PriMera Scientific Engineering
Volume 8 Issue 2 February 2026
DOI: 10.56831/PSEN-08-248
ISSN: 2834-2550

PriMera Scientific Publications

Application of PCA in the Frequency Domain using the Covariance Spectrum

Type: Research Article
Received: December 22, 2025
Published: February 03, 2026

Citation:
Rune Brincker., et al. "Application of PCA in the Frequency Domain using the Covariance Spectrum". PriMera Scientific Engineering 8.2 (2026): 04-16.

Rune Brincker[1]*, Sandro Amador[2], Manuel Aenlle Lopez[3], Guangli Du[4], Umberto Alibrandi[5] and Emmanouil Lydakis[6]

[1]*CEO of Brincker Monitoring ApS, former Professor of Structural Dynamics, Technical University of Denmark (DTU), Denmark*

[2]*Assistant Professor, Technical University of Denmark (DTU), Denmark*

[3]*Full Professor, University of Oviedo, Spain*

[4]*Associate Professor, Aalborg University, Denmark*

[5]*Research Fellow, National University of Singapore*

[6]*Lecturer, Technical University of Denmark*

***Corresponding Author:** Rune Brincker, CEO of Brincker Monitoring ApS, former Professor of Structural Dynamics, Technical University of Denmark (DTU), Denmark.

Abstract

In this paper we will consider a re-interpretation of principal component analysis (PCA) for the case of time series of correlated data where the principal components are partly covered in noise, so that only in a part of the considered frequency band, the principal components are visible. In this case it might be useful to consider a representation of the spectral density matrix that is a real and one-sided function of frequency, in this paper denoted the covariance spectrum, which is directly representing the covariance matrix as a function of frequency. This means that if a principal component is dominating in a narrow frequency band, it might not be visible on the time domain, but be visible in the narrow band. The covariance spectrum can then be added in the considered band to represent the covariance matrix of the principal component, and classical PCA can be used to illustrate the properties of the principal component. Basic theory is introduced, and the principle is illustrated on a case with two harmonics in white noise acting on an arbitrary mechanical system.

Introduction

In this paper we will consider the analysis of time series using Principal Component Analysis (PCA). Classical PCA is well known in statistics, see for instance in the well know text book by Jolliffe [1].

In much of PCA statistics it is assumed that the observations are stochastically independent. However, in his chapter 12 Jolliffe discusses the implications for PCA of non-independent data, and much of the chapter is concerned with PCA for time series data. This is the most common type of non-independent data, and also the kind of data that we will consider in the present paper.

In his chapter 12 Jolliffe consider aspects of PCA that arise only for such data, for example, principal components in the frequency domain. In his analysis he is pointing out that for time series data, dependence between the data is induced by their relative closeness in time, but he does not go further into the matter as we will do in this paper. He is mentioning a number of different applications where data correlation is considered, but is not considering correlation functions and spectral densities as we will do here.

Since Lolliffe published his book in 2002, some papers have been published on performing PCA in the frequency domain, and we will go shortly over some of them.

In [2] it is realized that PCA often does not achieve the most useful separation of physical and background noise components. They suggest to inspect physical spectral components in the frequency domain and amplify these by a weighting function to obtain better results. In [3] it is shown how to improve Full Waveform Inversion (FWI) in seismic imaging using PCA in the frequency domain, and it is claimed that this leads to significant speed up within the FWI procedure, especially at low frequencies. In [4] the difference between usual PCA and PCA in the frequency domain has been studied to illustrate proximity between the two PCAs. In [5], like in [3] the idea is to improve full waveform inversion (FWI) using PCA in the frequency domain, which leads to significant speed up, especially at low frequencies. In [6] a spectral domain method is developed for multivariate stationary time series that transforms the observed series into several groups of lower-dimensional multivariate subseries using PCA in the frequency domain. The methods are illustrated on wind speed forecasting. In [7] the authors develop the theoretical background necessary to implement PCA at high frequency making it possible for PCA to becomes feasible over short windows of observation. The tools are illustrated on data from the stock market.

However, the background of this paper is mainly to explain why the identification technique for operational modal analysis (OMA) developed by Brincker et al in 2000, see [8], has been so popular even though the initial theoretical background was not well described. One reason is that it was present as one of the two main techniques in the ARTeMIS software, see [10], that was one of the early commercial products in OMA, and as such became well-known. Another reason is that it works extremely well in difficult cases where data might be influenced by non-linearities and noise, that might disturb more classical OMA techniques that are based on linear models.

Another purpose of this paper is to make use of the contribution from Yi Liu et all from 2019, see [9], where it is shown that a more simple version of the classical spectral density function can be introduced where we only work with a one-sided and real-valued spectral density that is still exactly satisfying the Parseval theorem, and where each of the spectral density matrices at the different frequency lines can be considered as a covariance matrix, so that classical PCA can be applied directly.

For the further analysis, we are considering a general, but zero-mean and second order stationary vector response $\mathbf{y}(t)$ as a function of the time t. We can think about this response as generated by some kind of system that is loaded by some kind of forces $\boldsymbol{X}(t)$, but in general we will not assume anything about the system or the kind of forces that might be driving the system unless specifically mentioned in the following.

