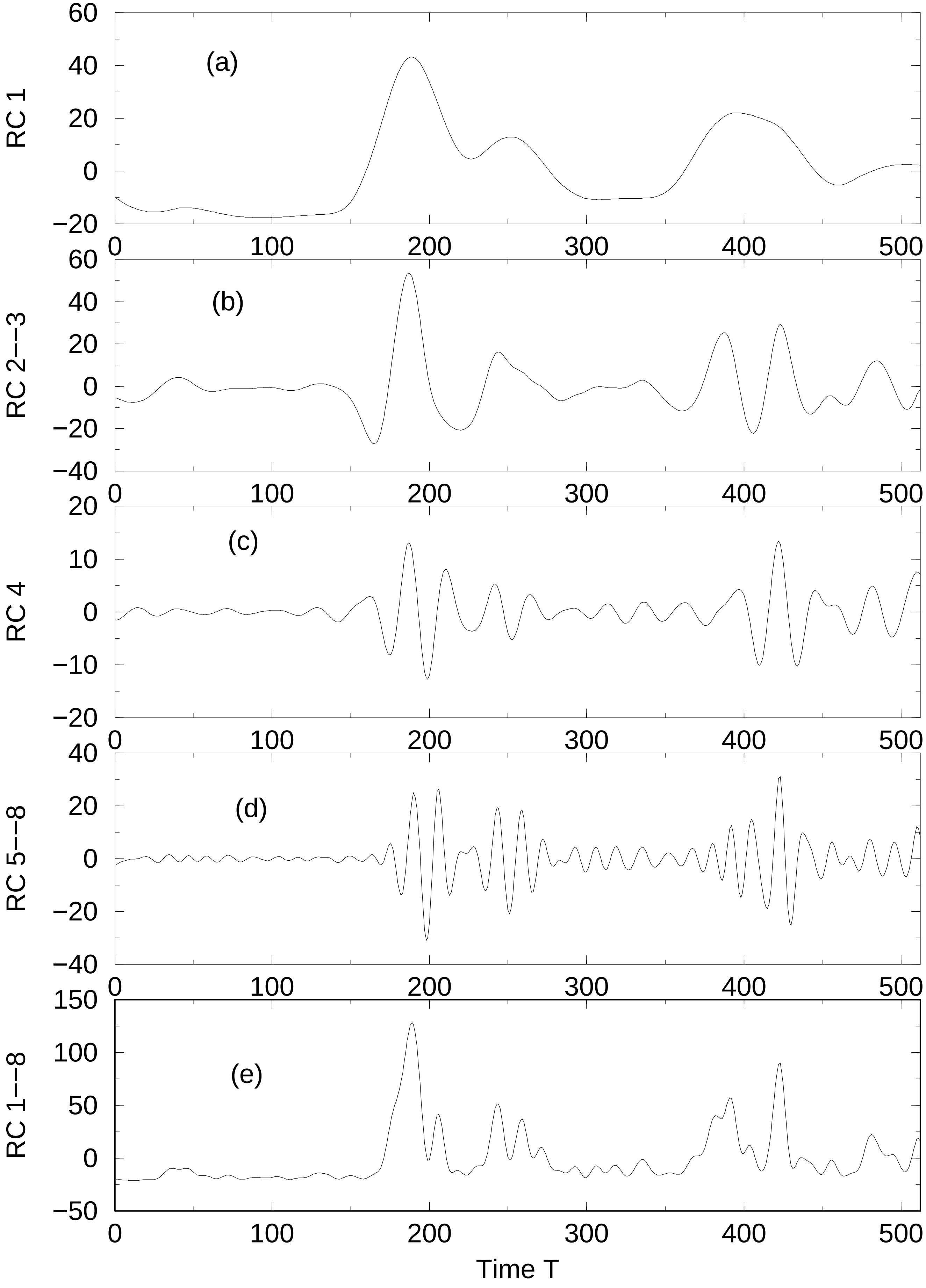

−0.40
−0.42
−0.44
−0.46
−0.48
EOF 1
16
32
64
128
256
(a)
1.0
0.5
0.0
−0.5
−1.0
EOF 2
(b)
2.0
1.0
0.0
−1.0
−2.0
EOF 3
(c)
2.0
1.0
0.0
−1.0
−2.0
EOF 4
(d)
0.0
0.2
0.4
0.6
0.8
1.0
Lag L

