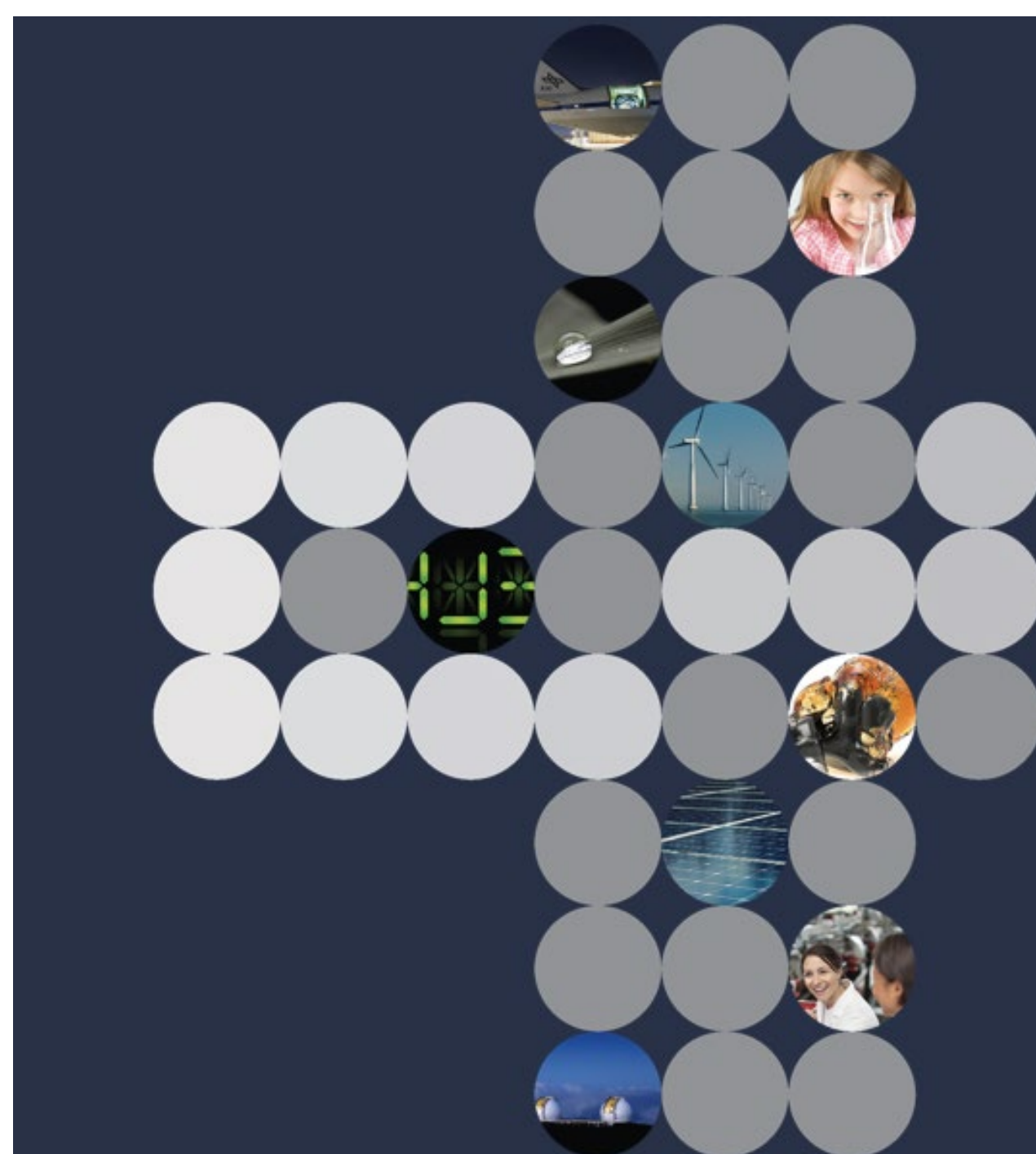


EE512 – Applied Biomedical Signal Processing

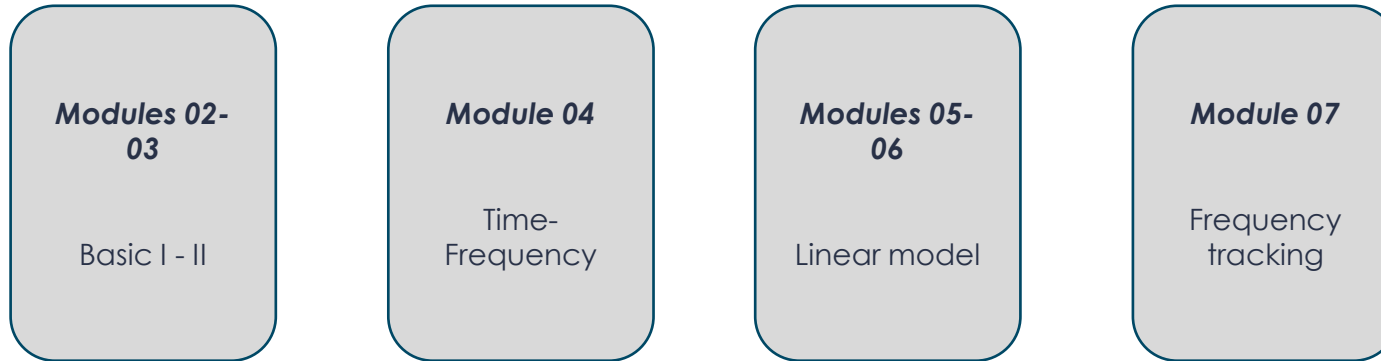
Singular Value Decomposition

Guillaume BONNIER
CSEM Signal Processing Group

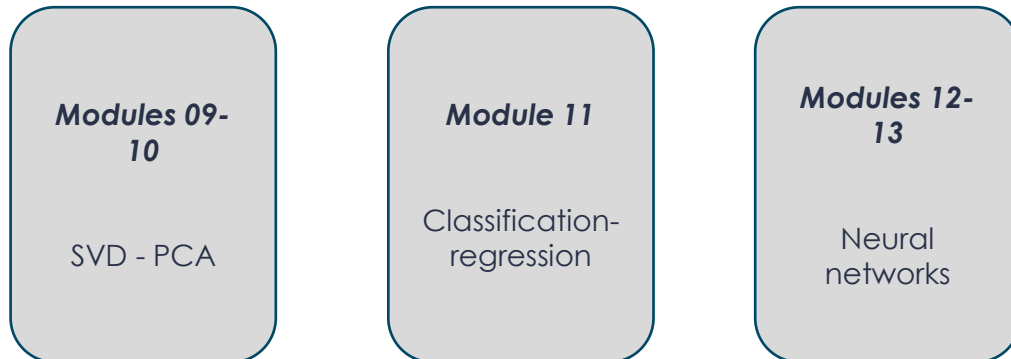


EE512 - Recap

- Modules 02-07: Signal processing



- Modules 08-13: Data driven approaches

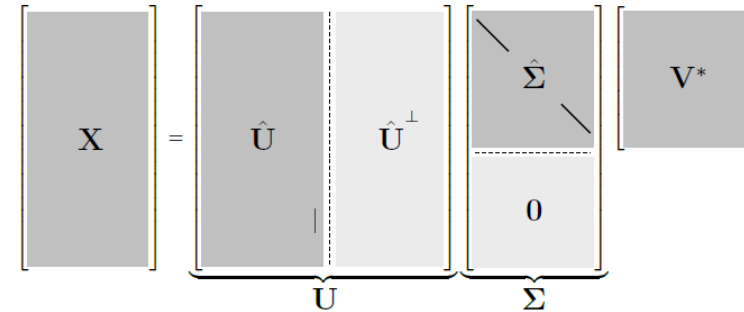


SVD module outline

Introduction

Singular Value Decomposition definition

Mathematical interpretation

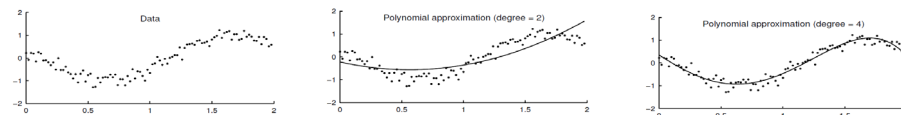
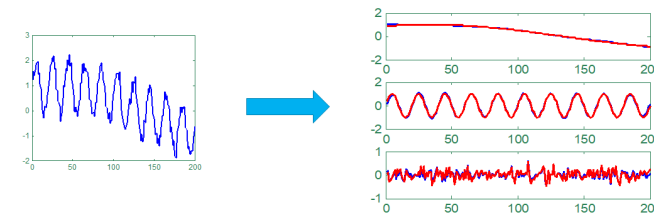
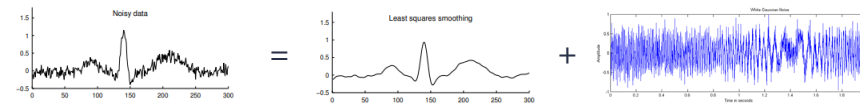


Practical application

Noise reduction

Singular spectrum analysis

Least square solution using SVD



Introduction (1): SVD overview

SVD is a **matrix factorization** technique which provides a systematic way to determine a **low-dimensional approximation** to high-dimensional data **in terms of dominant patterns**.

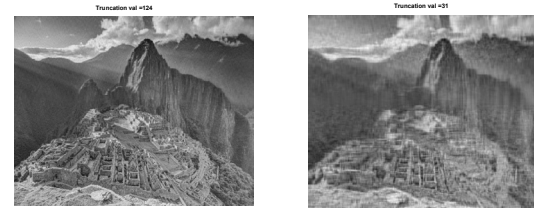
Mathematical **mapping** to a new coordinate system, similar to Fourier transform, **based on data**. “Tailored” solution to a specific problem.

Based on **simple** linear algebra, **interpretable and scalable**.

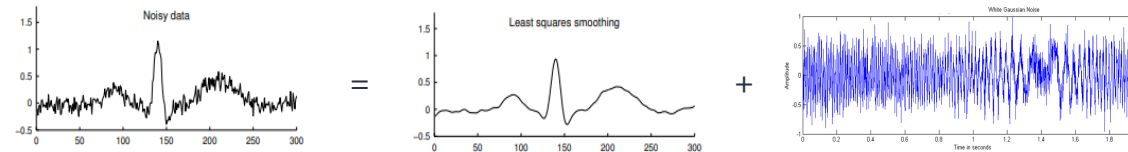
Introduction (2): SVD applied examples

One of the most widely used technique in **data processing**, **dimensionality reduction** and **high-dimensional statistics**.

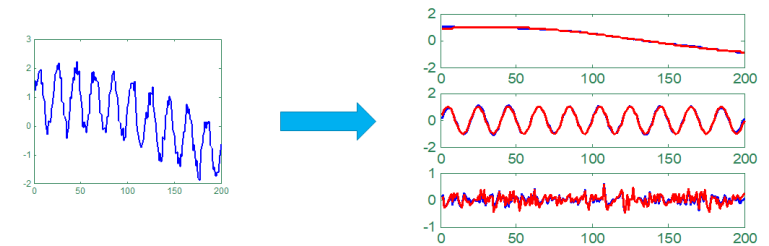
- Image compression



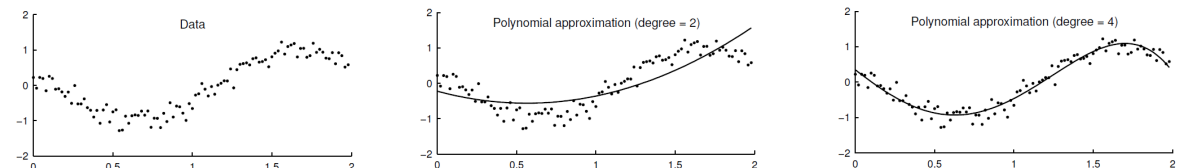
- Noise reduction



- Filterless signal decomposition



- Robust approach for **least square** estimation problems



Mathematical definition (1): SVD central theorem

SVD central theorem: **any** matrix \mathbf{X} of size $K \times M$ and of rank r can be decomposed as:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

with \mathbf{U} an **orthogonal** $K \times K$ matrix, \mathbf{V} an **orthogonal** $M \times M$ matrix, and $\mathbf{\Sigma}$ a specific $K \times M$ matrix.

$\mathbf{\Sigma}$ can be seen as the generalization of a **diagonal** matrix, with:

$$\Sigma_{ij} = 0, i \neq j \quad \text{and} \quad \Sigma_{ii} = \sigma_i > 0, i = 1, 2, \dots, r$$



`U, S, V = numpy.linalg.svd(X)`

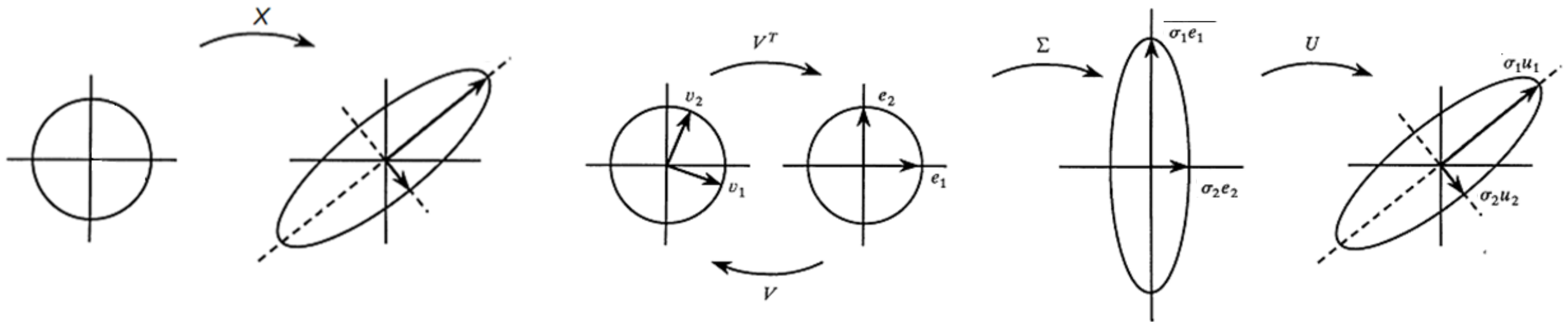


`[U, S, V] = svd(X);`

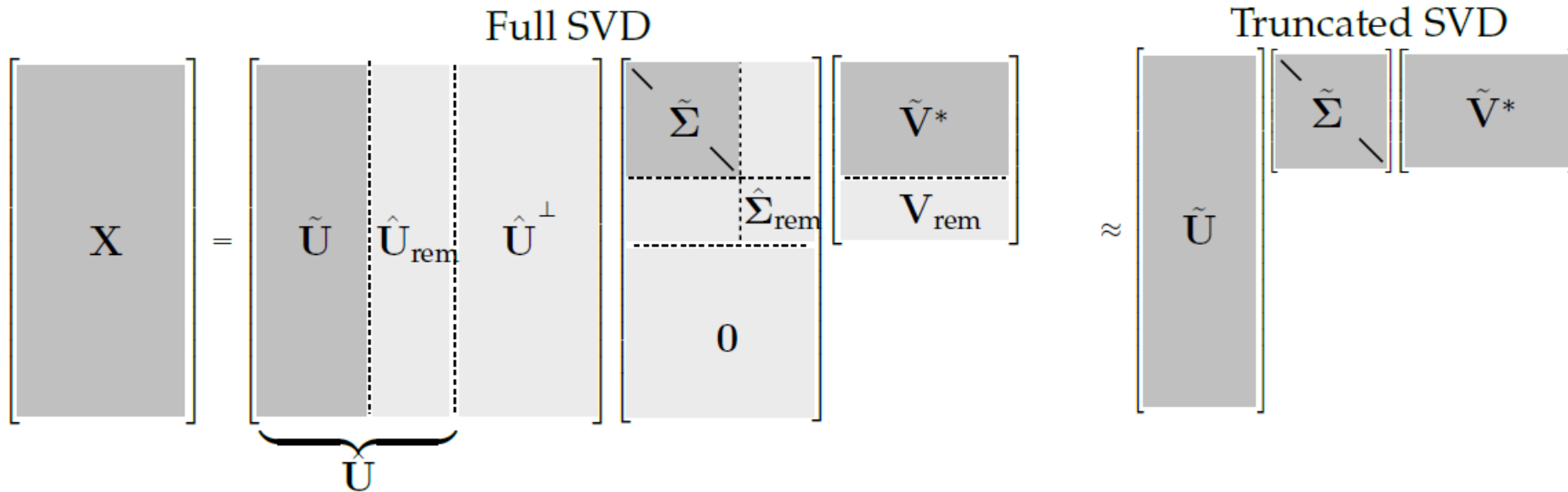
Mathematical definition (2): Geometrical interpretation

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

- \mathbf{U} , and \mathbf{V} are unitary matrices **rotating** vectors (preserve angle and length).
- $\mathbf{\Sigma}$ is rectangular diagonal matrix **scaling** vectors.