The idea is then to say, that since the spectral density is always well defined, but is also difficult to interpret since spectral densities are in general complex valued and 2 sided functions defined from minus to plus infinity, we simplify the spectral density to a one-sided and real-valued function denoted the covariance spectrum, so that the distribution over the positive frequency becomes an energy distribution. We do that using the Parseval theorem, because this theorem is already relating spectral density to the energy expressed by the covariance of the response.

As an example, in section 6 we consider two harmonics in broad banded white noise, and illustrate the advantages of using PCA in the frequency domain.

Spectral background

As it is well-known from the literature on time series and random responses, see for instance Newland [11], any zero-mean stationary response can be characterized by its correlation function matrix $\mathbf{R}_{\mathrm{y}}(\tau)$ defined as

$$\mathbf{R}_{\mathrm{y}}(\tau) = \mathrm{E}\left[\mathbf{y}(t)\mathbf{y}^{T}(t+\tau)\right] \qquad (1)$$

And the corresponding spectral density matrix $\mathbf{G}_{\mathrm{y}}(\omega)$ is then given by its Fourier transform

$$\mathbf{G}_{\mathrm{y}}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{R}_{\mathrm{y}}(\tau)e^{-i\omega\tau}d\tau \qquad (2)$$

Where ω is the angular frequency. From this equation it is easy to see that the spectral density at negative frequency is the complex conjugate of the corresponding spectral density at positive frequency

$$\mathbf{G}_{\mathrm{y}}(-\omega) = \mathbf{G}_{\mathrm{y}}^{*}(\omega) \qquad (3)$$

This means that adding the two spectral density matrices together we obtain the one-sided and real valued spectral density

$$\mathbf{C}_{\mathrm{y}}(\omega) = \mathbf{G}_{\mathrm{y}}(\omega) + \mathbf{G}_{\mathrm{y}}(-\omega) \qquad (4)$$

In Yi Liu et al, [9], the complete theory for the so defined one-sided and real valued spectral density is given for discrete time, the only new thing here is the interpretation of this function.

The key to understanding the importance of this more simplified function is the Parseval theorem that states that the initial value of the correlation function matrix is equal to the integral over definition interval of the spectral density matrix

$$\mathbf{R}_{\mathrm{y}}(0) = \int_{-\infty}^{\infty} \mathbf{G}_{\mathrm{y}}(\omega)d\omega \qquad (5)$$

Where the initial value matrix $\mathbf{R}_{\mathrm{y}}(0)$ is the covariance matrix

$$\mathbf{C}_{\mathrm{y}} = \mathrm{E}\left[\mathbf{y}(t)\mathbf{y}^{T}(t)\right] \qquad (6)$$

That is indeed a constant matrix because we have assumed the signal $\mathbf{y}(t)$ to be stationary. This means due to Eq. (4) and (5), that the one-sided and real valued spectral density $\mathbf{C}_{\mathrm{y}}(\omega)$ is actually a distribution of the energy of the system, so that the covariance matrix is given by the integral of the one sided function $\mathbf{C}_{\mathrm{y}}(\omega)$

$$\mathbf{C}_{\mathrm{y}} = \int_{0}^{\infty} \mathbf{C}_{\mathrm{y}}(\omega)d\omega \qquad (7)$$

This also means, that the one-sided spectral density $\mathbf{C}_{\mathrm{y}}(\omega)$, *w* that we will denote “the covariance spectrum” for each frequency represents the contribution to the covariance matrix $\mathbf{C}_{\mathrm{y}}$ at that frequency.

Principal component analysis (PCA)

The idea of the PCA is to reduce the number of components in the response $\mathbf{y}(t)$ with N components by projecting the response onto a subspace of principal vectors

$$\mathbf{A} = \left[\mathbf{a}_1, \mathbf{a}_2, \cdots\right] \qquad (8)$$

That are unit length and orthogonal, creating the *N* x *N* orthogonal matrix **A** so that the inner product is equal to the identity matrix

$$\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T = \mathbf{I} \qquad (9)$$

The projected response $\hat{\mathbf{q}}(t)$ also denoted the vector of principal components is then created as

$$\hat{\mathbf{q}}^T(t) = \mathbf{y}^T(t)\mathbf{A} \qquad (10)$$

That is constituting the basic PCA equation. However, it is important to be aware, that because $\hat{\mathbf{q}}(t)$ is an estimate depending on the choice of the principal vectors, we have indicated that with a hat. Taking the transpose of Eq. (10), and using Eq. (9) we obtain the corresponding equation for the synthesized signal

$$\hat{\mathbf{y}}(t) = \mathbf{A}\hat{\mathbf{q}}(t) \qquad (11)$$

Reduction of the number of components can then be made excluding the components with small amplitude (noise components), until the remaining components are believed to be of physical importance.

