Multivariate singular spectrum analysis and the road to phase synchronization

Andreas Groth\(^{1,2,*}\) and Michael Ghil\(^{1,2,3}\)

\(^{1}\)Geosciences Department, Ecole Normale Supérieure, Paris, France
\(^{2}\)Environmental Research & Teaching Institute, Ecole Normale Supérieure, Paris, France
\(^{3}\)Department of Atmospheric & Oceanic Sciences and Institute of Geophysics & Planetary Physics, University of California, Los Angeles, USA

(Dated: October 13, 2010)

We illustrate the ability of multivariate singular spectrum analysis (M-SSA) to automatically identify multiple oscillatory modes and their transition to phase synchronization. Without using any a priori definition of a suitable phase, M-SSA is able to detect phase synchronization in a chain of coupled chaotic oscillators. We demonstrate that a rotation of the M-SSA eigenvectors is necessary and thereby allows a correct identification of clusters of synchronized oscillators.

PACS numbers: 05.45.Xt, 02.50.Sk

Over the last two decades, singular spectrum analysis (SSA) and multichannel SSA (M-SSA) have proven their usefulness in the temporal and spatio-temporal analysis of short and noisy time series in several fields of the geosciences and of other disciplines. M-SSA provides insight into the unknown or partially known dynamics of the underlying system by decomposing the delay-coordinate phase space of a given multivariate time series into a set of data-adaptive orthonormal components. These components can be classified essentially into trends, oscillatory patterns and noise, and allow one to reconstruct a robust “skeleton” of the dynamical system’s structure [1, 2]. In this Letter, we show that M-SSA can greatly help phase synchronization analysis and provides considerable insight into the mechanisms of rhythm adjustment in a chain of coupled chaotic oscillators.

Several ways of measuring phase synchronization are in use, and the robust definition of a reasonable phase for each oscillator is critical in all of them [3, 4]. In general, the phase is defined as an angle around an origin of rotation in phase space. This approach, however, requires an a priori information about the dynamical system, e.g. by visual inspection, and a reasonable definition can become difficult to formulate for a high-dimensional system.

SSA and M-SSA rely on the classical Karhunen-Loève spectral decomposition of time series. Broomhead and King [5, 6] introduced them into dynamical systems analysis, as a more robust version of the Mañé-Takens idea to reconstruct dynamics from a single time series. We focus here on M-SSA, which we summarize for completeness.

Let \(X = \{x_d(n) : d = 1 \ldots D, n = 1 \ldots N\}\) be a multivariate time series with \(D\) channels of length \(N\). We assume that each channel has been centered and normalized. Following Mañé and Takens [5, 6], the starting point of M-SSA is to embed each channel into an \(M\)-dimensional phase space, by using lagged copies \(X_d(n) = (x_d(n), x_d(n+M-1)), n = 1, \ldots, N-M+1\). From this we form the full augmented trajectory matrix \(X = (X_1, X_2, \ldots, X_D)\), which has \(DM\) columns of length \(N-M+1\).

The M-SSA procedure then computes the covariance matrix \(C = X^\top X\) of \(X\) and its eigendecomposition; here \((\cdot)^\top\) is the transpose. Due to finite-size effects, the sample \(C\) may deviate slightly from symmetry, but appropriate estimation methods appear in Ref. [2]. The covariance matrix \(C\) combines all auto- as well as cross-covariances, up to a time lag equal to \(M-1\). Next, one diagonalizes the appropriately symmetrized covariance matrix

\[
\Lambda = E^\top C E \tag{1}
\]

to yield a diagonal matrix \(\Lambda\) that contains the real eigenvalues \(\lambda_k\) of \(C\), and a matrix \(E\) whose columns are the associated eigenvectors \(e_k\). The \(e_k\)s form a new orthogonal basis in the embedding space of \(X\), and the corresponding \(\lambda_k\)s give the variance in the direction of \(e_k\).

The spectral decomposition in Eq. (1) determines the directions of greatest variance successively, from largest to smallest, subject to the condition that each new direction be orthogonal to all the preceding ones.