Mathematical definition (3): Low rank approximation



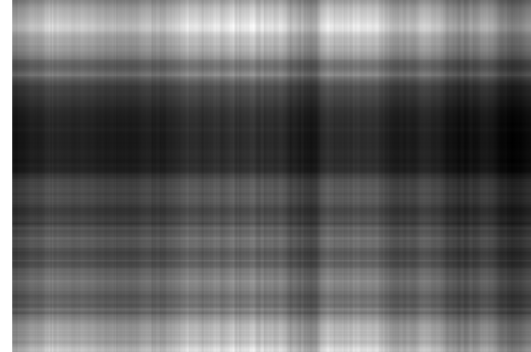
- The **SVD provides** an **optimal low-rank p** approximation **to the matrix X** by keeping the leading p singular values and vectors and discarding the rest.

$$\tilde{X} = \sum_{i=1}^p \sigma_i u_i v_i^T$$

SVD application: Data compression



r=1 (storage=0.08%)



r=20 (storage=1.67%)



r=100 (storage=8.34%)

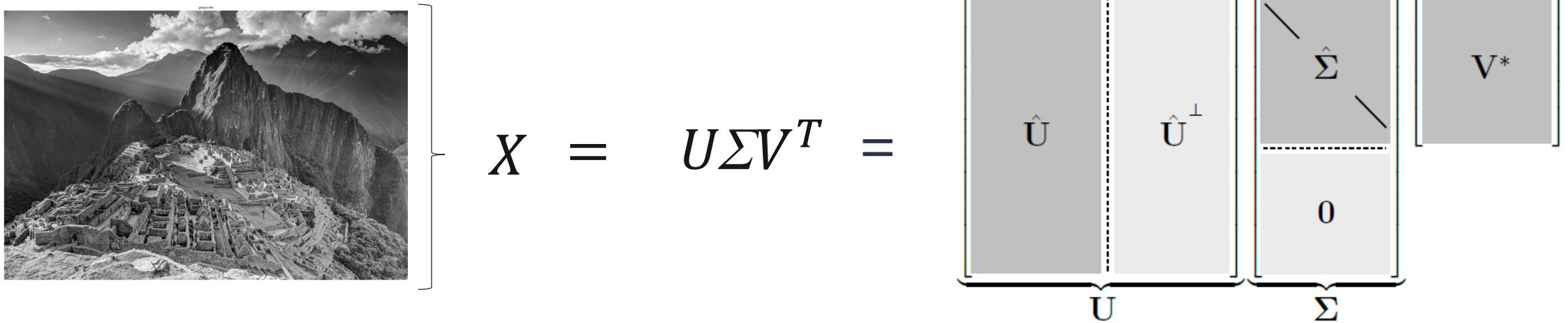


r=200 (storage=16.68%)



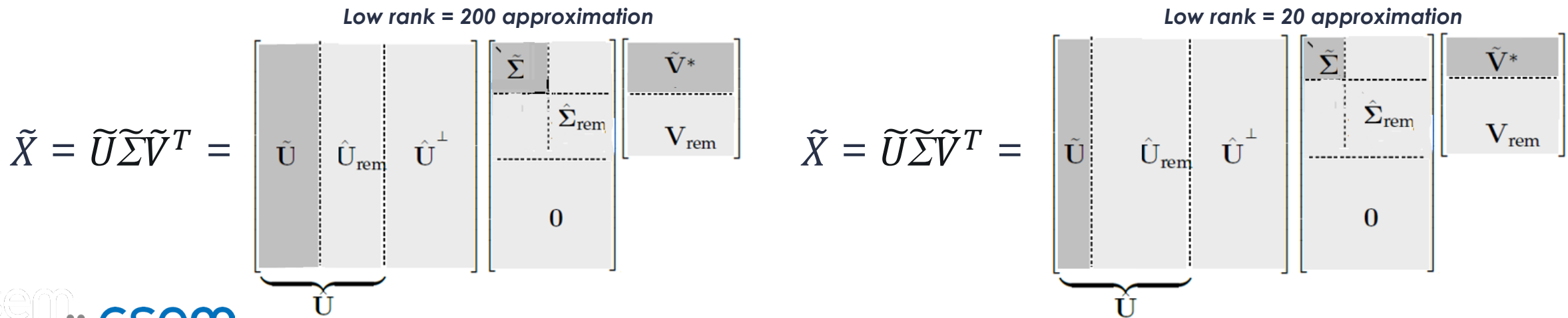
SVD application: Data compression

- Decompose the image using SVD:



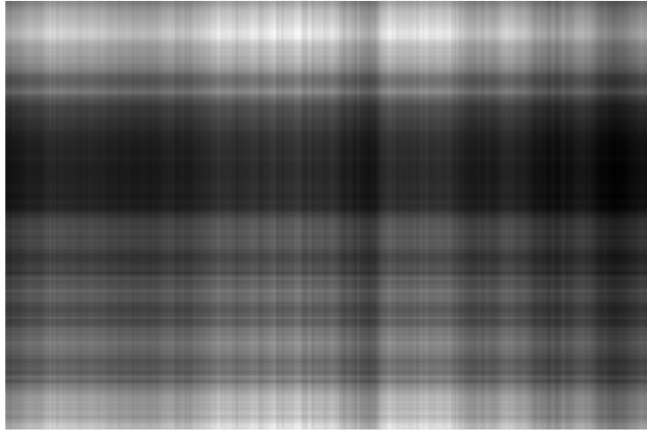
13

- Compute low-rank approximations of X : $\tilde{X} = \tilde{U}\tilde{\Sigma}\tilde{V}^T$



SVD application: Data compression

$r=1$ (storage=0.08%)



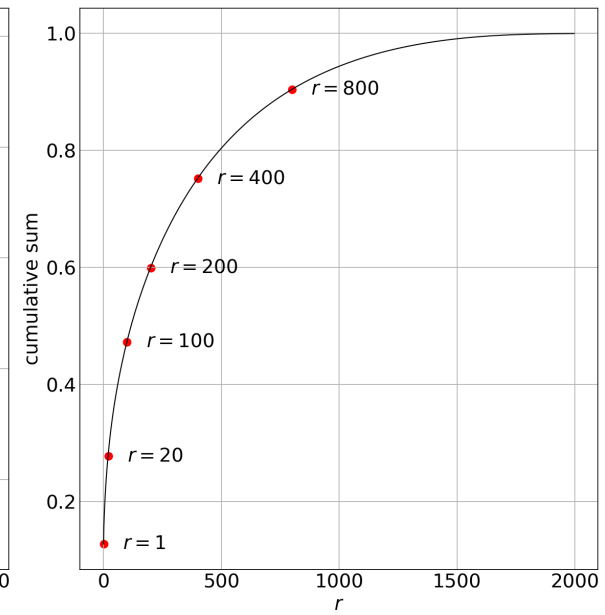
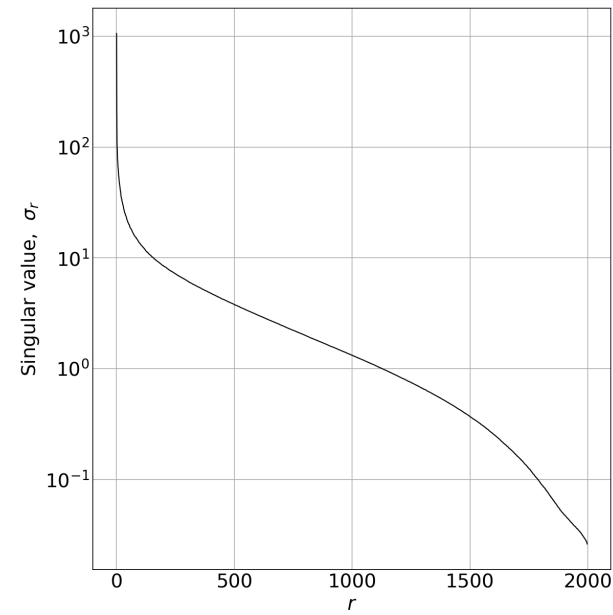
$r=20$ (storage=1.67%)



$r=100$ (storage=8.34%)



$r=200$ (storage=16.68%)

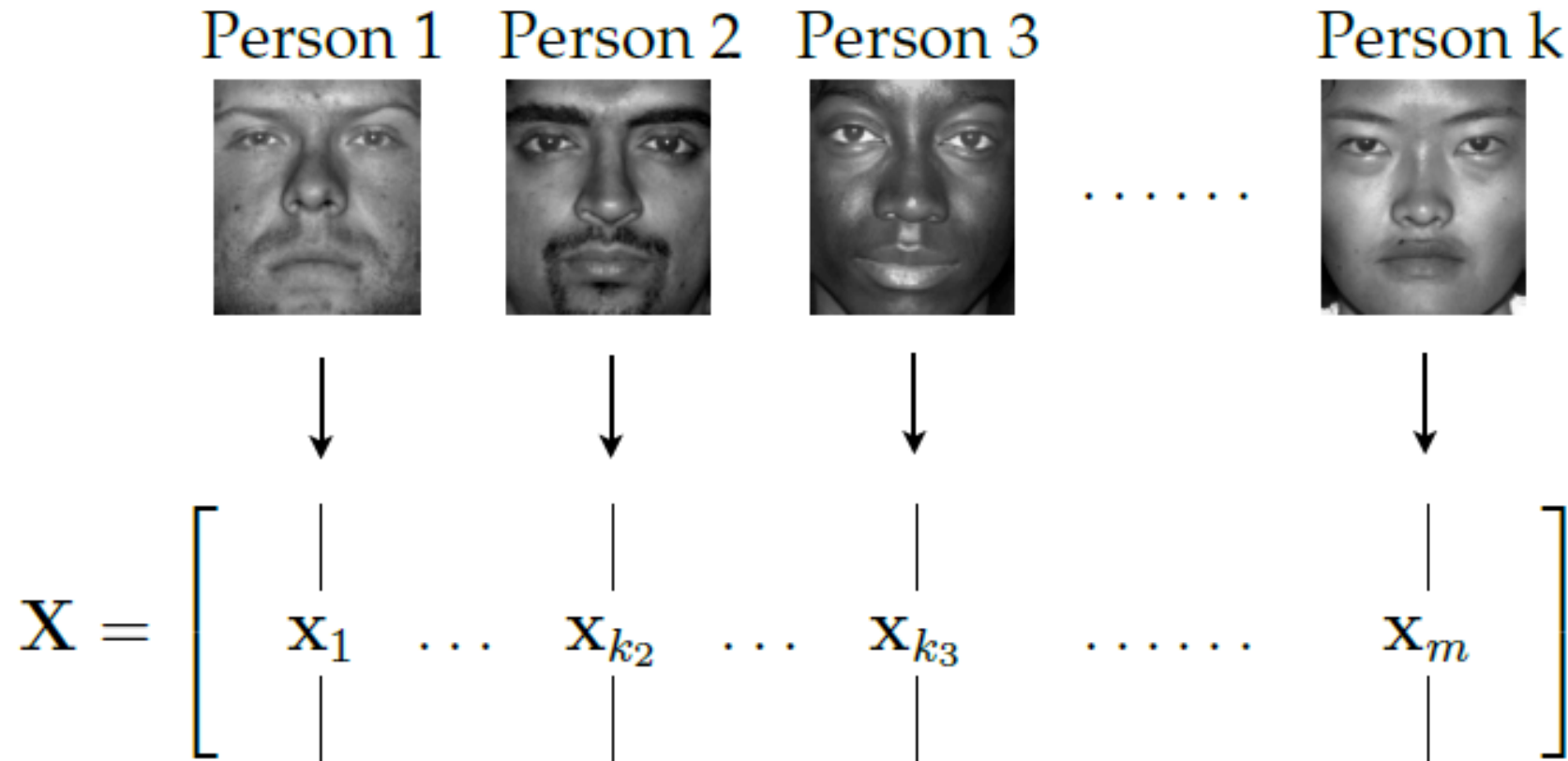


Mathematical definition (3): Dominant correlation interpretation

$$X = U\Sigma V^T \Rightarrow \begin{cases} XX^T U = U\Sigma^2 \\ X^T X V = V\Sigma^2 \end{cases}$$

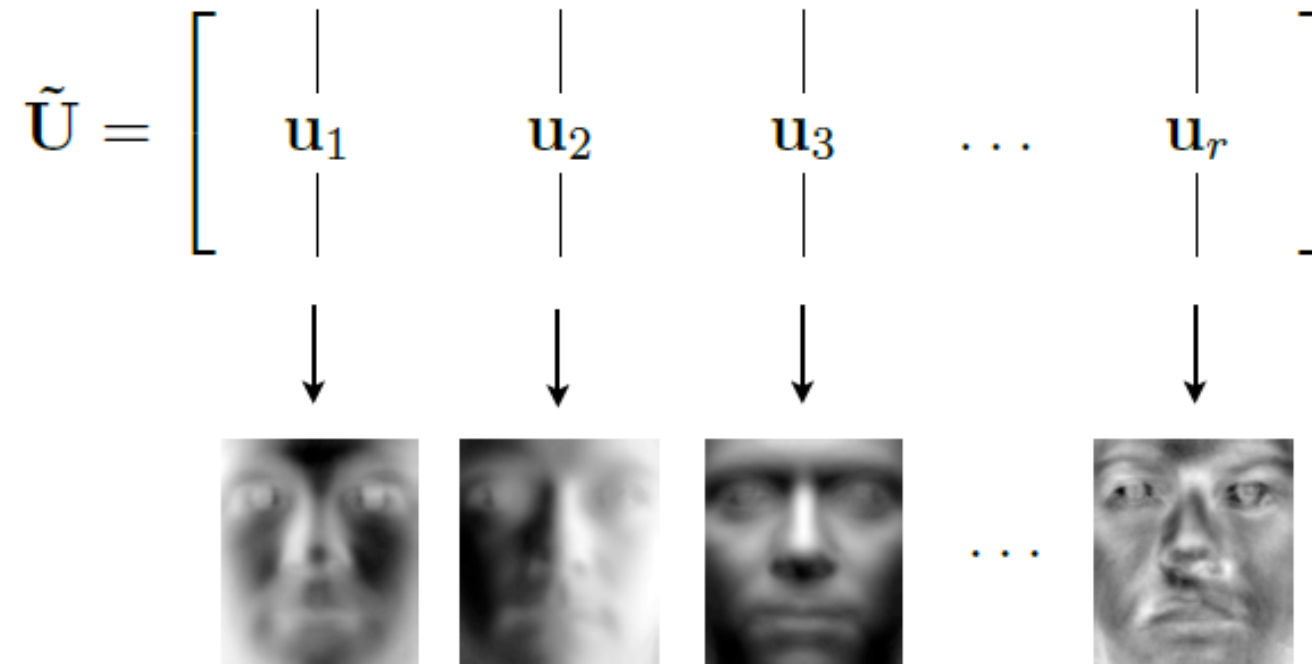
- Columns of U are the eigenvectors of the correlation matrix XX^T
- Columns of V are the eigenvectors of the correlation matrix $X^T X$
- Columns of U are hierarchically ordered by **how much correlation they capture** in the columns of X ; V similarly captures correlation in the rows of X .
- **The amount of correlation / information** captured by U or V component is **characterized by** their corresponding **singular values**.