It is normal when writing these equations to stack the transposed sampled signals of the time series in the response matrices

$$\mathbf{Y}^T = \begin{bmatrix} \mathbf{y}^T(\Delta t) \\ \mathbf{y}^T(2\Delta t) \\ \vdots \end{bmatrix}; \quad \mathbf{Q}^T = \begin{bmatrix} \mathbf{q}^T(\Delta t) \\ \mathbf{q}^T(2\Delta t) \\ \vdots \end{bmatrix} \qquad (12)$$

Where Δt is the sampling time step. This provides the basic PCA equation corresponding to Eq (10) in its classical PCA matrix form

$$\mathbf{Q}^T = \mathbf{Y}^T\mathbf{A} \qquad (13)$$

And the corresponding equation providing the estimate $\hat{\mathbf{Y}}$ of the data matrix of the synthesized signals

$$\hat{\mathbf{Y}} = \mathbf{A}\mathbf{Q} \qquad (14)$$

However, let us go back to continuous time and consider Eq. (11) where on the left we have the synthesized data expressed as a classical time series, and to the right we have the PCA model expressing the synthesized time series $\hat{\mathbf{y}}(t)$ as a linear combination of the principal vectors $\mathbf{a}_i$ and the principal components $q_i(t)$

$$\hat{\mathbf{y}}(t) = \mathbf{A}\mathbf{q}(t) = \mathbf{a}_1 q_1(t) + \mathbf{a}_2 q_2(t) + \cdots \qquad (15)$$

Since the principal vectors are here considered constant (describing the physical behavior of the system), the time dependence can only be described by the principal components $q_i(t)$.

A natural way to look at Eq. (15) is that we have decoupled the response into a linear combination of constant principal vectors and time-dependent principal components, that is very common in classical dynamics, see for instance Newland, [11]. In the classical dynamics the principal vectors are the so-called "mode shapes" and the principal components are the corresponding "modal components". However, in random vibrations, the principle of linear superposition does not come easy, and built on an assumption of stochastically independent modal coordinates, see [8], section 6.3.

Finally, we should estimate the principal vectors in such a way that we minimize the error between the original and the synthesized signal

$$\mathbf{e}(t) = \mathbf{y}(t) - \hat{\mathbf{y}}(t) \qquad (16)$$

Decomposition in time domain

In the classical PCA the principal vectors are found as the eigen-vectors of the covariance matrix see Eq. (6). Using the synthesized data from Eq. (11) the covariance matrix becomes

$$\mathbf{C}_{\mathbf{y}} = \mathrm{E}\left[\mathbf{A}\mathbf{q}(t)\mathbf{q}^T(t)\mathbf{A}^T\right] = \mathbf{A}\mathbf{C}_q\mathbf{A}^T \quad (17)$$

Multiplying from the right by the principal vector matrix **A** we get

$$\mathbf{C}_{\mathbf{y}}\mathbf{A} = \mathbf{A}\mathbf{C}_q \quad (18)$$

Which we recognize as an Eigen Value Decomposition (EVD). Since the covariance matrix is real and positive definite, the eigen vector matrix **V** = **A** is real, it fulfils Eq. (9) and all eigen values d_i are positive and gathered in the diagonal of the eigen value matrix $\mathbf{D} = \mathbf{C}_q$ which means that we have forced the principal components to be independent. The reason for introducing the matrices **V** and **D** is only that we want to distinguish between the principal vectors in the matrix **A** and the corresponding principal components in the matrix $\mathbf{C}_q$ and the similar quantities **V** and **D** (eigen vectors and eigen values) estimated by the eigen value decomposition.

Further, we also see from Eq. (17) that the right-hand side is a Singular Value Decomposition (SVD) of the real quadratic matrix $\mathbf{C}_{\hat{\mathbf{y}}}$, so that the general SVD

$$\mathbf{C}_{\mathbf{y}} = \mathbf{U}_1\mathbf{S}\mathbf{U}_2^H \quad (19)$$

However, let us reduce to the case where for the singular matrices $\mathbf{U_1} = \mathbf{U_2} = \mathbf{U} = \mathbf{A}$, and the singular value matrix $\mathbf{S} = \mathbf{C}_q$ as shown in Eq. (18). In order to secure that this also means that the principal components estimated by the EVD are the same as the principal components estimated by the SVD, we just have to sort the principal values like the SVD in descending order from the upper left to the right of the the eigen value matrix **D**. Again, the reason for introducing the matrices **U** and **S** is only that we want to distinguish between the principal vectors in the matrix **A** and the corresponding principal components in the matrix $\mathbf{C}_q$; likewise for the similar quantities **U** and **S** (singular vectors and singular values) estimated by the singular value decomposition.

Decomposition in frequency domain

We will now explain the advantages involved in having defined the covariance spectrum, so that the above-mentioned decomposition can also be carried out in the frequency domain.

In section 1, we have concluded that the covariance spectrum for each frequency represents the contribution to the covariance matrix at that frequency. This also means that the covariance spectrum can be decomposed at any frequency just like the covariance matrix given by Eq. (17).

$$\mathbf{C}_{\mathbf{y}}(\omega) = \mathbf{A}^T\mathbf{C}_q(\omega)\mathbf{A} \quad (20)$$

Where the matrices **A** and $\mathbf{C}_q(\omega)$ just as before can be found by EVD or SVD. As a result, the principal components in the diagonal matrices $\mathbf{C}_q(\omega)$ can be plotted as a function of frequency. It should be mentioned here that all terms in this equation are functions of the frequency, so also the principal vector matrices are functions of frequency $\mathbf{A} = \mathbf{A}(\omega)$.