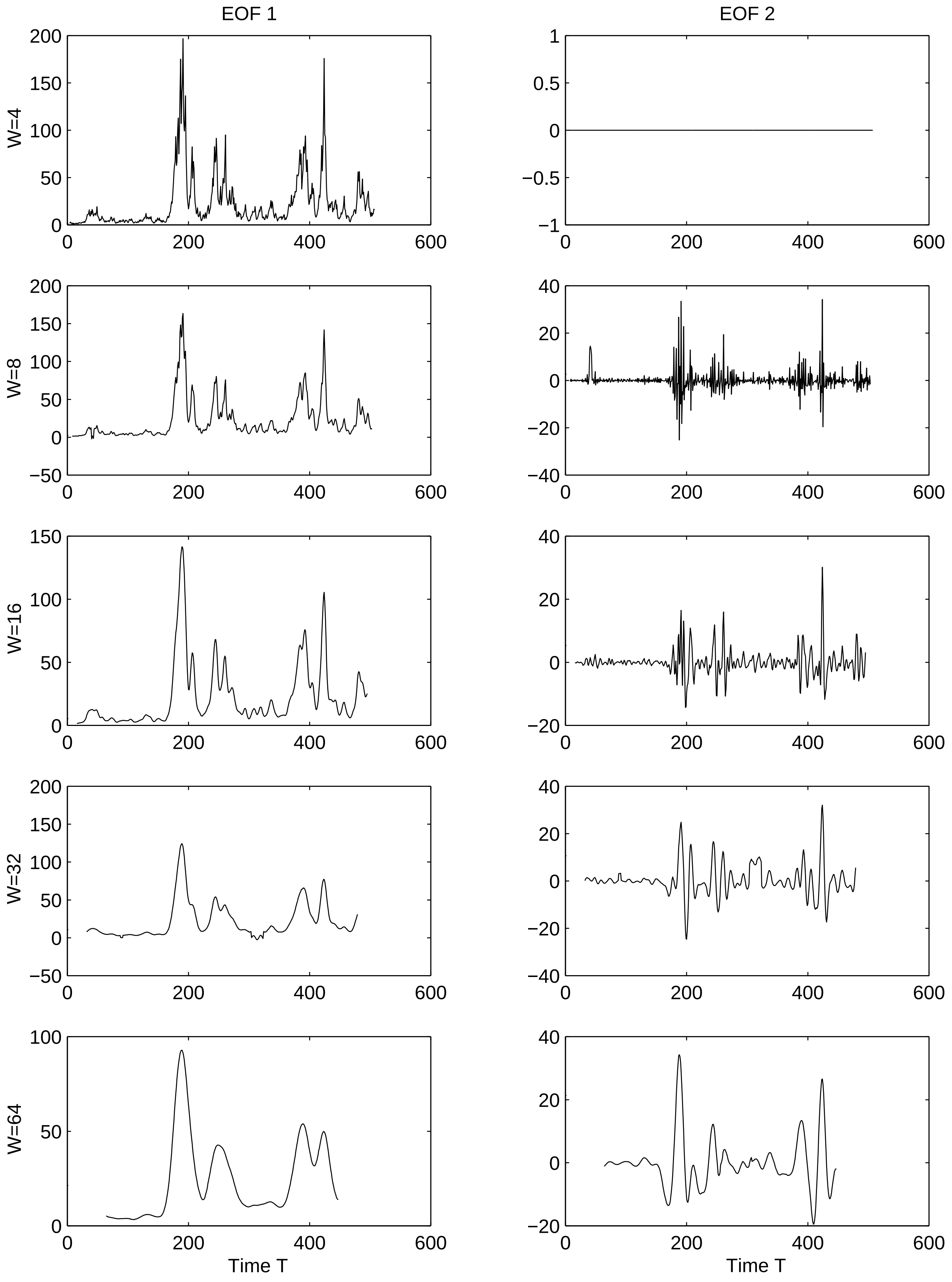

EOF 1
EOF 2
W=16
W=32
W=64
W=128
W=256
2
0
−2
−4
−6
−8
0.10
0.05
0.00
−0.05
−0.10
0.04
0.02
0.00
−0.02
−0.04
0.1
0.1
0.0
−0.1
−0.1
0.1
0.0
−0.1
−0.2
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
0
200
400
600
800
1000
Time T
Time T