SVD application: Face recognition example



SVD application: Face recognition example

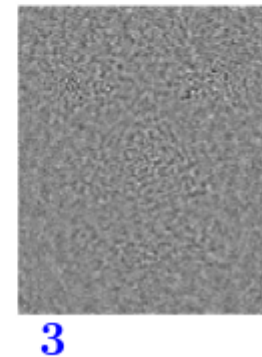
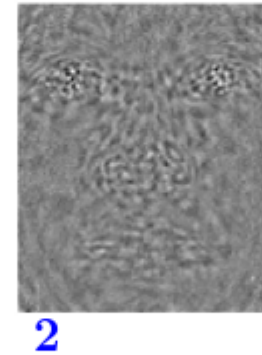
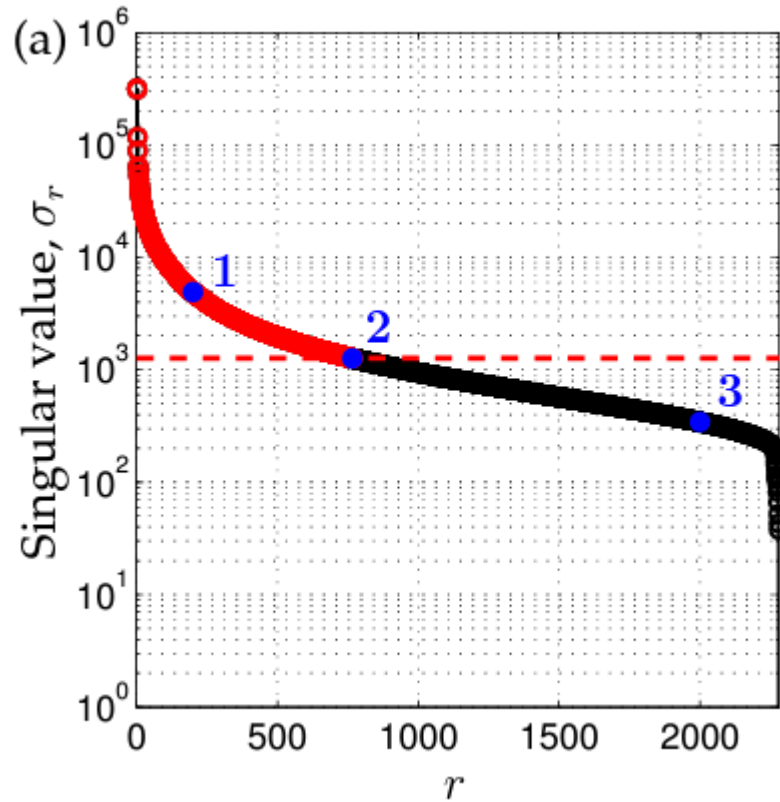
$$X = U\Sigma V^*$$



Eigenfaces

SVD application: Face recognition example

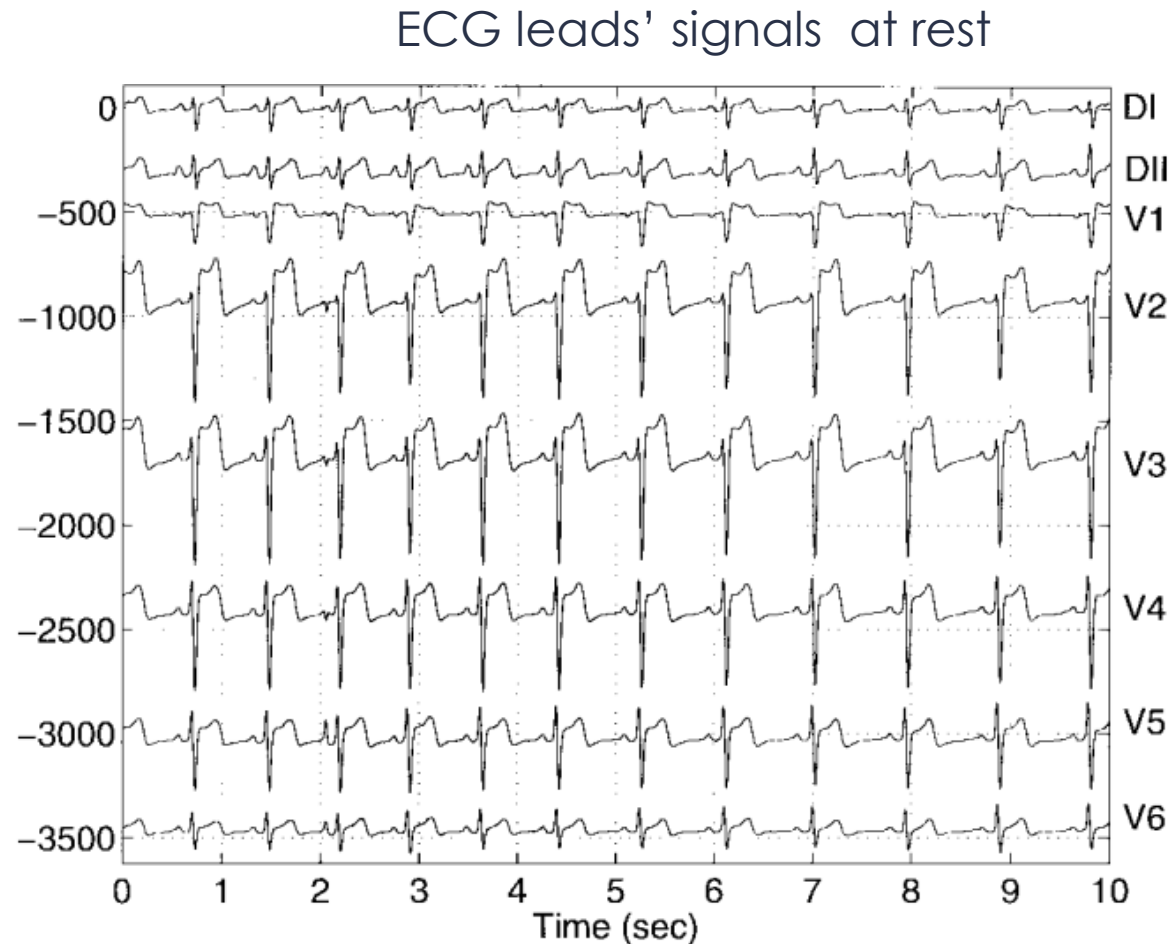
$$X = U\Sigma V^*$$



Eigenfaces

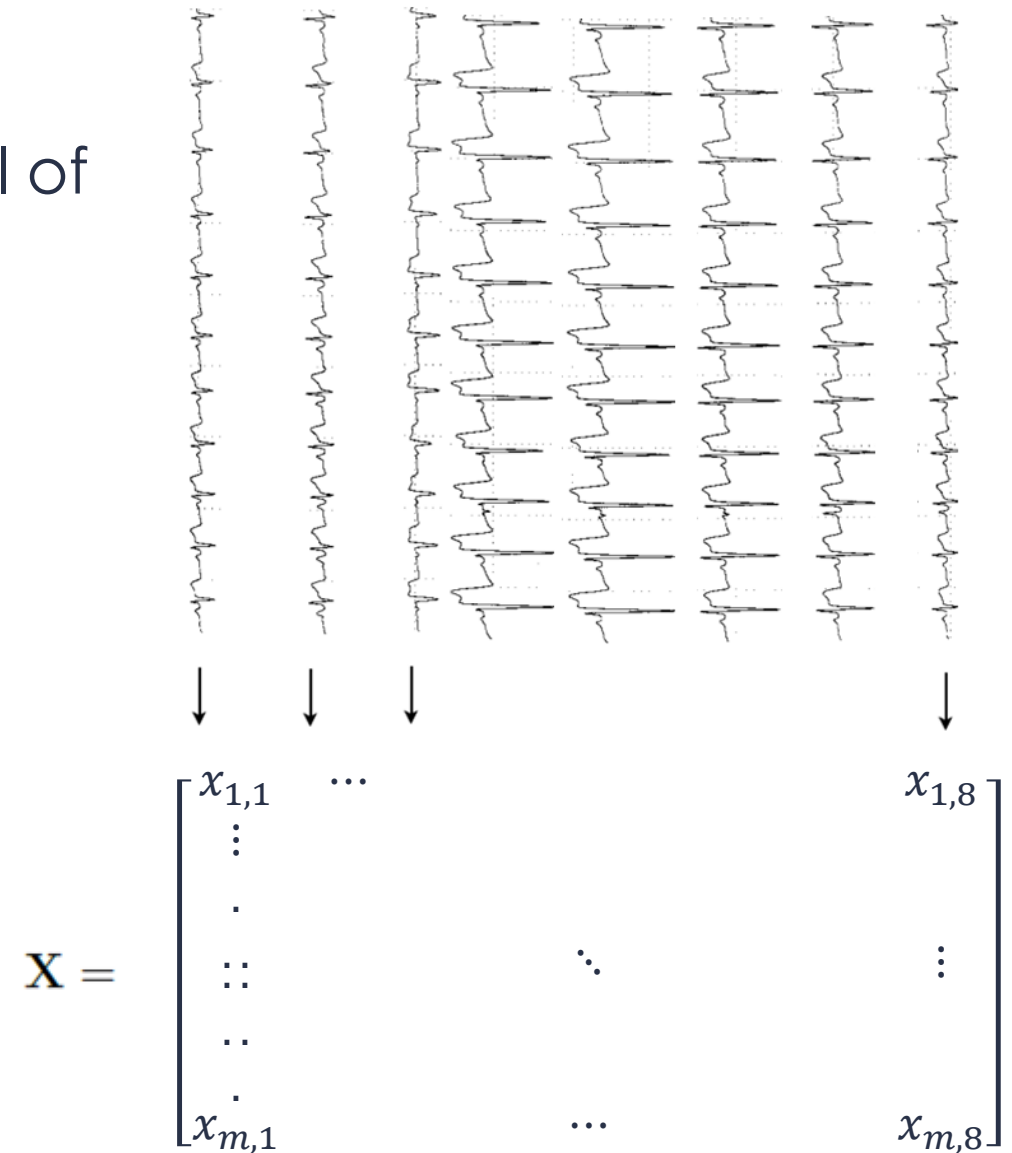
SVD application: Noise reduction

- Example of 8-leads ECG decomposition:



SVD application: Noise reduction

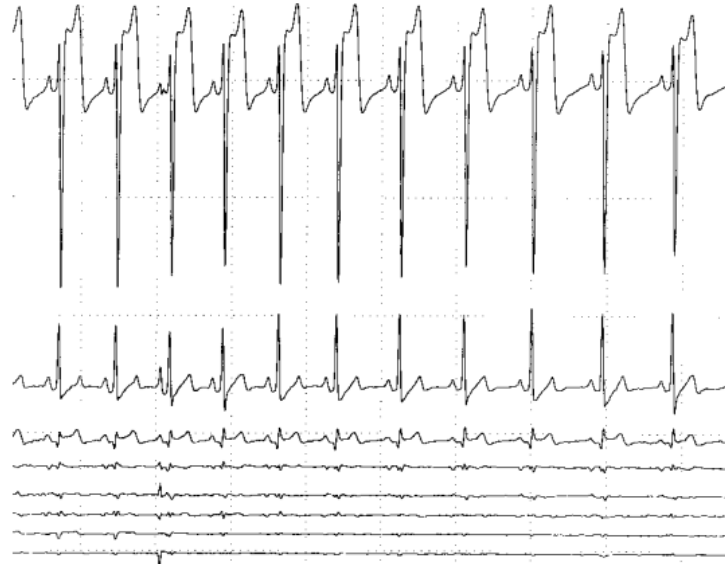
- We fill the matrix X with one lead signal of m samples in each column, giving X , a matrix $m \times 8$.
- One observation / measurement corresponds to one lead signal of m samples.