If we have a dominant principal component at a certain frequency, let us say ω_1, then we might like to see how the spectral density $C_{\omega 1}(\omega)$ for that dominating component is looking around and then we might use the projection

$$C_{\omega 1}(\omega) = \mathbf{a}_{\omega 1}^T\mathbf{C}_{\mathbf{y}}(\omega)\mathbf{a}_{\omega 1} \quad (21)$$

Where the projection vector $\mathbf{a}_{\omega 1}$ is taken as the most dominant principal vector, that is the first column of the matrix **A** estimated at the frequency ω_1, so that using Matlab index notation

$$\mathbf{a}_{\omega 1} = \mathbf{A}(:,1,\omega_1) \quad (22)$$

Then, when the dominating component has been identified in the frequency domain, we can take it back to time domain to illustrate its correlation function properties by the inverse Fourier transform, see Eq. (2). We will be illustrating the principle in the following example.

Example of two harmonics in white noise

The idea is now to consider harmonics in heavy broad-banded noise. If we just go ahead and use the PCA in the classical way, we get the noise from the whole frequency band into the analysis, and as a result we might not be able to see or identify the two harmonics.

However, looking at it in the frequency domain, we might still be able see the two harmonics, and using the covariance spectrum, we can pick the covariance matrices of the two harmonics directly from the covariance spectrum, and perform the PCA on these covariance matrices. This will give us a better estimate as we shall see in the following where we will compare classical PCA in time domain, with PCA in the frequency domain.

We are considering a mechanical system, where the geometry and physical properties of the structure can be considered as arbitrary.

Simulation approach

We will use a sampling rate of 10 Hz giving us a Nyquist frequency of 5 Hertz and place the two harmonics close to the middle of the frequency band, and so that the physical signal matrix without noise is given by

$$\mathbf{Y}_0 = \left[\mathbf{y}_0(t_n)\right] = \left[\mathbf{a}_1 \cos(2\pi f_1 t_n) + \mathbf{a}_2 \cos(2\pi f_2 t_n)\right] \quad (23)$$

Where $f_1 = 2.50$ Hz and $f_2 = 2.75$, and t_n is the discrete time. The principal vectors $\mathbf{a}_1$, $\mathbf{a}_2$ that should be considered as the Operating Deflection Shapes (ODS) of the two harmonics are chosen as one constant value and one antisymmetric linear function respectively, in such a way that we fulfill Eq. (9). The number of signals in the data vector is 10, and number of data points in the time series is taken as a reasonable radix 2 number to be $N_p = 8192$. For the Fast Fourier Transform (FFT) analysis we use a data segment size of 512, and a Hannig window with 50 % overlap. The data matrix is then created as

$$\mathbf{Y} = \mathbf{Y}_0 + \mathbf{Y}_n \quad (24)$$

Where $\mathbf{Y}_n$ is the noise matrix consisting of uncorrelated white noise with the standard deviation σ_n. We will now consider the three cases:

1. Moderate heavy noise $\sigma_n = 1/3$ with where both PCA techniques works fine.
2. Medium heavy noise with $\sigma_n = 1$ where the classical PCA becomes problematic.
3. Heavy noise with $\sigma_n = 3$ where the classical PCA becomes nearly useless but the PCA based on the covariance spectrum works fine.

Case 1, moderate heavy noise

A short clip of the simulated data for this case where the standard deviation of the noise is $\sigma_n = 1/3$ is shown in Figure 1. We see from the top plot, that the physical signal has clear indications of its periodic behavior, and that the final signal has a somewhat larger amplitude due to the noise.

Figure 2 shows the plot that is the most important innovation in this paper - the plot of the eigenvalues of the covariance spectrum. This plot illustrates the value of the covariance spectrum, because it shows where important narrow band signals are present.

Since the data have 10 signals, the covariance spectral matrices are all 10 x 10, and thus, we have 10 eigenvalues at each frequency line in Figure 2. In the most of the frequency band we just see the constant levels of the eigenvalues due to the white noise, but the narrow-banded signals from the two harmonics are penetrating the noise floor in the middle of the frequency band, and both harmonics become visible, with a local signal-to-noise ratio in the frequency band of about 27 dB, see Figure 2.

For the classical PCA case the empirical covariance matrix is calculated as

$$\mathbf{C}_y = \mathbf{Y}\mathbf{Y}^T / (N_p - 1) \qquad (25)$$

Where N_p is the number of data points, and then the eigenvalues d_i are extracted using the eigenvalue decomposition in Eq. (18)

$$\mathbf{C}_y \mathbf{V} = \mathbf{V}\mathbf{D} \qquad (26)$$

Where **V**, **A** are the eigenvector matrix and eigenvalue matrices respectively, and we can then plot $\sqrt{d_i}$ in descending order as shown in the left top plot of Figure 3.