To illustrate the insights provided by M-SSA into phase synchronization, we consider a chain of diffusely coupled and detuned Rössler oscillators [7]:

\[
\begin{align*}
\dot{x}_j &= -\omega_j y_j - z_j + c(x_{j+1} - 2x_j + x_{j-1}), \\
\dot{y}_j &= \omega_j x_j + 0.15y_j, \quad \dot{z}_j = 0.1 + z_j(x_j - 8.5).
\end{align*} \tag{2}
\]

The position in the chain is given by the index \(j = 1, \ldots, J\); \(\omega_j = \omega_1 + 0.02(j - 1)\) are the associated natural frequencies, with \(\omega_1 = 1\), and we assume free boundary conditions: \(x_0(n) = x_1(n), x_{J+1}(n) = x_J(n)\).

We first integrate the uncoupled system, with \(c = 0\), in order to see whether M-SSA is able to distinguish between the different oscillators. The individual oscillators, while being chaotic, are in a phase-coherent regime, with a well-defined center of oscillation in the \((x-y)\)-plane [3].

We consider \(J = 5\) oscillators, integrate Eq. (2) with variable step size, and sample the solution at equal intervals \(\Delta t = 0.4\). The observed time series \(x\) has \(D = 3J = 15\) channels, and we take \(N = 2500\).

To cover more than one oscillation period, we select a window length of \(M = 30\), and from Eq. (1) we thus
derive $DM = 450$ eigenvalues and eigenvectors, respectively; the 40 largest eigenvalues are plotted in Fig. 1. The leading 10 eigenvalues are clearly significant and fall into five pairs of nearly equal ones. These pairs correspond to an oscillatory mode, provided the associated eigenvectors correspond to the same period [2, 8]. Such pairs are the data-adaptive equivalent of sine-and-cosine pairs in Fourier analysis.

The eigenvalue spectrum in Fig. 1 thus suggests that our M-SSA has correctly identified the five uncoupled oscillators in Eq. (2), each of them described by an oscillatory pair, adapted to the natural frequency. An inspection of the corresponding eigenvectors in Fig. 2, however, shows a more ambiguous picture: The eigenvectors do not distinguish between the uncoupled systems, and a projection of $X$ onto the eigenvectors, given by $A = XE$ — the columns of $A$ being referred to as principal components (PCs) — would mix up distinct systems.

This mixture is a general shortcoming of PC analysis (PCA) in general and of M-SSA in particular. These methods are designed to capture a maximum of the variance in the dataset with a minimum of PCs [2, 9]: while good for signal compression purposes, they limit the physical interpretation of the underlying dynamical system. In the example at hand, all five modes have the same variance, and thus signal compression does not necessarily reveal the correct underlying structure, only a mixture of degenerate eigenvectors [9].

Recently, PCA has been applied to help detect synchronized clusters of coupled oscillators [10]. The results suffer precisely from these mixing problems and the authors discuss a heuristic approach of setting eigenvector coordinates to zero and rerunning the analysis. We rely here instead on eigenvector rotation [11], as widely used in PCA, and focus on varimax orthogonal rotation [12].

Varimax attempts to simplify the eigenvectors’ structure and thus reduce mixture effects by finding an appropriate orthogonal rotation $E^* = ET$, where the $e_{mk}^*$ are the coordinates of the rotated eigenvectors $e_k^*$. One wishes to bring the coordinates of each $e_k^*$ either close to one or to zero and to suppress intermediate values that get in the way of a clear classification.

In PCA as applied to random fields [11] or to factor analysis [12], one simply maximizes the variance of the squared coordinates,

$$V_1 = \frac{S}{\sum_{k=1}^{S} \sum_{m=1}^{DM} \left( \frac{e_{mk}^*}{h_m^*} \right)^2} - \left( \frac{1}{DM} \sum_{m=1}^{DM} h_m^* \right)^2,$$

where $S$ is the number of rotated $e_k^*$’s, and $h_m^* = \sum_{k=1}^{S} e_{mk}^* \neq 0$ is the corresponding normalization. Kaiser [12] gives an explicit equation for the sequential rotation of pairs of $e_k^*$ that shows this algorithm’s superiority over more sophisticated optimization procedures, such as independent component analysis.