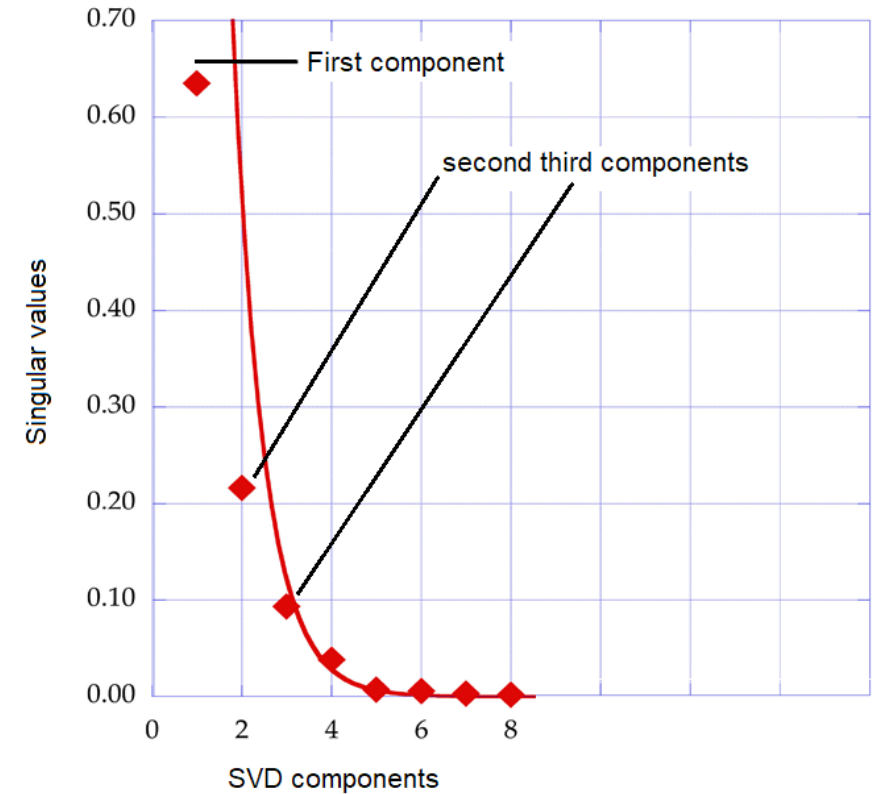
SVD application: Noise reduction

$$X = U\Sigma V^*$$

$$U^* X =$$

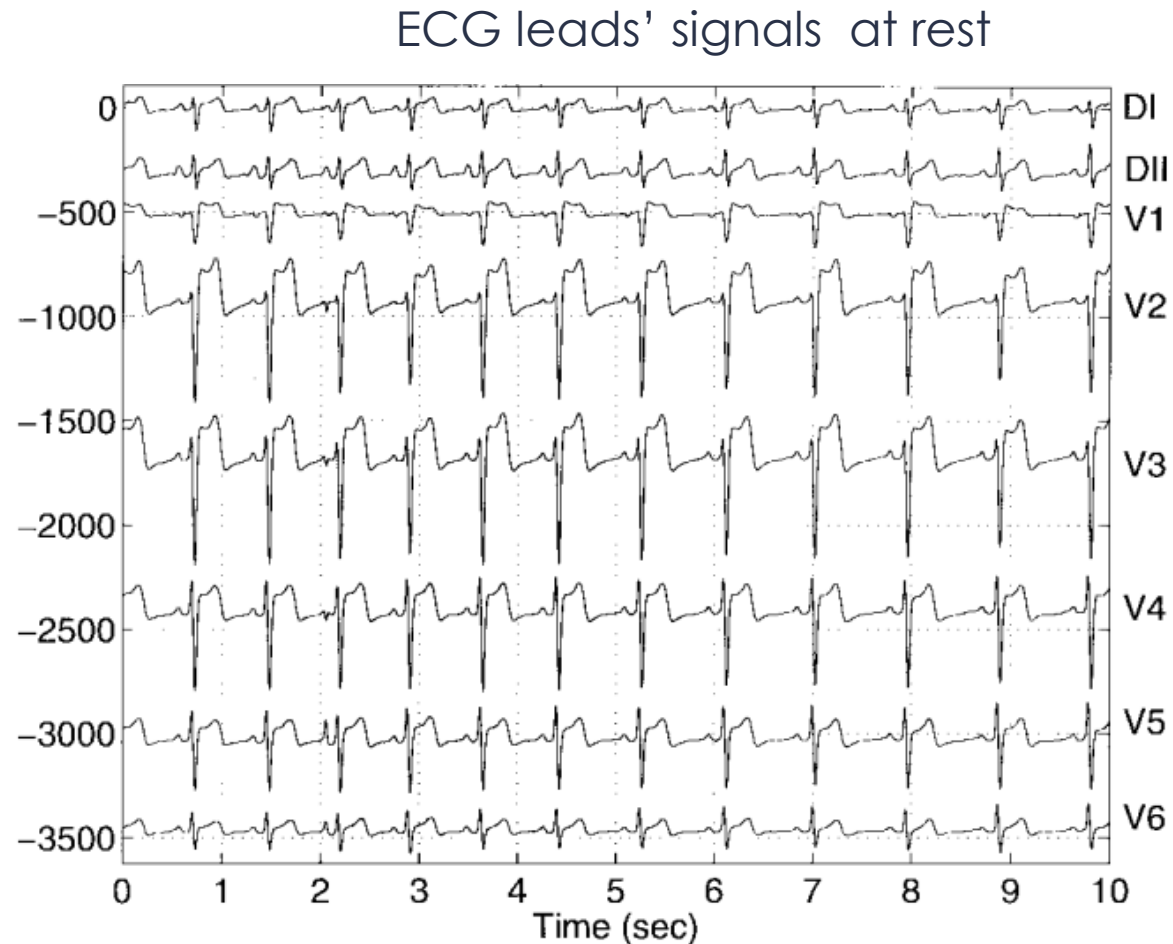


Projections of signal leads (matrix X)
in the SVD components (columns of matrix U)
subspace



SVD application: Noise reduction

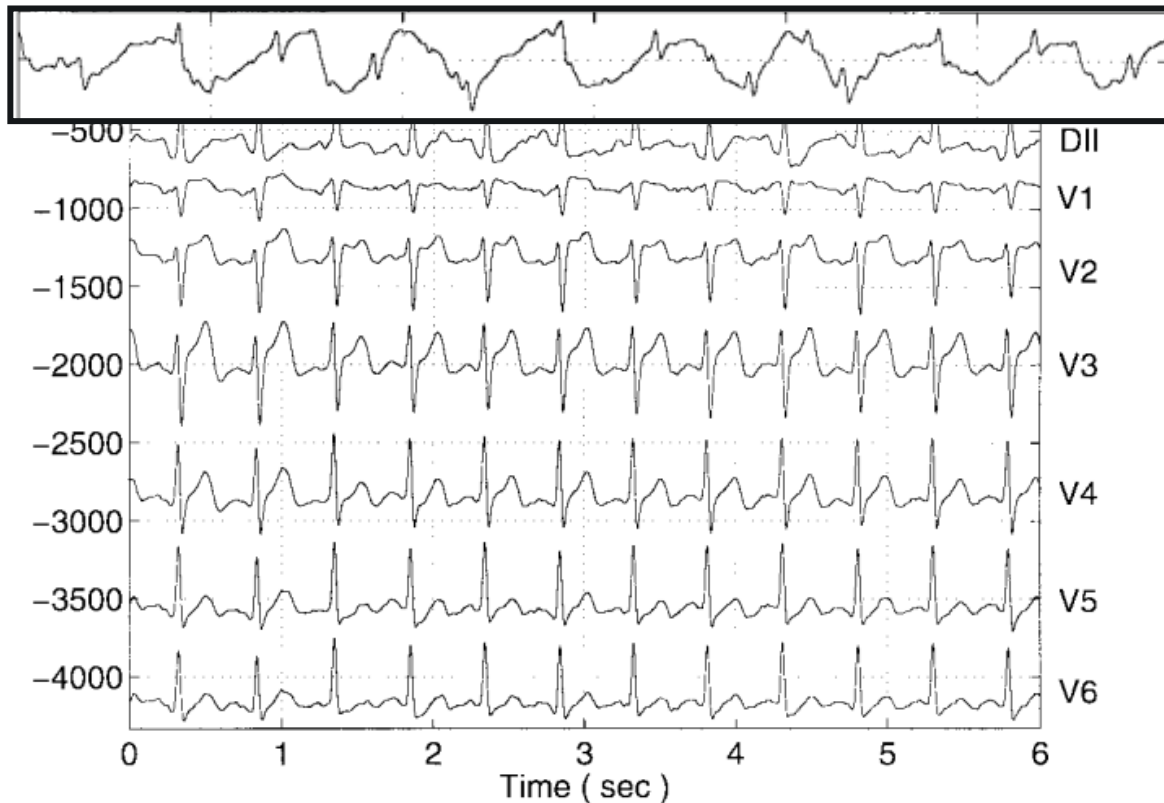
- Example of 8-leads ECG decomposition:



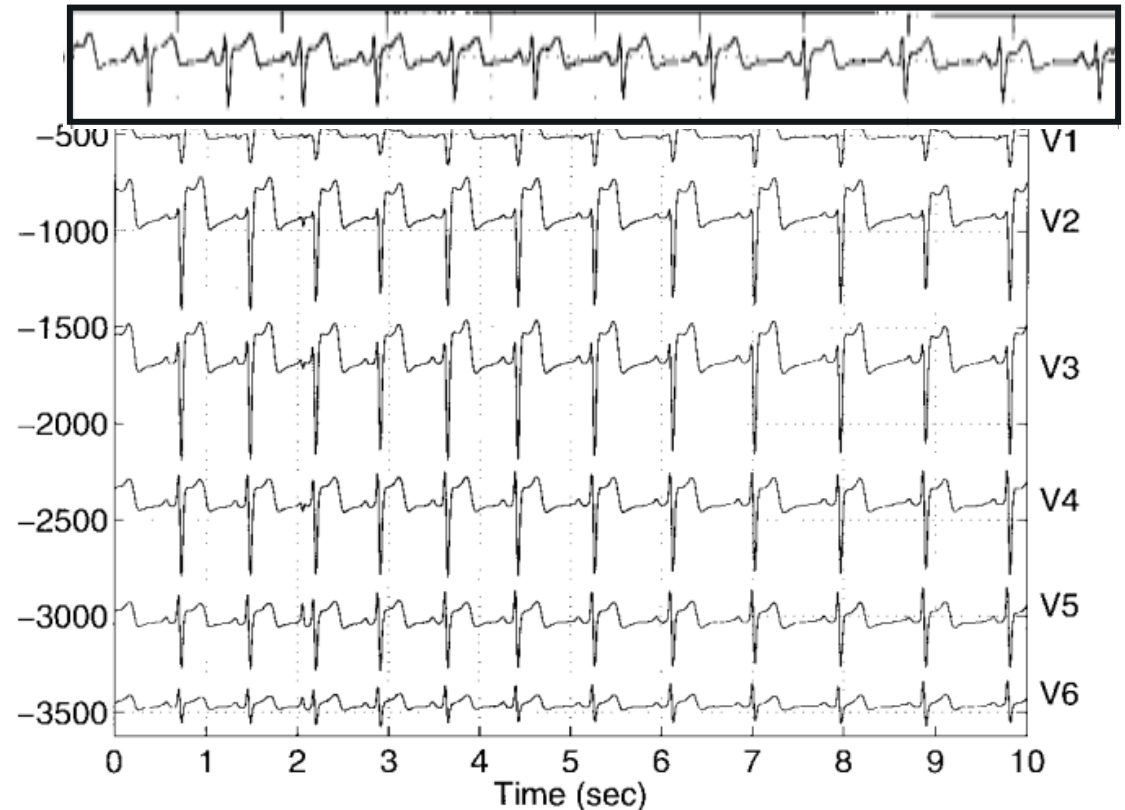
SVD application: Noise reduction

Noise : ECG with baseline wandering (DI)

ECG leads' signals during exercise

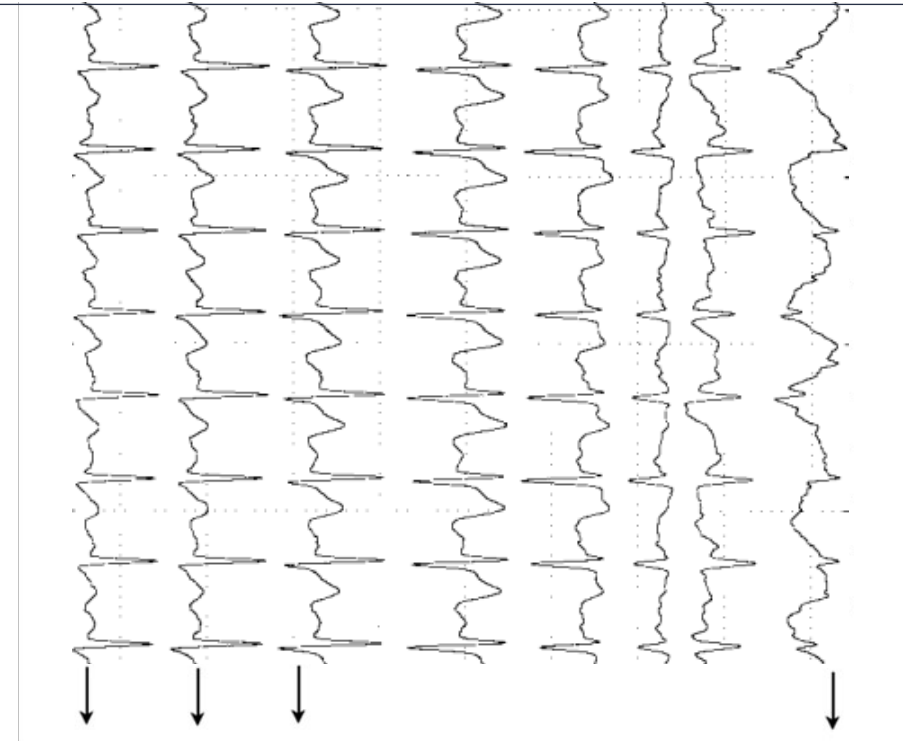


ECG leads' signals at rest



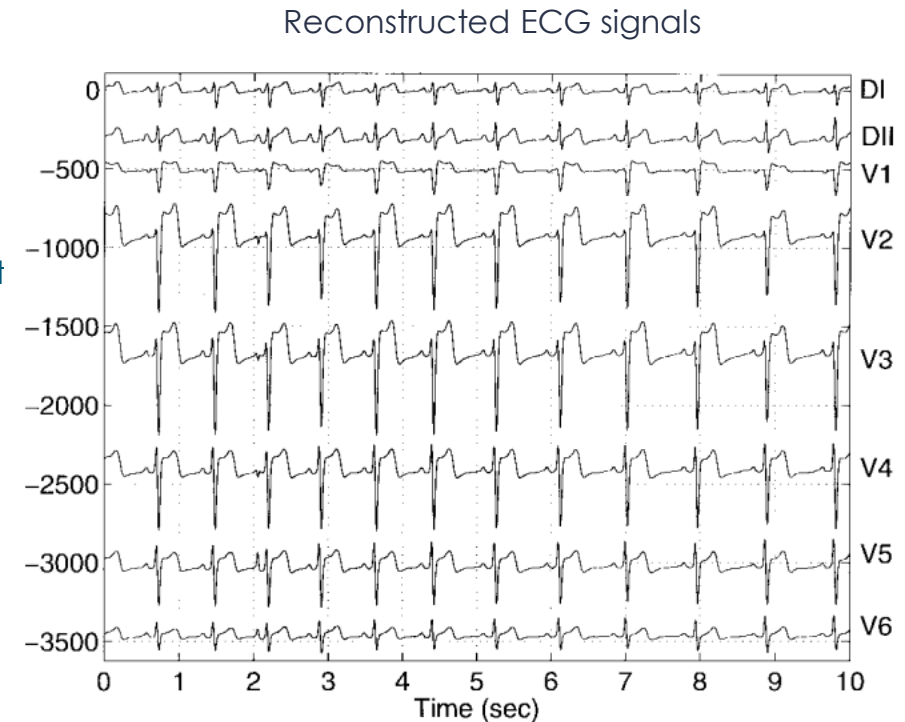
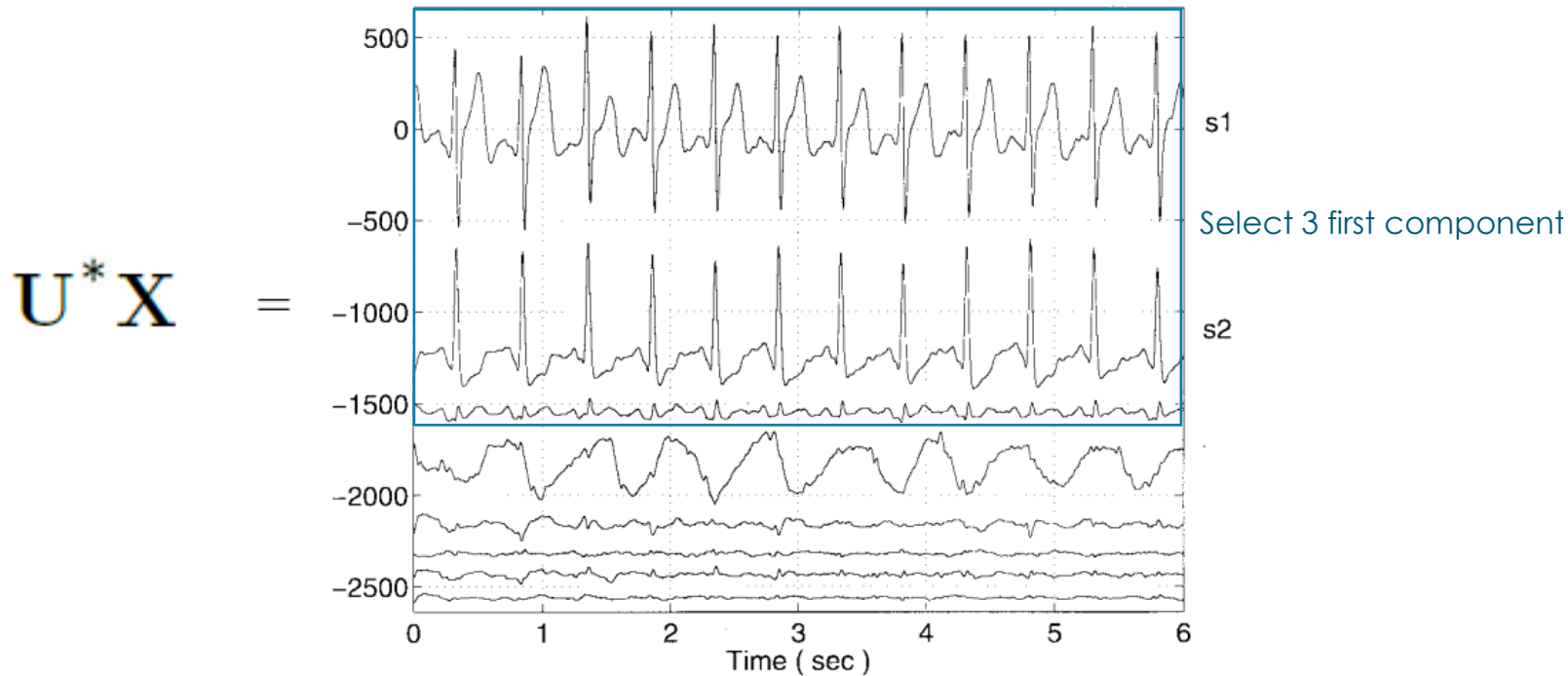
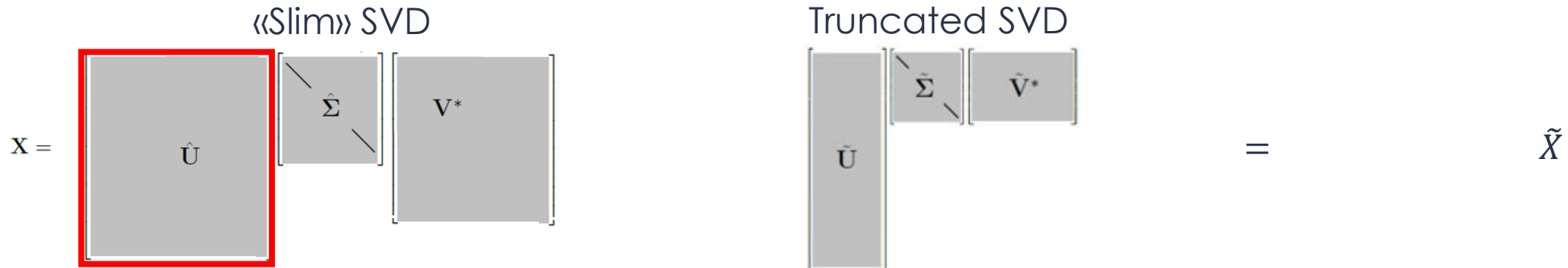
SVD application: Noise reduction