This type of plot is the classical key plot that makes it possible for us to choose the number of principal vectors from matrix $\mathbf{V}_c$. In this case we have ourselves chosen the two principal components, so we know that we have to pick two principal vectors, and we also clearly see in the plot, that the first two eigenvalues are significantly larger than the rest, indicating that two principal signals are in fact present. Thus, in order to estimate the principal vector matrix, we chose the two first vectors in the eigenvector matrix

$$\hat{\mathbf{A}}_c = \mathbf{V}_c(:,1:2) \qquad (27)$$

If we want to use the frequency domain PCA, we can follow two approaches. One is to pick the two covariance matrices $\mathbf{C}_{f1} = \mathbf{C}_y(f_1)$, $\mathbf{C}_{f2} = \mathbf{C}_y(f_2)$ at the two frequencies $f_1 = 2.50$ Hz and $f_2 = 2.75$ at the top of the peaks shown in Figure 2, and then perform eigenvalue decomposition on the two covariance matrices individually, see right top plot and the left bottom plot of Figure 3 that is showing the 10 eigen values from the two covariance matrices $\mathbf{C}_{f1}$ and $\mathbf{C}_{f2}$. To estimate the principal vector matrix we pick the first vector from the two corresponding eigenvector matrices $\mathbf{V}_{f1}$, $\mathbf{V}_{f2}$

$$\hat{\mathbf{A}}_{f12} = \left[\mathbf{V}_{f1}(:,1), \mathbf{V}_{f2}(:,1)\right] \qquad (28)$$

If we do it this way, we have to realize that the matrix $\hat{\mathbf{A}}_{f12}$ is not orthogonal, because the two vectors $\mathbf{V}_{f1}(:,1)$ and $\mathbf{V}_{f2}(:,1)$ are picked from two difference covariance matrices.

However, In order to secure orthogonality, we can follow a second approach where we add the two covariance matrices together $\mathbf{C}_{f1+2} = \mathbf{C}_y(f_1) + \mathbf{C}_y(f_2)$, and then perform the eigenvalue decomposition on the resulting single covariance matrix, see the bottom right plot in Figure 3. In this case, we do the PCA taking the first two vectors in the corresponding eigenvector matrix $\mathbf{V}_{f1+2}$ to estimate the principal vector matrix

$$\hat{\mathbf{A}}_{f1+2} = \mathbf{V}_{f1+2}(:,1:2) \qquad (29)$$

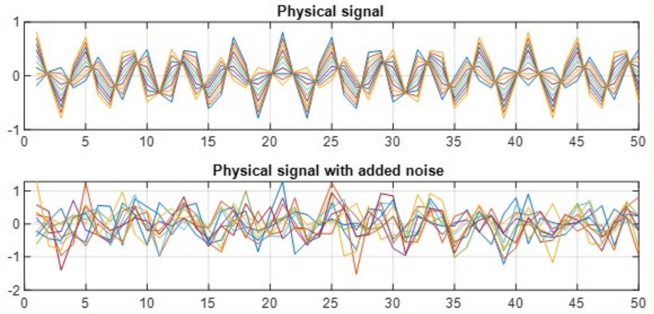

Figure 1: The first 50 realizations of a simulated time series for case 1, moderate heavy noise with noise standard deviation $\sigma_n = 1/3$.

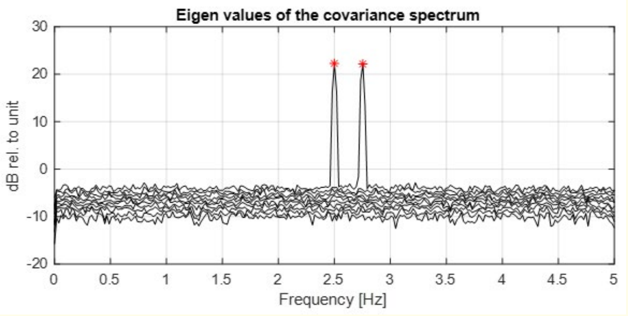

Figure 2: The eigen values $\mathbf{C}_q(\omega)$ according to Eq. (20) for case 1, moderate heavy noise with noise standard deviation $\sigma_n = 1/3$. We see the two harmonics indicated by the red asterisks with a signal-to-noise ratio in the frequency domain of approximately 27 dB.

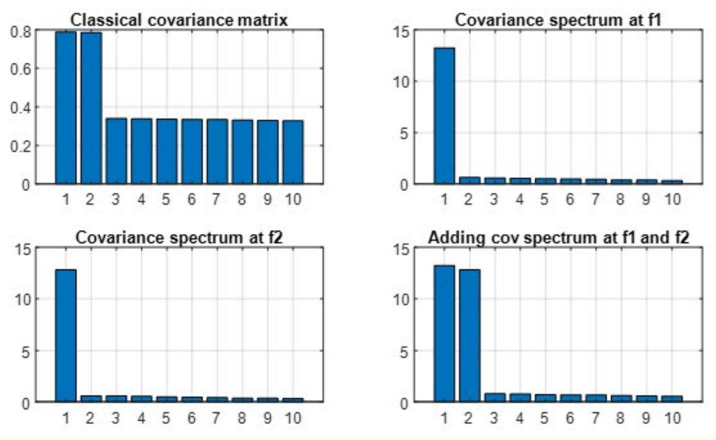

Figure 3: The eigen value plots for case 1, moderate noise with noise standard deviation $\sigma_n = 1/3$. We see that the classical PCA in the time domain (top to the left) works fine together with the PCA based on the covariance spectrum.