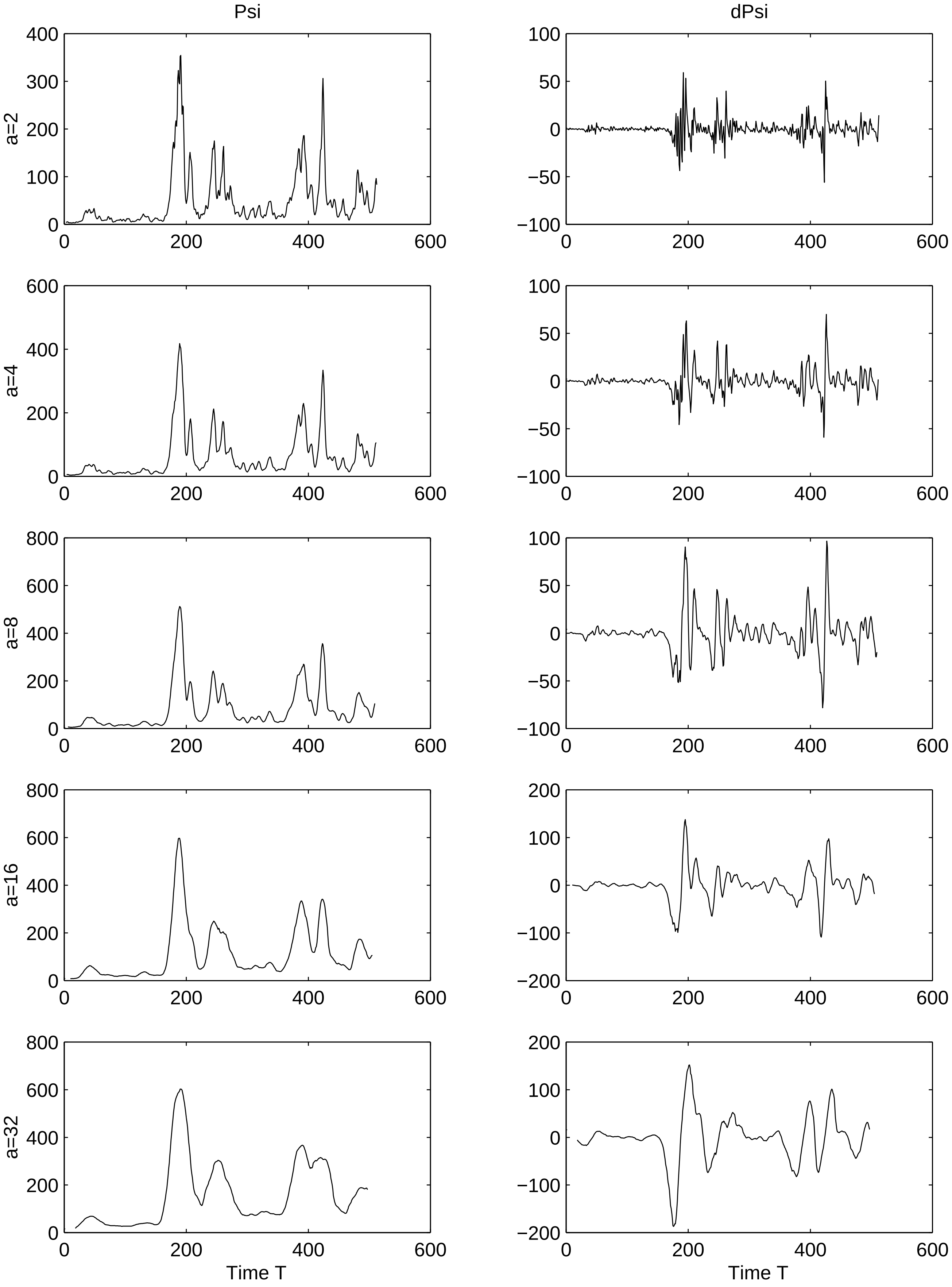


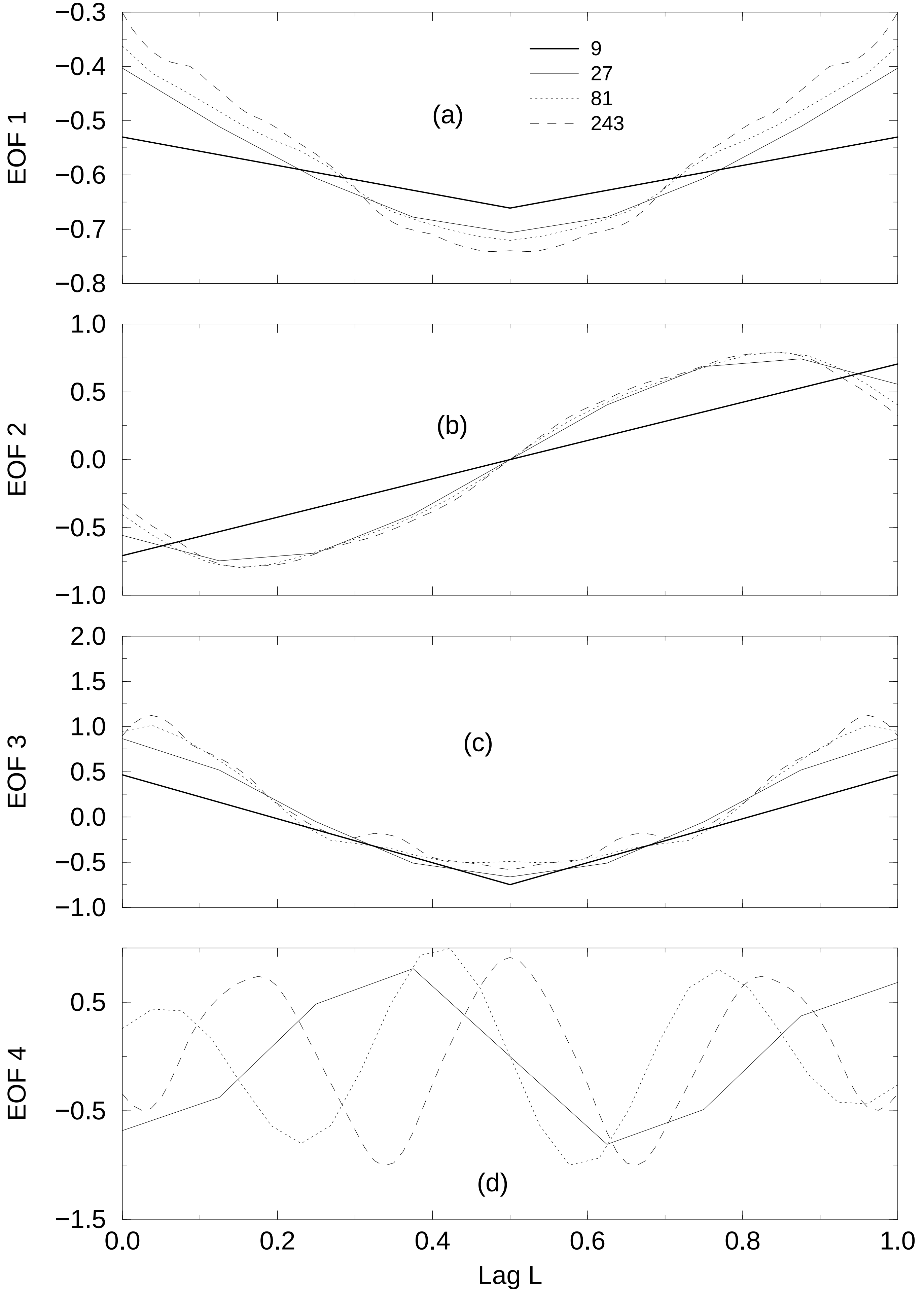
2
0
−2
−4
−6
−8
Log−periodic
0
200
400
600
800
1000
Time T


(a)
(b)
P3
Time T
Count
0
50
100
150
200
0
100
200
300
400
500
0.00
0.05
0.10
0.15
0
50
100
150
200