SVD decomposition of X: $X = U\Sigma V^T$



$$X = \begin{bmatrix} x_{1,1} & \dots & & & & & & x_{1,8} \\ \vdots & & & & & & & \vdots \\ \cdot & & & & & & \ddots & \vdots \\ \ddots & & & & & & & \vdots \\ \ddots & & & & & & & \vdots \\ \cdot & & & & & & & \vdots \\ x_{m,1} & & & & \dots & & & x_{m,8} \end{bmatrix}$$

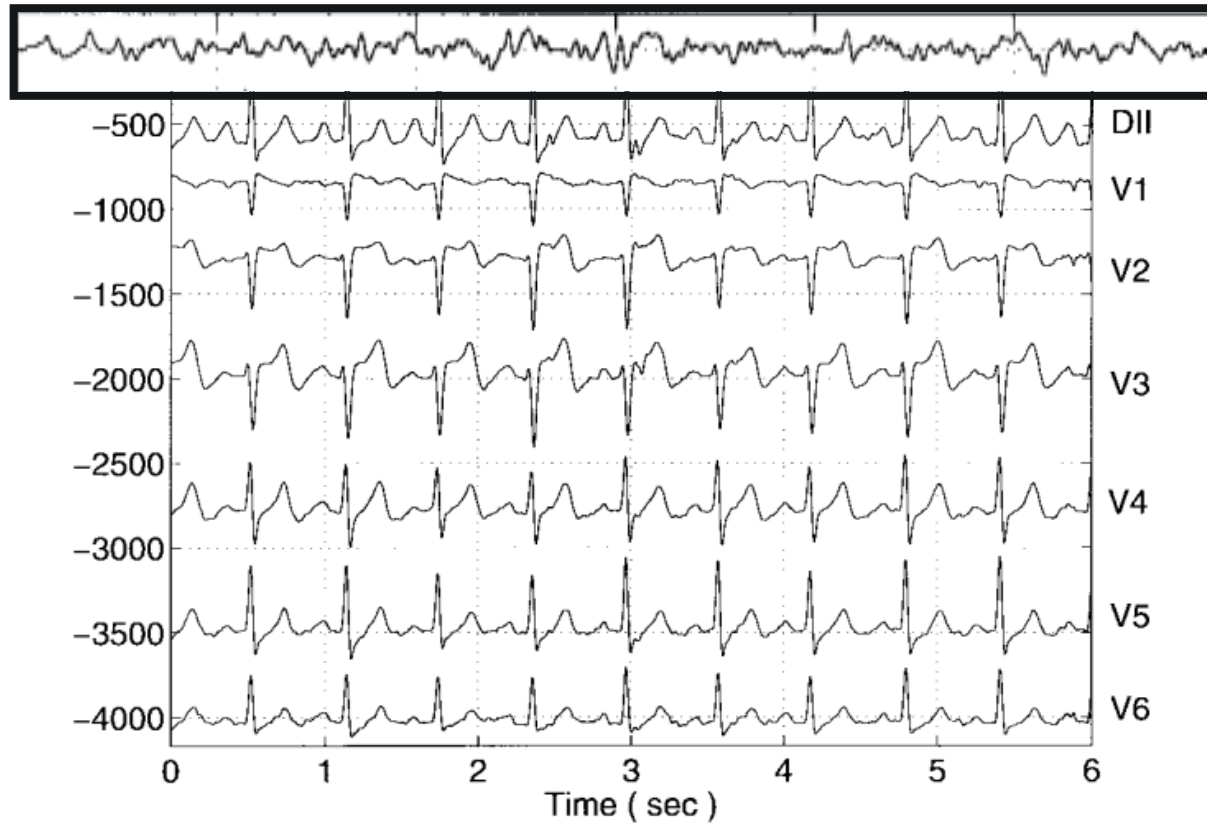
SVD application: Noise reduction



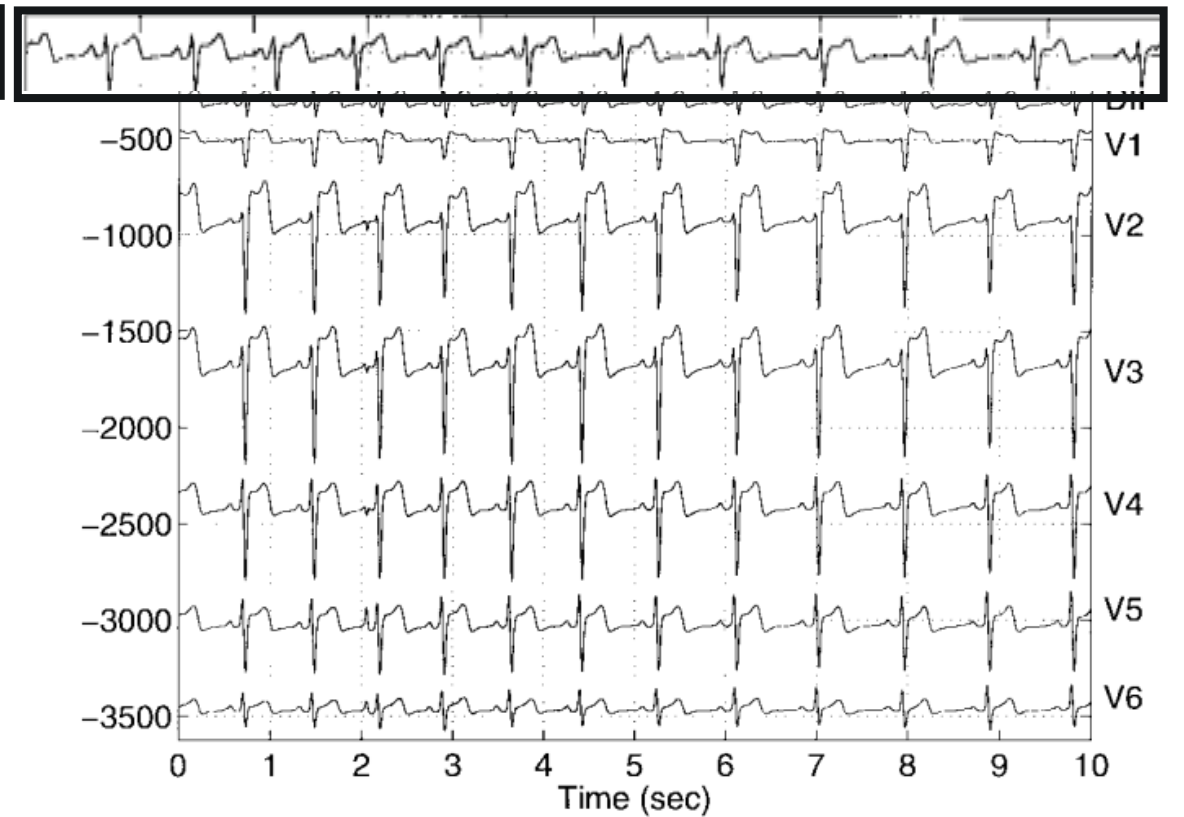
SVD application: Noise reduction

Noise : ECG with electromyogram noise

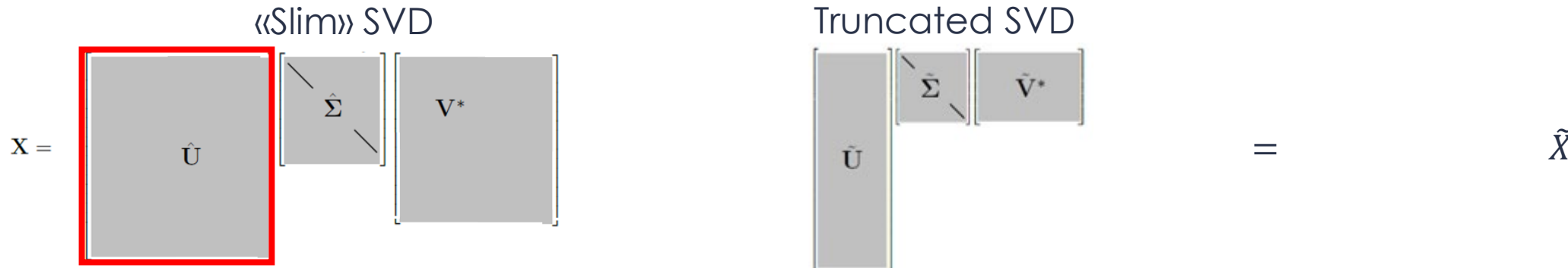
ECG leads' signals during exercise



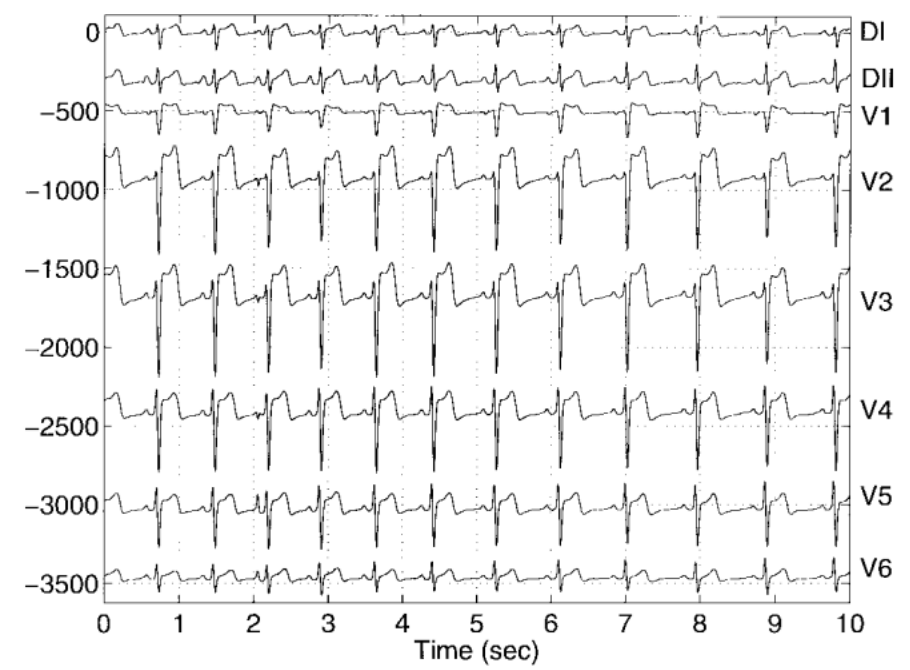
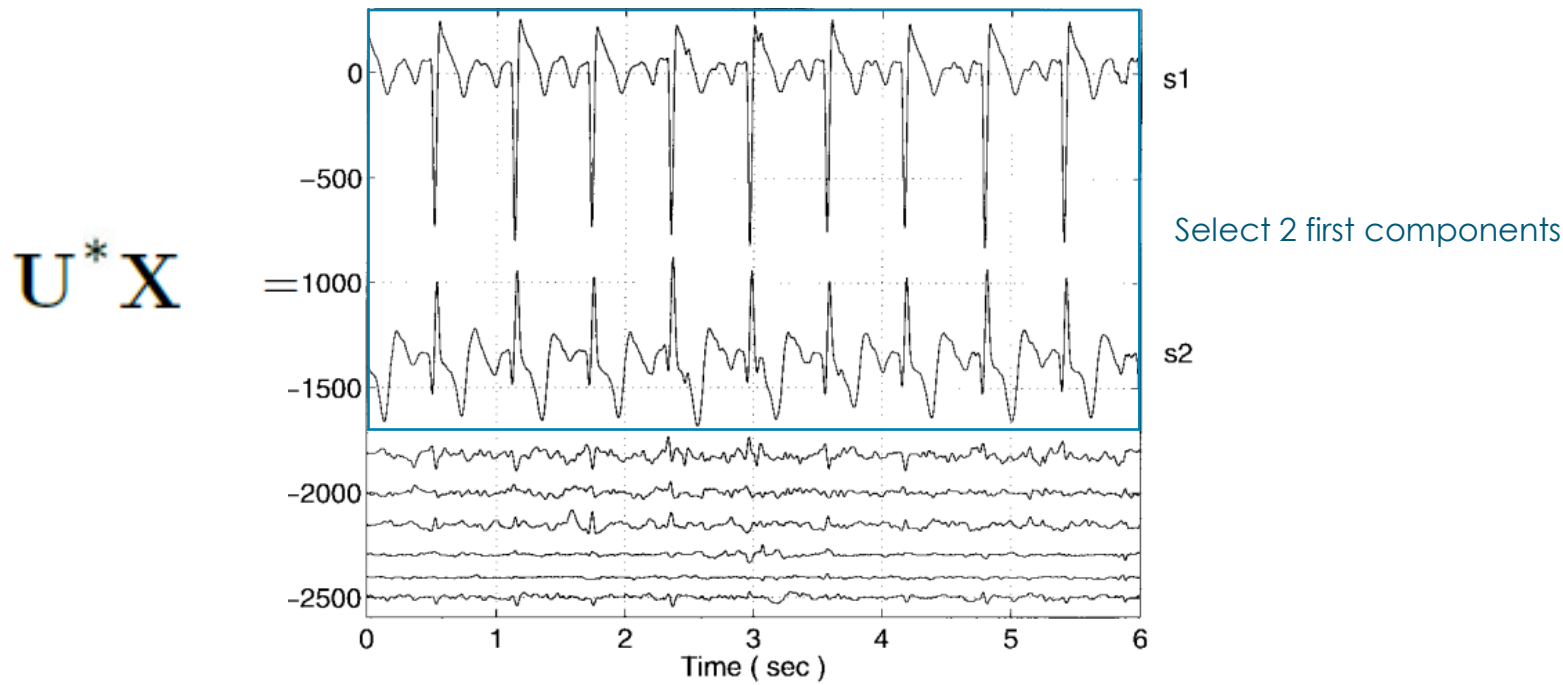
ECG leads' signals at rest