Case 2, medium heavy noise

Results of this case where the standard deviation of the noise is now $\sigma_n = 1$ are given in Figures 4 and 5.

Figure 4 shows that the noise has raised with about 10 dB so that the signal-to-noise ratio in the frequency domain is now about 17 dB.

Figure 5 shows that the classical PCA is now facing problems because the drop-down from eigen values 1 and 2 to the remaining eigenvalues is now so small, that it might not seem to be a meaningful choice to pick only two principal vectors. However, since we know that two is in fact the right number, we can again estimate the principal vector matrix as given by Eq. (27).

The rest of Figure 5 shows us that the frequency domain PCA is doing well, clearly pointing out the right principal vectors and we can estimate the principal vector matrix as given by Eq. (28) and (29).

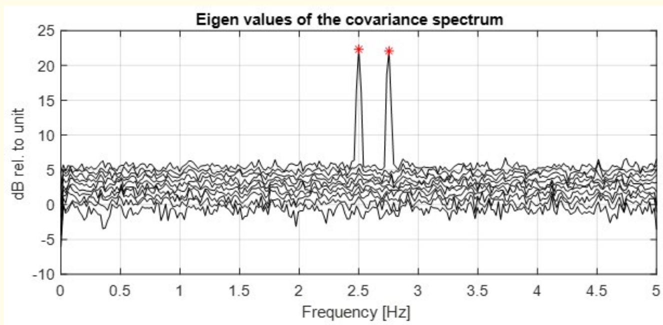

Figure 4: The eigen values $\mathbf{C}_q(\omega)$ according to Eq. (20) for for case 2, medium heavy noise with noise standard deviation $\sigma_n = 1$. We see the two harmonics indicated by the red asterisks with a signal-to-noise ratio in the frequency domain of approximately 17 dB.

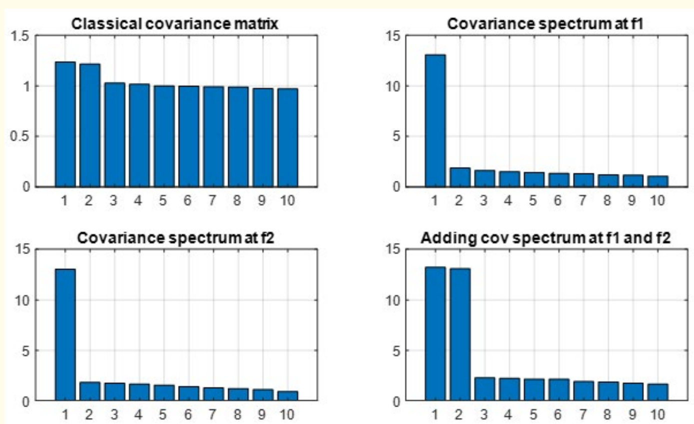

Figure 5: The eigen value plots for case 2, medium heavy noise with noise standard deviation $\sigma_n = 1$. We see that the classical PCA in the time domain (top to the left) is now facing problems, whereas the covariance spectrum based PCA is doing well.

Case 3, heavy noise

We are now at the case where the standard deviation of the noise is $\sigma_n = 3$ and the results are given in Figures 6 and 7.

Figure 6 shows that the noise has raised again with about 10 dB so that the signal-to-noise ratio in the frequency domain is now about 7 dB.

Figure 7 shows that the classical PCA is now useless because the drop-down from eigen values 1 and 2 to the remaining eigenvalues is so small, that it is impossible to argue to pick only two principal vectors. However, since we know that two is in fact the right number, we can again estimate the principal vector matrix as given by Eq. (27).

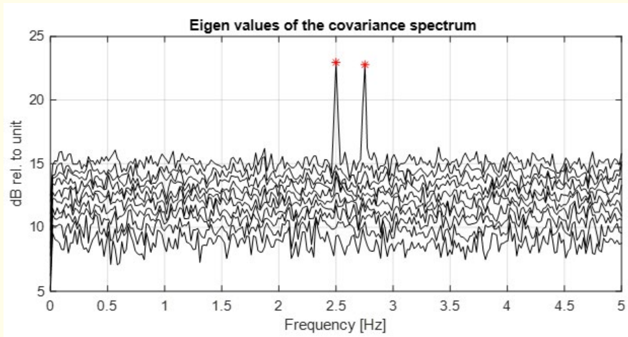

Figure 6: The eigen values $\mathbf{C}_q(\omega)$ according to Eq. (20) for case 3, heavy noise with noise standard deviation $\sigma_n = 3$. We see the two harmonics indicated by the red asterisks with a signal-to-noise ratio in the frequency domain of approximately 7 dB.