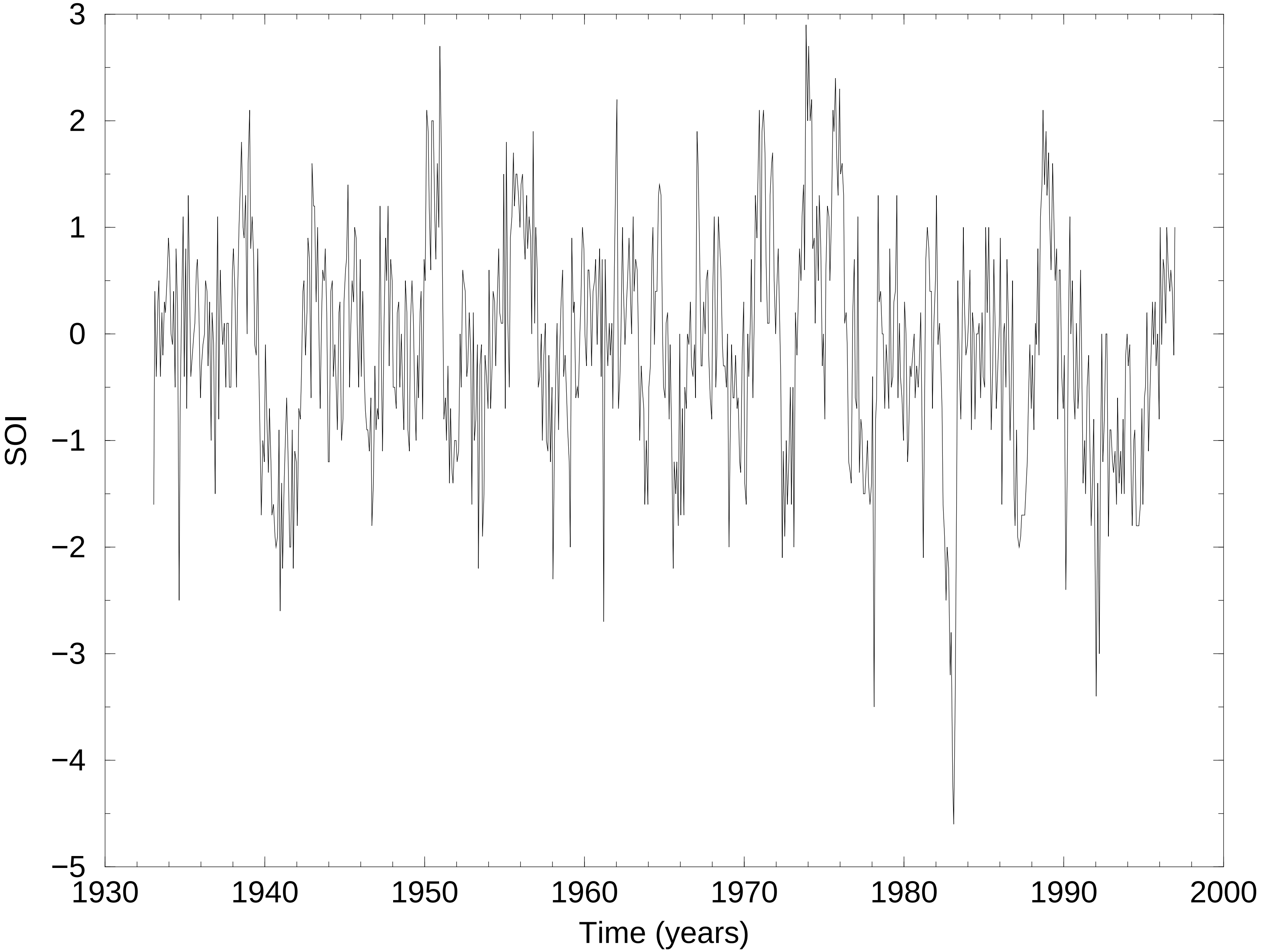

(a)
RC 1
60
40
20
0
−20
0
100
200
300
400
500
(b)
RC 2−−3
60
40
20
0
−20
−40
0
100
200
300
400
500
(c)
RC 4
20
10
0
−10
−20
0
100
200
300
400
500
(d)
RC 5−−8
40
20
0
−20
−40
0
100
200
300
400
500
(e)
RC 1−−8
150
100
50
0
−50
0
100
200
300
400
500
Time T