SVD application: Noise reduction



Reconstructed ECG signals

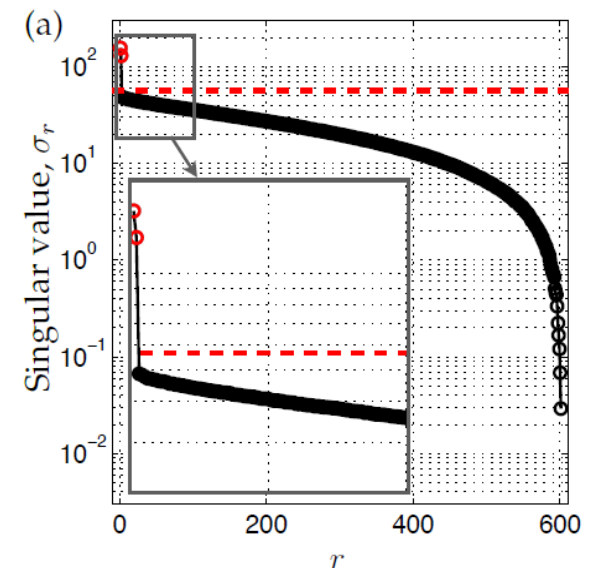
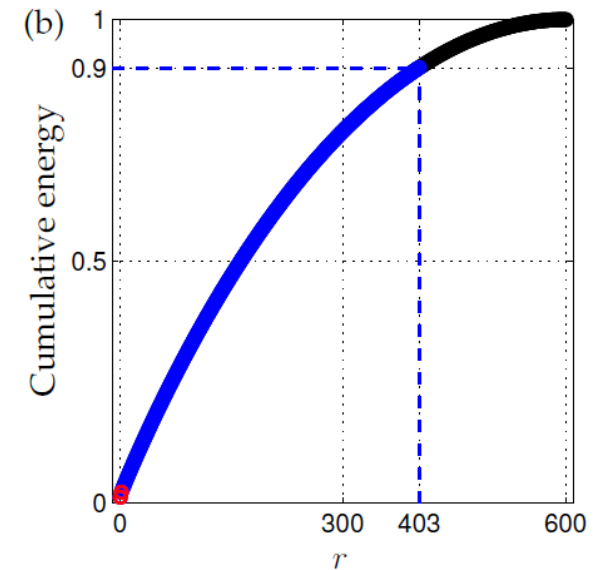


SVD application: Truncation techniques

Often, one truncates the SVD at an effective rank r_e as the smallest i such that:

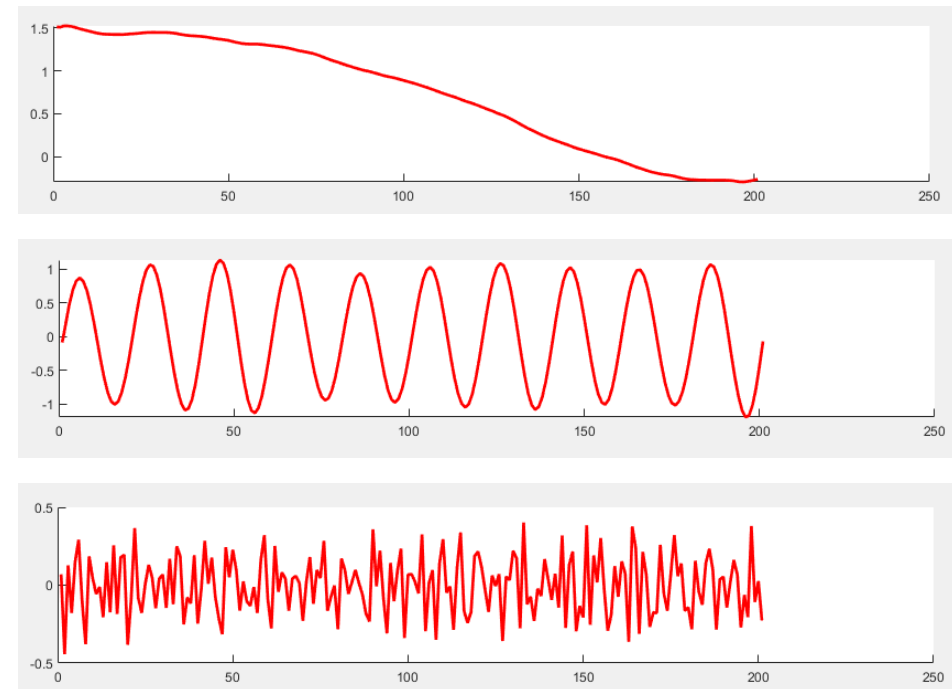
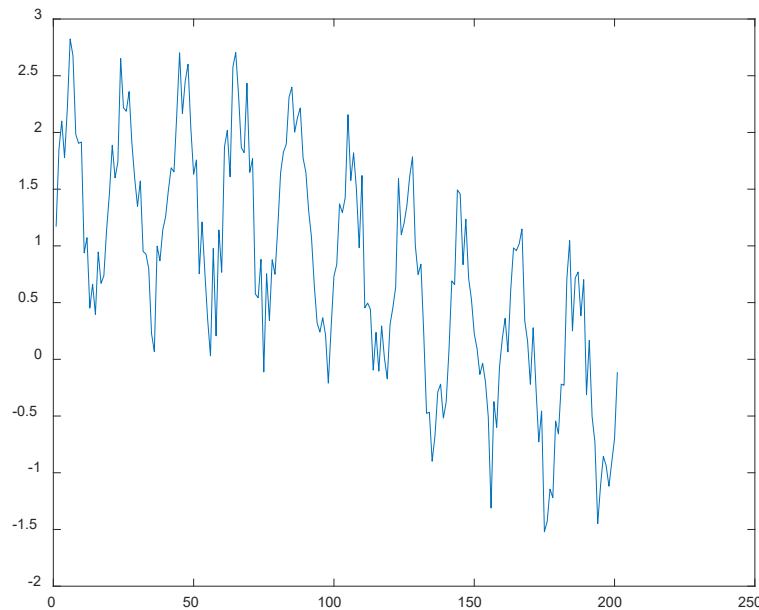
$$\frac{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_i^2}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_r^2} > t, \text{ (e.g. } t = 0.98\text{)}$$

Other techniques involve identifying “elbows” or “knees” in the singular value distribution.

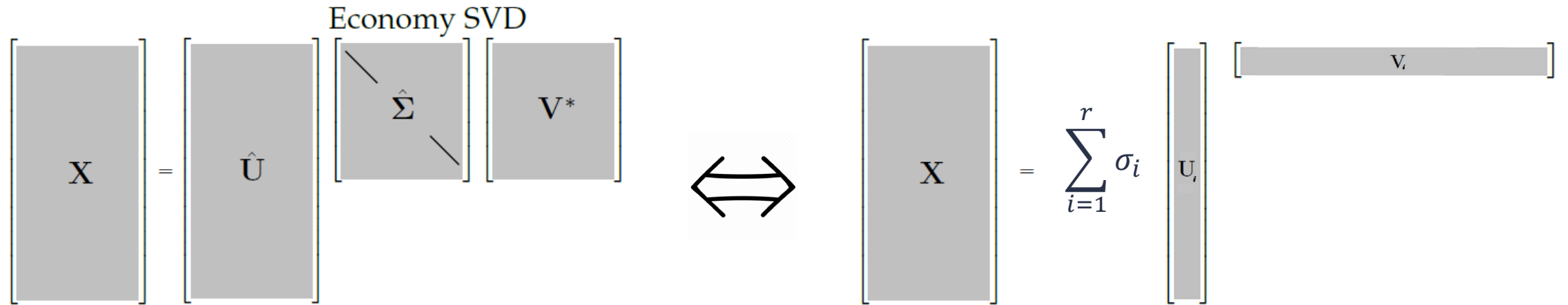


SVD application: Non parametric spectral estimation

- Singular spectrum analysis (SSA) is a principal component analysis that is very suitable for the study of one-dimensional nonlinear time-series data.
- It provides a **decomposition** of a signal **into bandpass components** **without** the need **to define passbands** a priori.



SVD application: Non parametric spectral estimation



31

- The SVD can be expressed as follow with r the rank of matrix X :

$$X = \sum_{i=1}^M \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

SVD application: Non parametric spectral estimation

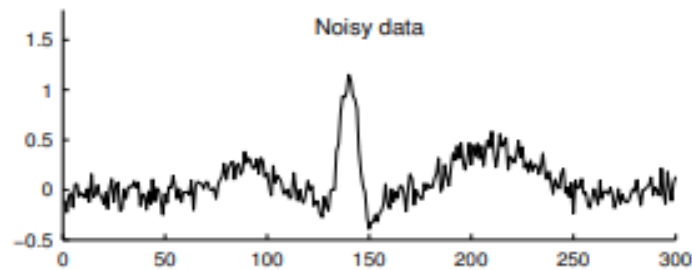
- Dimensionality reduction => Noise reduction

$$X = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_{r-1} u_{r-1} v_{r-1}^T + \sigma_r u_r v_r^T$$

- Each $u_i v_i^T$ is a basis direction of the matrix X, and σ_i the coordinate in this basis. 32
- The amount of X pointing in $u_i v_i^T$ direction is characterized by σ_i .
- Considering a white noise (same in all direction), the first components have higher SNR than the last components.

SVD application: Non parametric spectral estimation

- From time serie signal to a matrix:



$$X = \begin{bmatrix} x(1) & x(2) & x(3) & \cdots & x(L) \\ x(2) & x(3) & x(4) & \cdots & x(L+1) \\ x(3) & x(4) & x(5) & \cdots & x(L+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x(N-L+1) & x(N-L) & \cdots & \cdots & x(N) \end{bmatrix}$$

33

- Matrix X has Hankel-type structure: it has the same values on all its anti-diagonals

SVD application: Non parametric spectral estimation

- Let's decompose X using SVD:

$$X = \sum_{i=1}^L \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

- The exterior products $\mathbf{u}_i \mathbf{v}_i^T$ correspond to matrices with the same size as X , that is $(N-L+1) \times L \dots$

...However, these matrices do not necessarily have a Hankel-type structure.

SVD application: Non parametric spectral estimation

- Singular spectrum analysis central idea: Transform $\mathbf{u}_i \mathbf{v}_i^T$ matrices in Hankel-type matrices, with the weighted sum of the matrices equal to X .

Replace each coefficient on the anti-diagonal in each matrix by the average of the coefficients on this anti-diagonal, and one still gets X .

$$\begin{bmatrix} 5 & 10 \\ 10 & 6 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 4 & 2 \\ 4 & 3 \end{bmatrix} + \begin{bmatrix} 3 & 8 \\ 6 & 4 \\ 2 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3 \\ 3 & 3 \\ 3 & 3 \end{bmatrix} + \begin{bmatrix} 3 & 7 \\ 7 & 3 \\ 3 & 0 \end{bmatrix}$$

SVD application: Non parametric spectral estimation

- Now the transformed matrices are Hankel-type, and signals $w_i(n)$, $i = 1, \dots, L$, can be extracted. One has:

$$x(n) = \sum_{i=1}^L \sigma_i w_i(n)$$

- Note that in the resulting decomposition the component variances depend on the singular values of \mathbf{X} . The less \mathbf{v}_i , \mathbf{u}_i , 'explain' \mathbf{X} , the smaller the component.

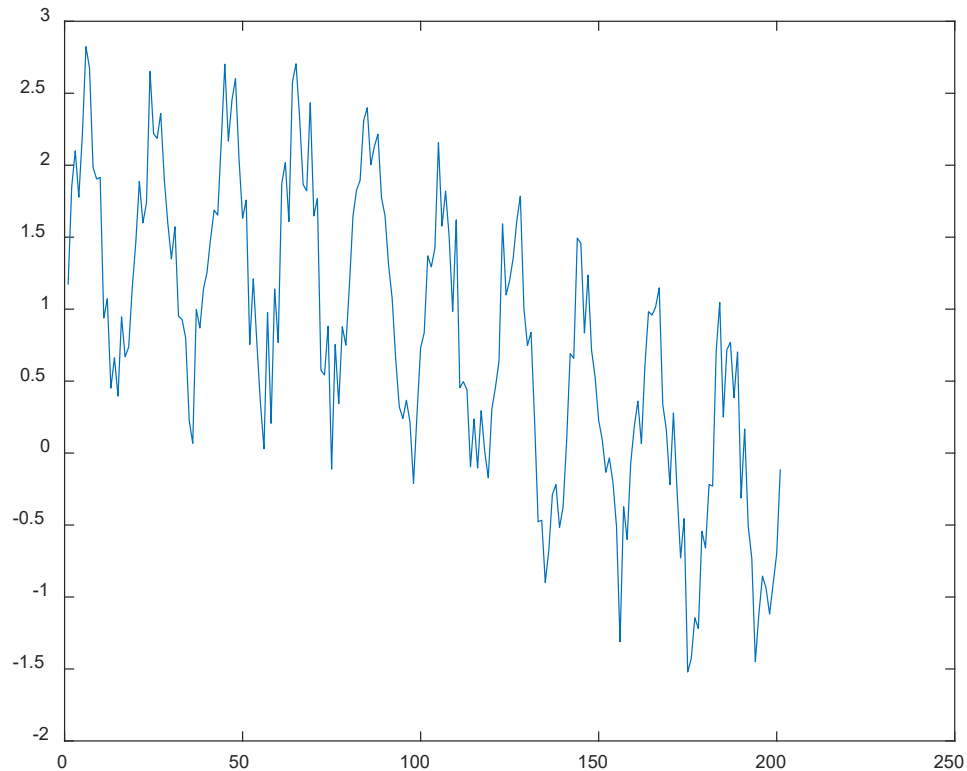
SVD application: Non parametric spectral estimation

- The only **parameter to select is L** , the number of columns of **\mathbf{X}** . This choice is mostly based on trial-and-error.
- The **components w_i are** usually presented in decreasing order of variance, which directly **follows** the decreasing **order of the singular values**.
- These **components are** usually **bandpass signals**, and their peak frequency can be estimated using non-parametric or parametric spectral estimation.