In the criterion $V_1$ of [11, 12], the variance is maximized over all the coordinates of the $e_k^*$’s. An important difference in the structure of classical PCA eigenvectors and M-SSA ones, however, impedes the direct application of varimax to the latter: in classical PCA, each coordinate in an eigenvector of dimension $D$ is related to a different channel in $X$, while in M-SSA, the eigenvectors have length $DM$ and are composed of $D$ consecutive segments of length $M$, each of which is associated with a channel.

Each segment of length $M$ is thus a filter that extracts a part of the time series’ power spectrum. M-SSA eigenvectors are, therefore, able to adapt to oscillatory behavior in terms of oscillatory pairs, and it is especially this property that makes M-SSA superior to PCA in phase synchronization analysis. Varimax, however, attempts also to increase the variance within each of these filters, which would lead to a loss of spectral properties, such as oscillatory pairs [2, 8].

To avoid this loss of a key M-SSA property, we propose here a simple but effective modification of the varimax criterion. Prior to the calculation of the variance, we sum over the individual filters $e_{dk}^* = \sum_{m=1}^{dM} e_{mk}^*$,
the PCs adapted to the natural frequencies of the oscillators and

e Fig. 3. In contrast to the unrotated

tion. The rotated

e rotated ones are clearly associated pairwise to each of

the first $S = 10$ eigenvectors.

and the criterion becomes

\[
V_M = \sum_{k=1}^{S} \left\{ \frac{1}{D} \sum_{d=1}^{D} \left( \frac{\bar{e}_{dk}}{h_d^*} \right)^2 - \left( \frac{1}{D} \sum_{d=1}^{D} \bar{e}_{dk}^2 \right) \right\},
\]

with the normalization $h_d^{*2} = \sum_{k=1}^{S} e_{dk}^2$; for $M = 1$ we recover the original criterion $V_1$. With the modified criterion above, and following the appendix in Ref. [12], we have simply to replace — in the equation of pairwise rotation of two eigenvectors $e_i$ and $e_j$, — each $\bar{u}_{mj} = e_{mj}^2$ by its averaged version $\bar{e}_{di}^2 + \bar{e}_{dj}^2$, as well as $\bar{u}_m^2 = 2e_{mi}e_{mj}$ by $2\bar{e}_{di}\bar{e}_{dj}$.

Typically, the rotation is restricted to the first $S$ eigenvectors, which capture most of the systems’ variance. In order to find an appropriate subset of eigenvectors, it is common to look for a gap in the eigenvalue spectrum. Such an obvious gap as in Fig. 1, however, cannot be expected in general, especially in applications to short and noisy time series [2]. In our experience with real data, dividing each eigenvector $e_k$ by the singular value $\lambda_k^{1/2}$, prior to varimax, stabilizes the rotation results over a large range of S-values. Without this scaling, varimax would be susceptible to “overrotation” [13] and could end up with too many clusters for large $S$.

The scaling is only used to derive the rotation matrix $T$; we apply this $T$ to rotate the leading $S$ original eigenvectors, namely $E_S^* = E_S T$. There are two reasons for doing so: first, scaling, rotating and rescaling leads to an oblique rotation, which performs worse than an orthogonal rotation; second, an oblique rotation does not preserve the variance captured by $E_S$.

The result of the eigenvector rotation is shown in Fig. 3. In contrast to the unrotated $e_k$’s in Fig. 2, the rotated ones are clearly associated pairwise to each of the uncoupled oscillators and allow a unique identification. The rotated $e_k$’s form oscillatory pairs that are well adapted to the natural frequencies of the oscillators and the PCs $A_S^* = XE_S^*$ fit very well the cyclical motion.

Next, in order to obtain the variance that the $e_k$’s capture, we project the covariance matrix $C$ onto them, $A_S^* = E_S^T C E_S^*$, like in Eq. (1). Since the rotated eigenvectors $E_S^*$ are not part of the singular value decomposition of the covariance matrix $C$, the matrix $A_S^*$ is not diagonal. From the equivalent formulation $A_S^* = A_S^* T A_S^*$ we see, however, that $A_S^*$ does comprise the variance of the rotated PCs along its diagonal, which we denote by $\lambda_k^*$. Furthermore, the equivalent formulation $A_S^* = T^T A_S T$ is a similarity transformation to $A_S$, the diagonal matrix of the $S$ leading eigenvalues of $A$; hence the trace (i.e., the sum of variances) is preserved.