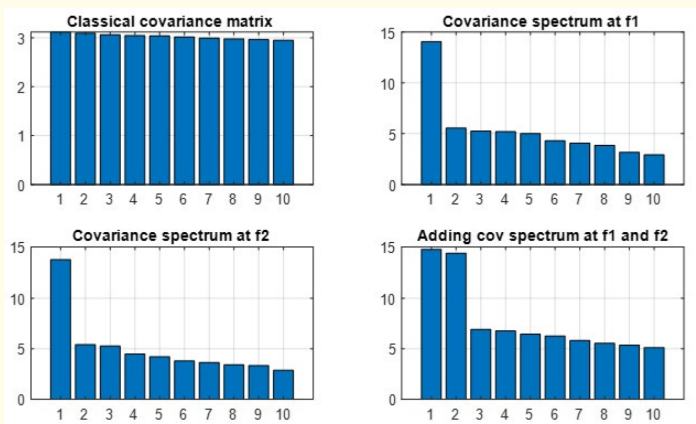

Figure 7: The eigen value plots for case 3, heavy noise with standard deviation $\sigma_n = 3$. We see that the classical PCA in the time domain (top to the left) is now useless, whereas the covariance spectrum based PCA is doing well.

The rest of Figure 4 shows us that the frequency domain PCA is doing well, clearly pointing out the right principal vectors and we can estimate the principal vector matrix as given by Eq. (28) and (29).

Principal components in the frequency domain

We will now further illustrate the frequency domain decomposition on the most difficult case 3 with heavy noise with standard deviation $\sigma_n = 3$.

We will estimate the two principal components in the frequency domain using the principles in Eq. (20) except that we will force the two estimated singular vectors to be orthogonal. This is addressed adding the covariance matrices from the two peaks of the two harmonics. The two principal components are then estimated together as the diagonal elements of the covariance matrices given by

$$\mathbf{C}_q(\omega) = \mathbf{A}_{f1+2}^T \mathbf{C}_y(\omega) \mathbf{A}_{f1+2} \quad (30)$$

Where the matrix $\hat{\mathbf{A}}_{f1+2}$ is given by Eq. (29), see the top plot of Figure 8, showing the estimate of the spectral densities of the two principal components. The bottom plot shows the corresponding correlation functions found by the inverse FFT.

As we can see from the bottom plot, we have a clear indication of the strong noise by the initial "delta function" peak, and after that, we see oscillating behaviour with constant amplitude indicating the presence of the two harmonics.

If we read the initial peak value and take the square root to estimate the noise level, we obtain a standard deviation typically close to 3, that over 10 simulations give us a mean value of 3.08 corresponding to an overestimation of the noise of about 3 %.

It is well known that the correlation function of a harmonic signal given by $a\cos(2\pi f t_n)$ as shown in Eq. (23) will be the same harmonic but with an initial value $a^2/2$ of equal to the variance of the estimated principal component. It means that we can estimate the amplitude of the harmonics reading the constant amplitude value from the bottom plot of Figure 8. Reading these values over 10 simulations give us a mean value of the amplitude of the harmonics of 1.08 corresponding to an overestimation of the amplitude of about 8 %.

These results are not bad taking into account the difficult case of the heavy noise.

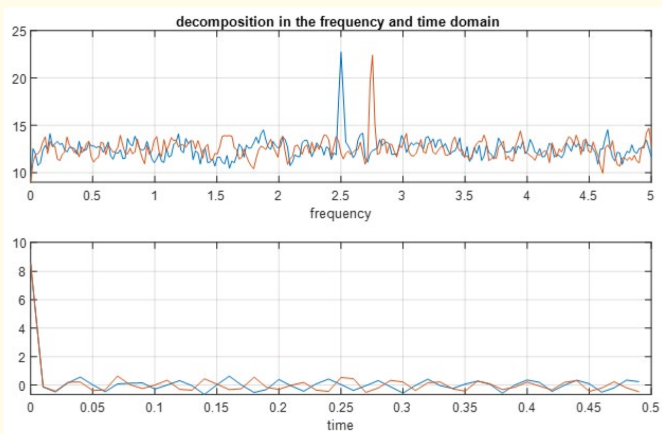

Figure 8: Top plot: Decomposition in frequency domain by plotting the two diagonal elements of $\mathbf{C}_q(\omega)$ given by Eq. (30) providing the spectral densities of the principal components, bottom plot: corresponding correlation functions of the principal components.

Example discussion

As illustrated in the preceding sections, the classical PCA works well in cases of low and moderate noise, but when the signal-to-noise ratio of the narrow band signals in the frequency domain gets below 20 dB, problems arise. It is no longer simple to detect the number of principal vectors to pick from the eigenvalue decomposition. Further, when the signal-to-noise ratio in the frequency domain falls down below 10 dB, it might be useless to use classical PCA simply because we can no longer use the eigenvalue plot to decide how many principal values we should use.

On the other hand, the PCA in the frequency domain, makes it much easier to pick the right number of principal vectors from the covariance matrices, that in this approach is picked directly at the peak of the narrow band signals.

As indicated above, the main information that is provided by the eigenvalue plot, is the drop-down from the last principal eigenvalue to the next one. Let us say, that we consider the square root of the last principal eigenvalue $\sqrt{d_n}$, it is then natural to consider the ratio ***R*** between this eigenvalue and next eigenvalue

$$R = \sqrt{d_n / d_{n+1}} \qquad (31)$$

We will consider this as the eigenvalue signal-to-noise ratio. Some typical values for this ratio are given in Table 1. From the results of Table 1 we see big difference between the classical PCA and the PCA based on the covariance matrix. The latter has much better signal-to-noise ratios in all cases.