SVD application: Non parametric spectral estimation

- Example: 200-sample signal

$$x(n) = \sin(2\pi*0.0025*n+\pi/3) + \sin(2\pi*0.05*n) + \text{white noise } (\sigma^2 = 0.04)$$



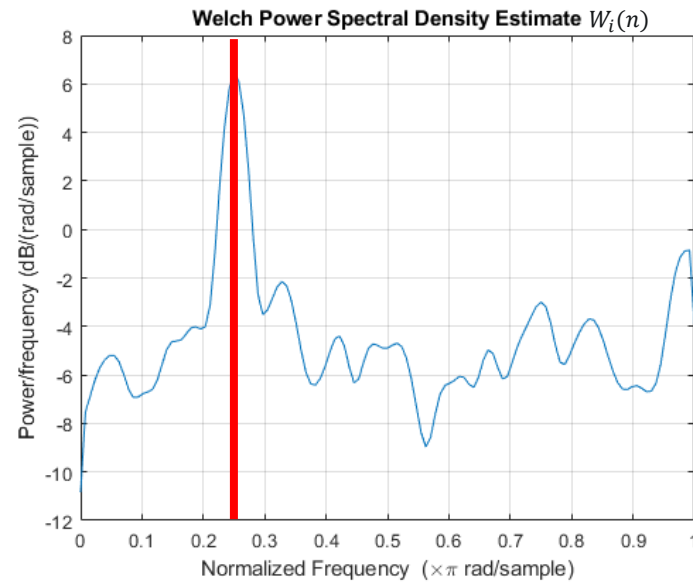
$$X = \sum_{i=1}^L \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

```
>> [Uhat, Shat, V] = svd(X, 'econ');
```

SVD application: Non parametric spectral estimation

- With $L = 20$, components variance versus frequency:

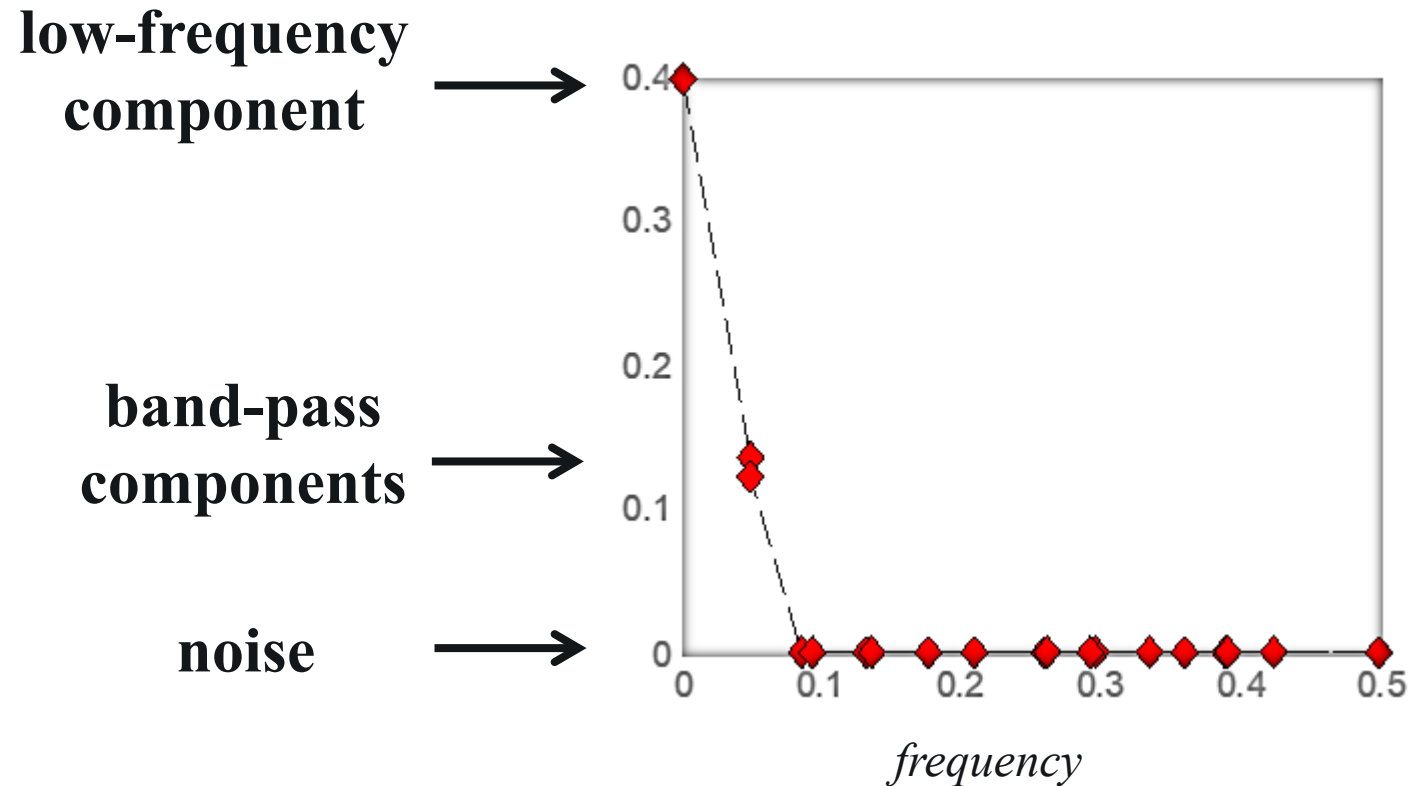
$$\mathbf{X} = \begin{bmatrix} x(1) & x(2) & x(3) & \cdots & x(L) \\ x(2) & x(3) & x(4) & \cdots & x(L+1) \\ x(3) & x(4) & x(5) & \cdots & x(L+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x(N-L+1) & x(N-L) & \cdots & \cdots & x(N) \end{bmatrix} \xrightarrow{\text{SVD}} \mathbf{X} = \sum_{i=1}^L \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$



$$\xrightarrow{\text{SSA}} x(n) = \sum_{i=1}^L \sigma_i w_i(n)$$

SVD application: Non parametric spectral estimation

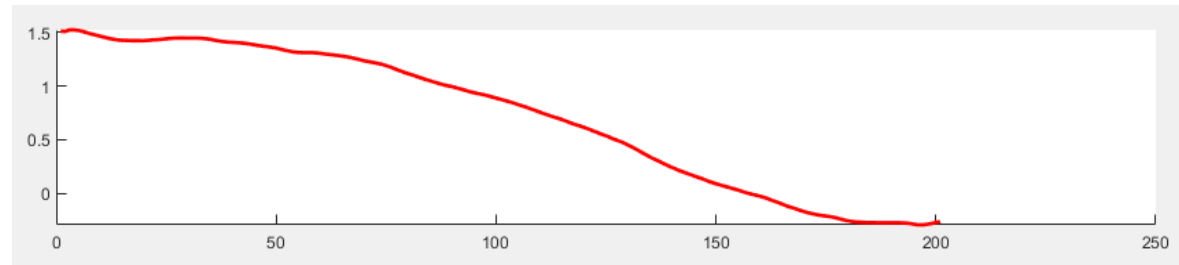
- components variance versus frequency:



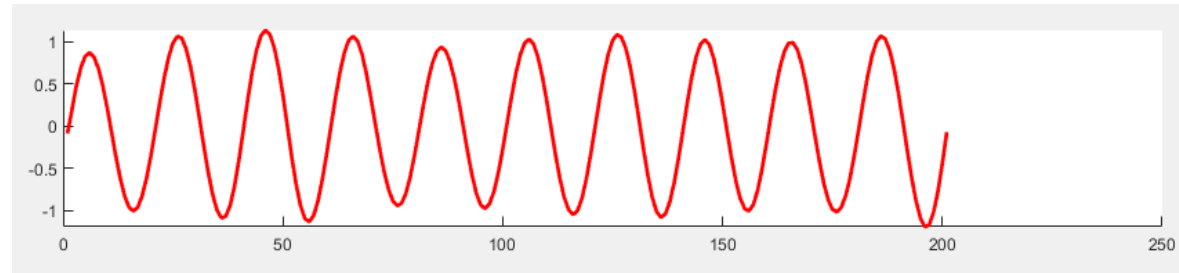
SVD application: Non parametric spectral estimation

- Decomposition:
$$x(n) = \sum_{i=1}^{20} \sigma_i w_i(n) = \sum_{i=1}^{20} y_i(n)$$

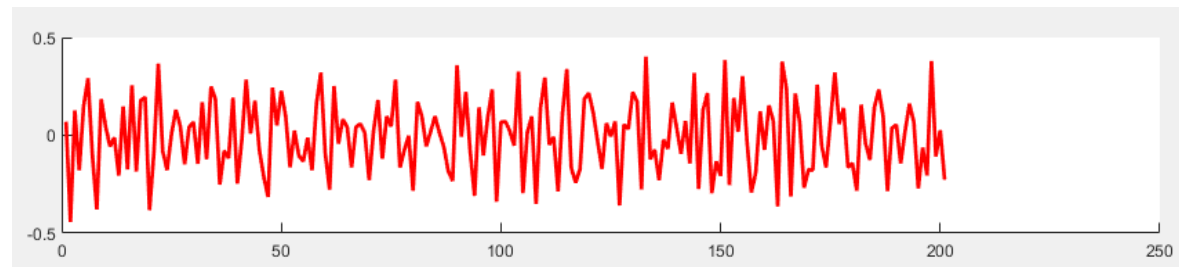
$$y_1 = \sigma_1 w_1 : \sin(2\pi * 0.0025 * n + \pi/3)$$



$$y_2 = \sigma_2 w_2 + \sigma_3 w_3 : \sin(2\pi * 0.05 * n)$$



$$y_3 = \sum_{i=4}^{20} \sigma_i w_i(n) : \text{white noise}$$



SVD application: Least square solution

- We can use SVD to obtain low-rank approximations to matrices and to perform pseudo-inverses of non-square matrices to solve with respect to \mathbf{c} the equation $\mathbf{X}\mathbf{c} = \mathbf{y}$ in the least-squares sense.

$$\mathbf{c}' = \min_{\mathbf{c}} \|\mathbf{X}\mathbf{c} - \mathbf{y}\|_2^2$$

SVD application: Least square solution

$$c' = \min_c \|Xc - y\|_2^2$$

Solution is :

$$c' = X^+ y$$

with X^+ the pseudo-inverse of X from the exact truncated SVD:

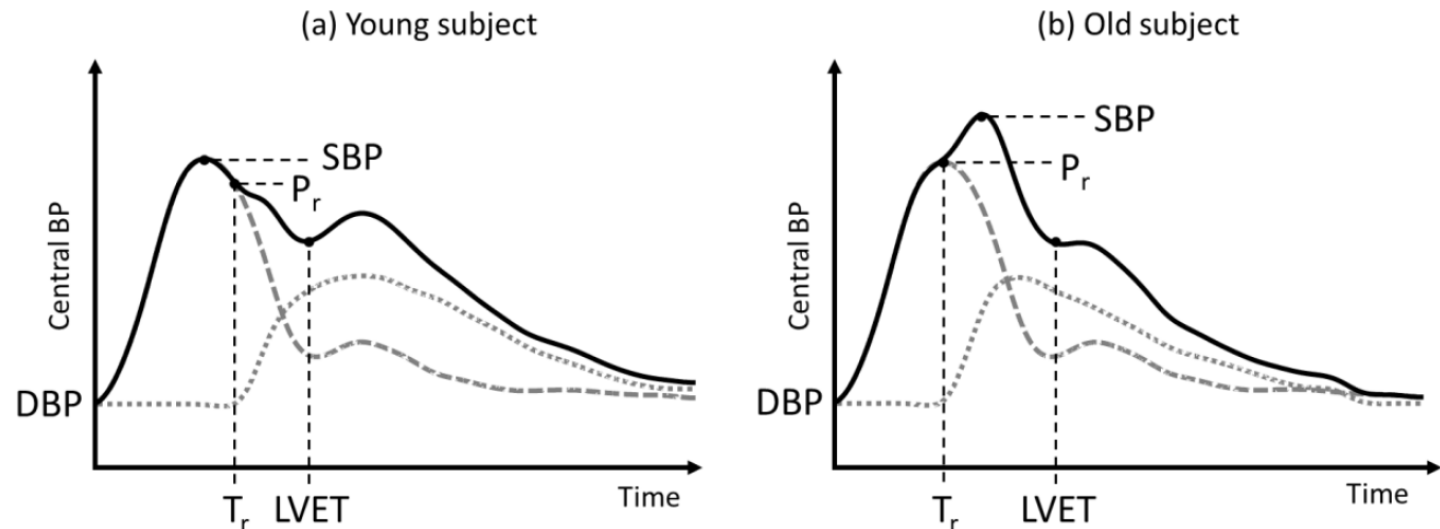
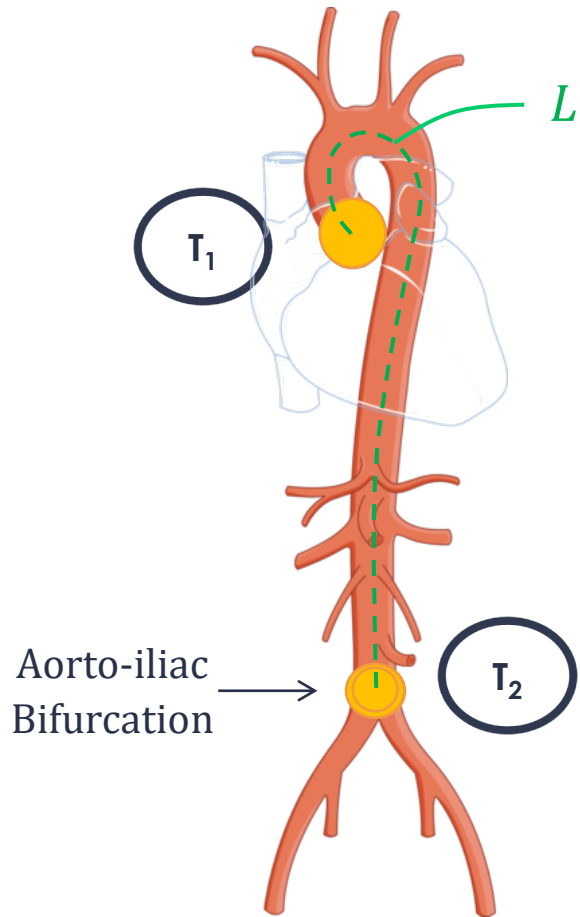
$$X^+ = \tilde{V} S^+ \tilde{U} = \sum_{i=1}^r \frac{1}{\sigma_i} v_i u_i^T \quad \text{with } S^+ = \begin{bmatrix} \tilde{\Sigma}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

SVD application: Least square solution

oBPM[®]: Physiological Background & Algorithm Pipeline

$$\text{Central Pulse Wave Velocity} = \frac{L}{T_2 - T_1} = \sqrt{1/(\rho\delta)} \propto \sqrt{1/\delta} \quad (\text{Bramwell-Hill equation})$$