Since M-SSA, like all PCA methods, tries to maximize variance captured by successive eigenvectors, it is also prone to artificial variance compression, i.e. it may account for too much (little) variance in the largest (smallest) eigenvalues. This is also reflected in the inequality of the $10$ largest $\lambda_k$’s in Fig. 1, although the oscillators have the same variance. The values $\lambda_k^*$, on the other hand, are almost perfectly balanced and reflect the correct situation. Even a rotation of many more eigenvectors, such as $S = 40$, yields the same picture (not shown).

So far, we discussed the case of uncoupled oscillators and showed the ability of M-SSA to correctly identify the distinct oscillatory modes. We present next its ability to correctly describe the transition to phase synchronization, and consider $J = 10$ Rössler oscillators (2). It is known that even a relatively large frequency mismatch of $\Delta \omega = 0.02$ yields a suppression of chaos as the coupling strength $c$ increases, and synchronization occurs via clustering [7]: the oscillators form clusters with the same frequency and a relatively large $\Delta \omega$ between clusters. As $c$ increases further, the number of clusters decreases in a cascade of vanishing frequencies; see Fig. 4(a).

This behavior is also reflected in a cascade of vanishing pairs of eigenvalues $\lambda_k$ (Fig. 4(b)). Each line in the figure is actually the superposition of two lines of nearly identical eigenvalues, and the number of significant pairs represents the number of clusters. The explanatory power of M-SSA, though, is quite limited in the absence of rotation. For weak coupling ($c \leq 0.15$), prior to any clustering, the spread among the significant $\lambda_k$’s increases with $c$. As clustering sets in at $c \approx 0.03$, the behavior at the successive bifurcation points is fairly smooth, with continuously decreasing eigenvalues and no distinct transitions; e.g., at $c = 0.05$ there is no obvious gap that would allow a clear counting of the clusters.

After rotation, however, the picture reflects much better the transition to phase synchronization, with sharp jumps at the bifurcation points (Fig. 4(c)). For intermediate coupling strength ($c \approx 0.04$), we observe fluctuating behavior that is almost absent in Fig. 4(a). These fluctuations in the $\lambda_k$’s reflect phase-locking fluctuations that do actually occur near the bifurcations (not shown), in spite of the frequency locking demonstrated by Fig. 4(a) [14].
The results above are robust with respect to observational noise. Even when noise contamination is high, say Gaussian noise with the same standard deviation as the signal is added, Figs. 4(b),(c) remain visually the same. It is, furthermore, a property of PCA in general that the eigenvalues are invariant with respect to an arbitrary orthogonal rotation of the time series $x$. Numerical tests also confirm the invariance of M-SSA eigenvalues; e.g., Figs. 4(b),(c) do not change visually when rotating $x$. The analysis in Fig. 4(a), on the contrary, is much more sensitive to observational noise, and is not invariant with respect to rotation.

In order to gain information about the participation of each of the oscillators in a cluster, we compute the so-called reconstructed components (RCs) by an inverse transformation of the PCs [2, 8]. In Fig. 5, there are three clusters, which comprise oscillators 1–3, 6–7, and 9–10, respectively. Within each cluster, the oscillators exhibit only a single peak of their spectral density, e.g. of the $x$–components. At the boundaries between clusters, the oscillators 5 and 8 are not uniquely classified: they show fairly intermittent behavior and the trajectories exhibit beats, while their power spectra show two peaks, with frequencies that agree with those of the two adjacent clusters, respectively (not shown).

A.G. has been supported by the Chaire Développement Durable of the Ecole Polytechnique and by a grant from the Réseau de Recherche sur le Développement Soutenable (R2DS) of the Région Ile-de-France, while M.G. received partial support from U.S. Department of Energy grant DE-FG02-07ER64439.

* Electronic address: andreas.groth@ens.fr

[14] Note that phase synchronization implies phase locking, as well as frequency locking.