Further, we see a smaller but significant improvement from the frequency domain PCA where the covariance is added from different peaks compared to the PCA where each peak provides a covariance matrix and one principal vector. The first approach however, should be considered in cases with closely spaced narrow-banded signals where the covariance matrices at the different peaks might interact.

Further, we should be aware, that providing principal vectors from individual points of the covariance spectrum will in general lead to a principal vector matrix given by Eq. (28) that is not orthogonal.

Case	***Classical PCA***	***Cov at f1***	***Cov at f2***	***Cov add f1, f2***
Case 1, moderate noise	2,27	22,22	20,16	15,65
Case 2, medium noise	1,18	7,21	6,82	5,42
Case 3, heavy noise	1,01	2,62	2,47	2,00

Table 1: Typical values of the signal-to-noise ratio for the eigenvalues defined in Eq. (31).

Another point interesting to investigate would be the noise present in the synthesized signals given by Eq. (16), in order to evaluate the difference of the different PCA approaches concerning their ability to perform noise reduction. The procedure is simple, first we use Eq. (13) to estimate the principal components, then we create the synthesized signals with reduced noise using Eq. (14) or (15), and then - since we have simulated the data and know the physical signals - calculate the noise on the synthesized signals according to Eq. (16).

It is important in many cases to perform noise reduction, and noise identification, but this is not the focus of this paper. However, in the considered cases, we only see small changes of the standard deviation of the noise on the synthesized signals which are all lying around $0.45 \times \sigma_n$, which means that the PCA have reduced the noise to less than the half.

Only for the classical PCA in case 3 with heavy noise, the noise on the synthesized signals is slightly higher than $0.45 \times \sigma_n$, and one might wonder why it works so well when we cannot see anything from the eigenvalue plots. The reason is that we have in fact chosen the two first eigenvectors in the eigenvector matrix, and even though these two vectors are not very good estimates of the original principal vectors that were chosen for simulation, they still define a reasonable basis for the subspace of the physical signals.

Finally, we have relatively good results performing the decomposition in the frequency domain and estimating the noise level and the amplitudes of the two harmonics.

Conclusions

We have combined the classical and well-known PCA where the covariance is estimated in the time domain with a known real-valued spectral density that we have denoted the covariance spectrum, and where the covariance matrices can be picked at the frequency lines where the covariance spectrum has large components that are believed to be important.

We have illustrated this new frequency domain PCA on an example with two harmonics in heavy noise, and we can conclude the following:

- One method for the PCA based on the covariance spectrum is based on using only one covariance matrix from each peak and then performing eigenvalue decomposition resulting in only one principal vector from each peak.
- Another method is based on picking one covariance matrix from a number of peaks, then add the covariance matrices together, and then performing eigenvalue decomposition resulting in as many principal vectors as the peaks we have chosen.
- Both the so defined methods for PCA based on the covariance spectrum work significantly better than the classical PCA, because the eigenvalue plot provides a better information about how many singular vectors should be used.
- The method based on adding the peak values from two or more points seem to work quite well even in the case of heavy noise to estimate both the noise level and the amplitude of the two harmonics.

In general, it can be concluded that the PCA based on the covariance spectrum provides a clearer picture of how many principal vectors that should be chosen, and as such helps the user to perform a better PCA than using classical PCA in the time domain.

References

1. IT Jolliffe. Principal components analysis, Second Edition, Springer (2002).
2. Terrence D Lagerlund, Frank W Sharbrough and Neil E Busacker. "Use of principal component analysis in the frequency domain for mapping electroencephalographic activities: comparison with phase-encoded Fourier spectral analysis". Brain Topography 17.2 (2004): 73-84.
3. Kyusoon Kim and Hee-Seok Oh. "Principal Component Analysis in the Graph Frequency Domain". arXiv:2410.08422v1 [stat.ME] (2024).
4. Alain Boudou and Sylvie Viguier-Pla. "On proximity between PCA in the frequency domain and usual PCA". Statistics, A Journal of Theoretical and Applied Statistic. 40.5 (2004): 447-464.
5. Chuncheng Liu., et al. "Application of principal component analysis for frequency-domain full waveform inversion". SEG Las Vegas 2012 Annual Meeting (2012).
6. Raanju R Sundararajan. "Principal Component Analysis using Frequency Components of Multivariate Time Series". arXiv:2010.04515v1 [stat.ME] (2020).
7. Yacine Aït-Sahaliaa and Dacheng Xiub. "Principal Component Analysis of High-Frequency Data". Journal of the American Statistical Association 114.525 (2019): 287-303.
8. R Brincker, L Zhang and P Andersen. "Modal identification from ambient responses using frequency domain decomposition". Proc. of the International Modal Analysis Conference (IMAC), San Antonio, Texas, USA, (2000).
9. Yi Liu, R Brincker and J MacDonald. "Consistent real-valued and one-sided spectral density functions". Proc of 8th International Operational Modal Analysis Conference (IOMAC), Copenhagen (2019): 409-420.
10. The ARTeMIS software for operational modal analysis". https://www.svibs.com/
11. DE Newland. "An Introduction to Random Vibrations, Spectral & Wavelet Analysis". Third Edition, Dover Publications (2005).