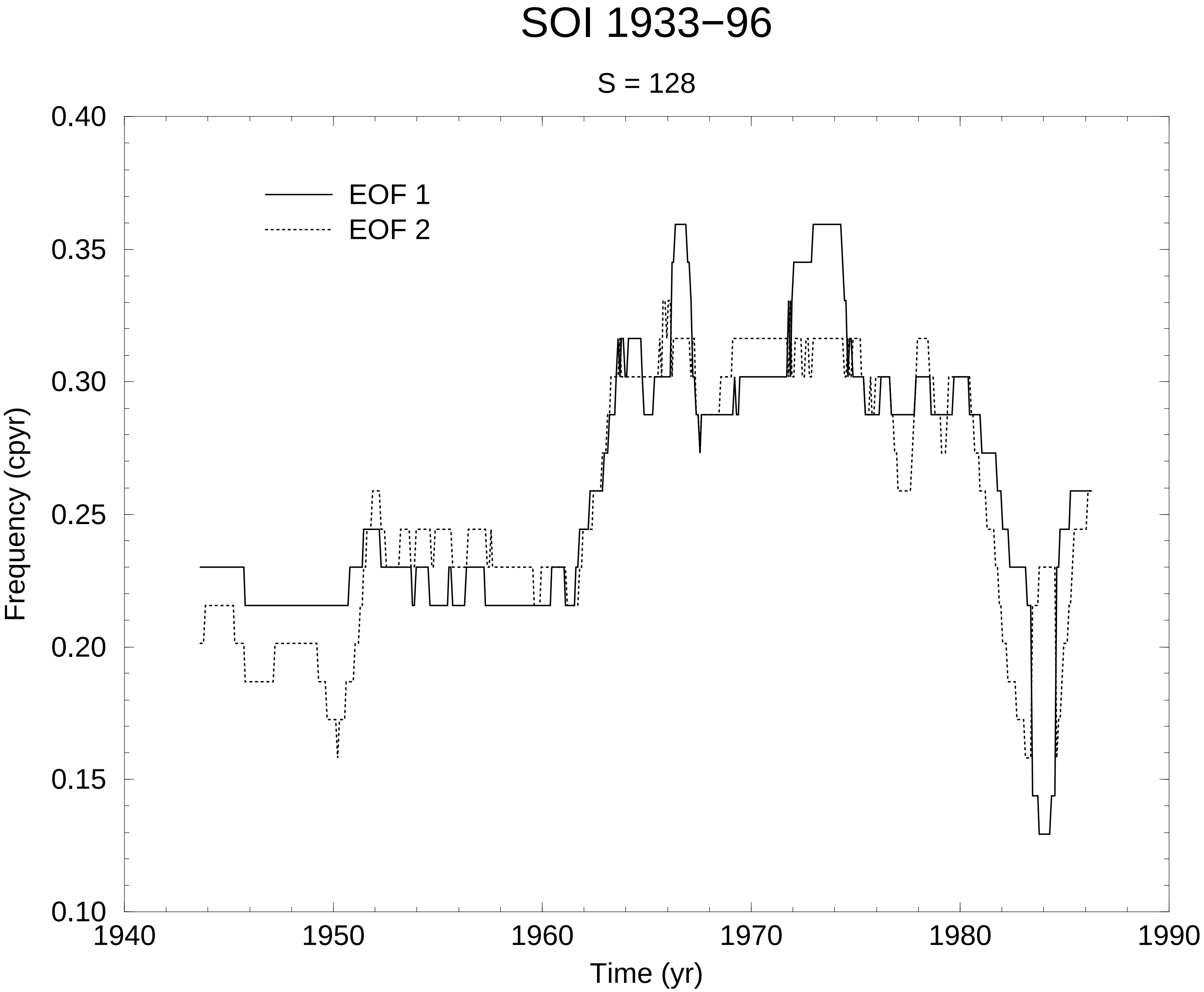

EOF 1
EOF 2
W=4
W=8
W=16
W=32
W=64
Time T
Time T
0
200
400
600
200
150
100
50
0
1
0.5
0
−0.5
−1
−50
40
20
−20
−40
150
100
50
0
40
20
0
−20
200
150
100
50
0
−50
40
20
0
−20
−40
100
50
0
40
20
0
−20


Psi
dPsi
a=2
a=4
a=8
a=16
a=32
Time T
Time T
0
200
400
600
100
300
800
50
−50
−100
−200



EOF 1
EOF 2
EOF 3
EOF 4
(a)
(b)
(c)
(d)
9
27
81
243
−0.3
−0.4
−0.5
−0.6
−0.7
−0.8
1.0
0.5
0.0
−0.5
−1.0
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
0.5
−0.5
−1.5
0.0
0.2
0.4
0.6
0.8
1.0
Lag L


3
2
1
0
−1
−2
−3
−4
−5
SOI
1930
1940
1950
1960
1970
1980
1990
2000
Time (years)



SOI 1933−96

S = 128