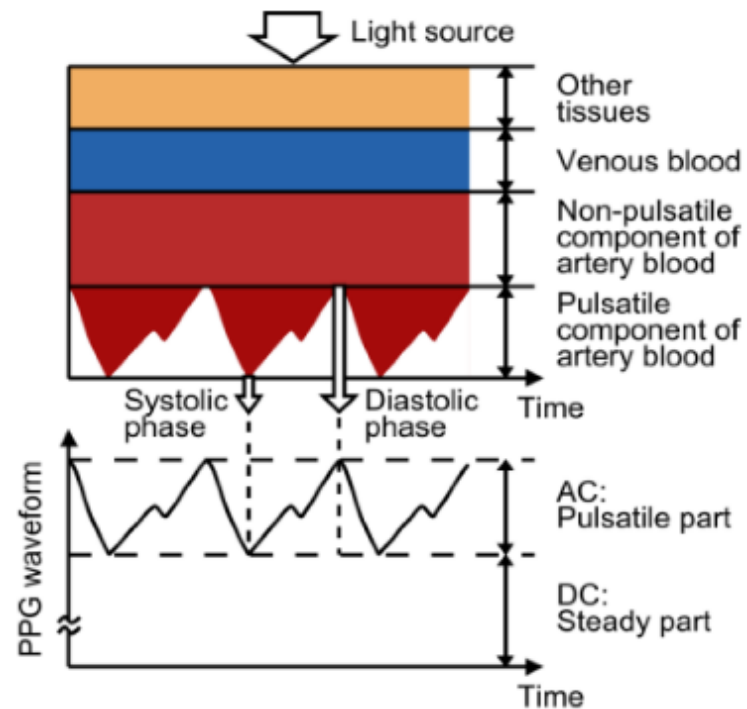
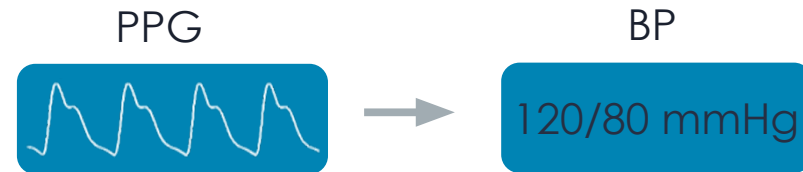
- δ = Aortic distensibility
- ρ = Density of blood (constant)
- L = Aorta length (constant)



$$T_r = 2(T_2 - T_1) = \text{Aortic time to reflection}$$

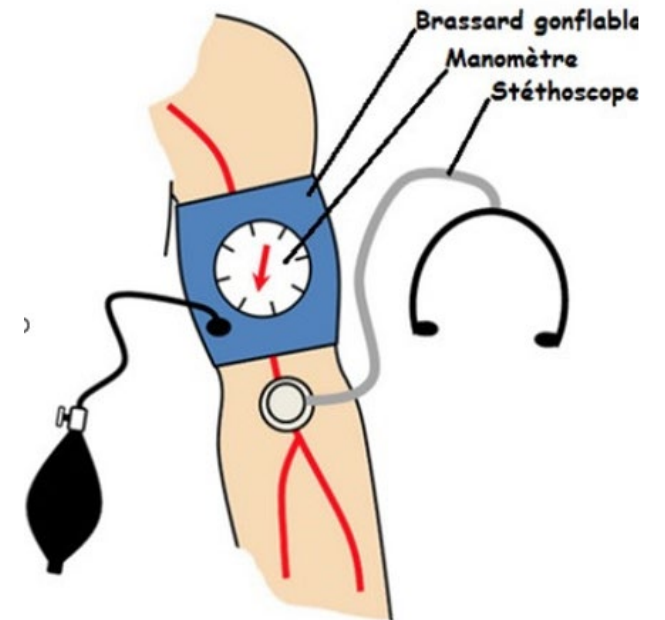
SVD application: Least square solution

- Least square problem example: Estimate blood pressure from PPG



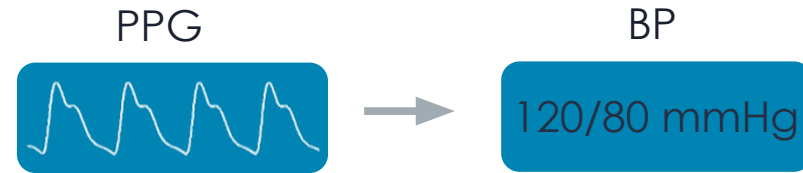
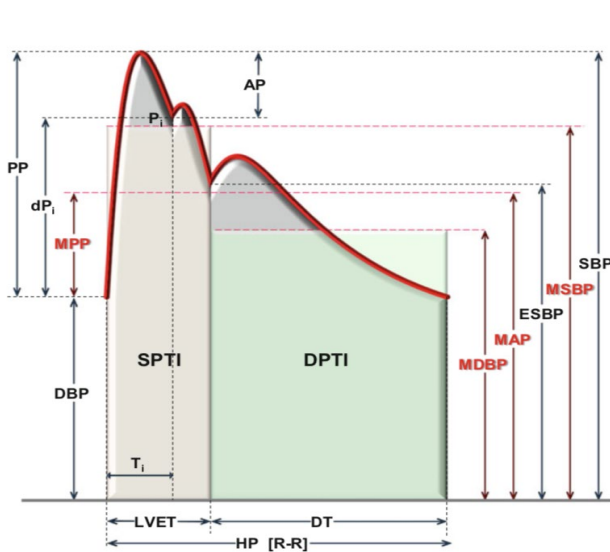
Sys: 120

Dia: 80

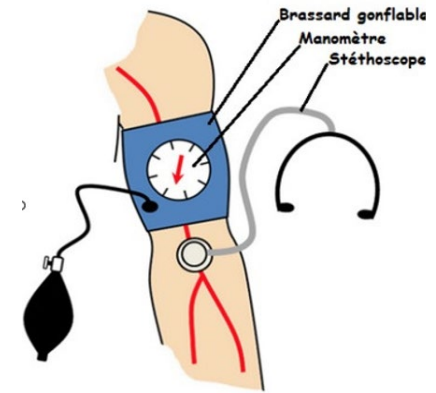


SVD application: Least square solution

- Least square problem example: Estimate blood pressure from PPG



$$Xc = y$$



$$X = \begin{matrix} x_{1,1} & \dots & x_{1,p} \\ \vdots & & \vdots \\ x_{m,1} & \dots & x_{m,p} \end{matrix} \begin{matrix} \leftarrow \text{Meas \#1: Pulse wave } p \text{ features} \\ \\ \leftarrow \text{Meas \#m: Pulse wave } p \text{ features} \end{matrix}$$

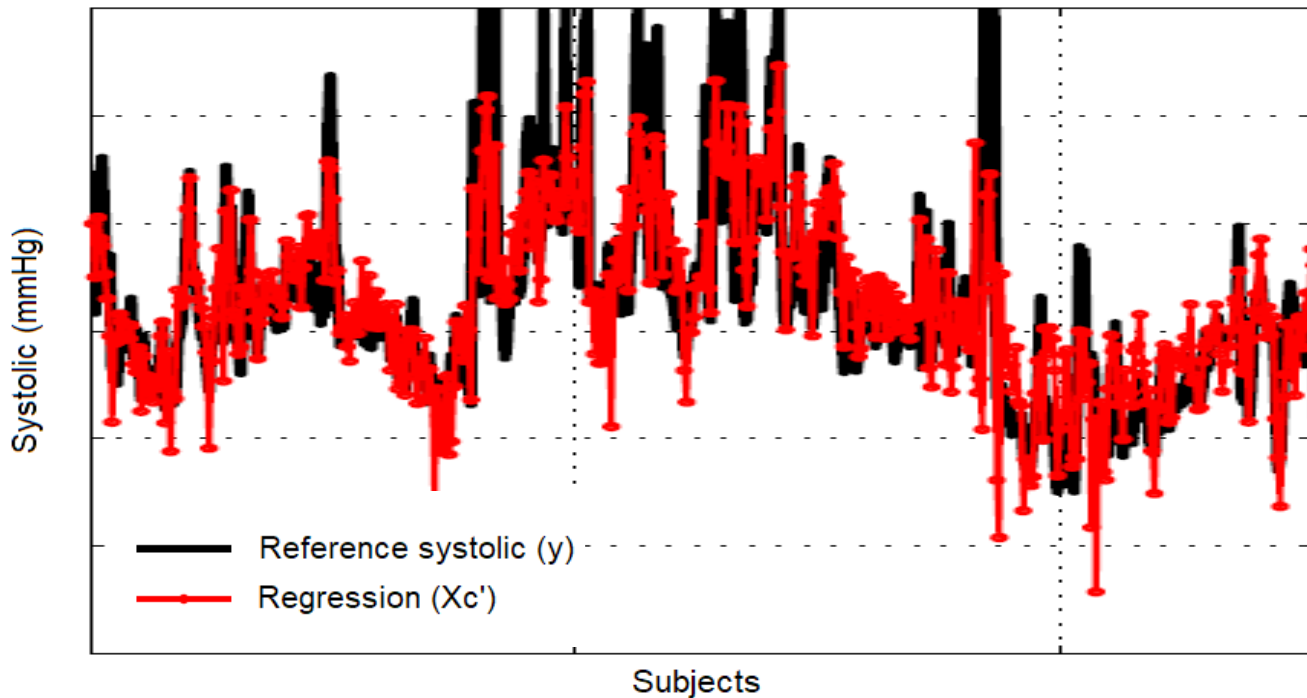
$$y = \begin{matrix} y_1 \\ \vdots \\ y_m \end{matrix} \begin{matrix} \leftarrow \text{Ref \#1: Systolic blood pressure} \\ \\ \leftarrow \text{Ref \#m: Systolic blood pressure} \end{matrix}$$

SVD application: Least square solution

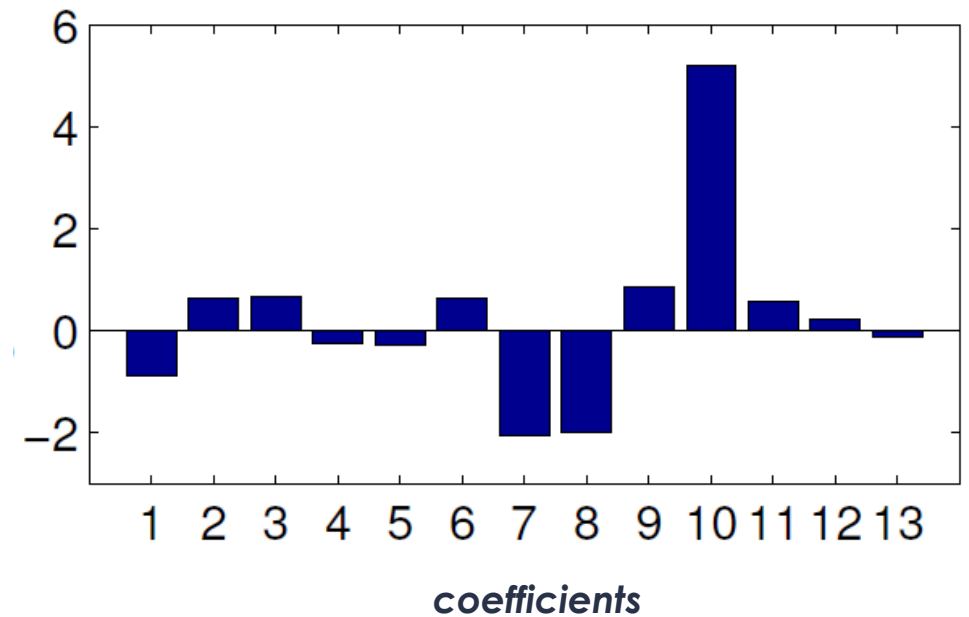
- Example: Estimate blood pressure using PPG.

$$c' = X^+ y$$

Xc'



c'



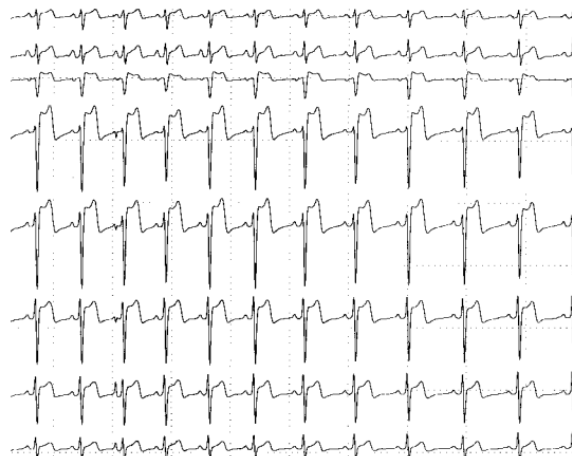
SVD data pre-processing: Data alignment

- The alignment of data significantly impacts the rank of the SVD approximation. The SVD essentially relies on a separation of variables between the columns and rows of a data matrix. In many situations, such as when analyzing misaligned data, this assumption breaks down, resulting in an artificial rank inflation.

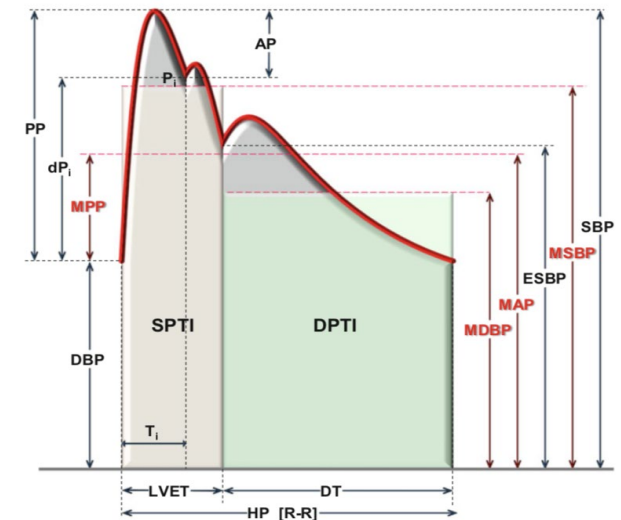
Example 1: Face Images



Example 2: ECG signals



Example 3: Pulse wave features



SVD data pre-processing: Data normalization

- SVD is sensitive to the average value of the input signals. When the derivations contain a nonzero average SVD algorithm decomposes these as orthogonal components which increases the rank of the data matrix and dimension of the signal space.
- Normalization, i.e. all signals scaled to unit variance, suppresses effects due to amplitude differences in the signals that could affect the singular values.

Conclusions

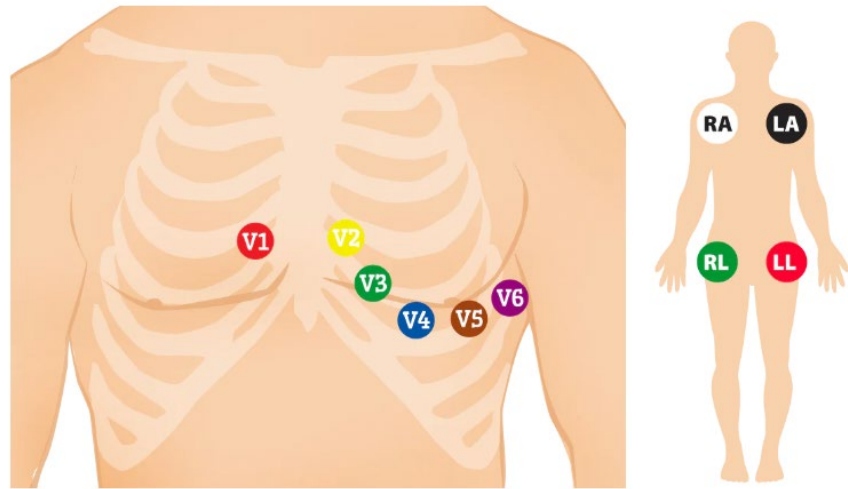
- SVD provides systematic approach to reduce dimensionality and resolve linear system of equations.
- Truncation, one of the most important decisions when using the SVD, is based on many factors, including specifications on the desired rank of the system, the magnitude of noise, and the distribution of the singular values.
- Singular spectrum analysis (based on SVD) provides a decomposition of a signal into bandpass components without prior knowledge.
- Data pre-processing, including alignment and normalization, is a key step to perform a SVD.

References

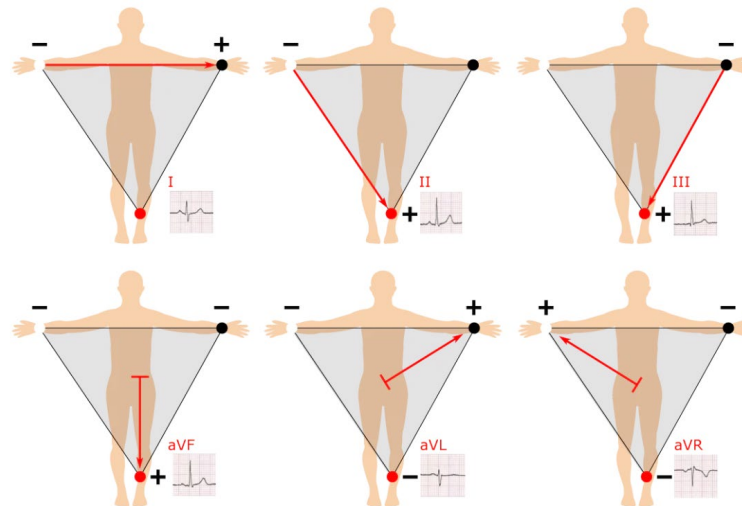
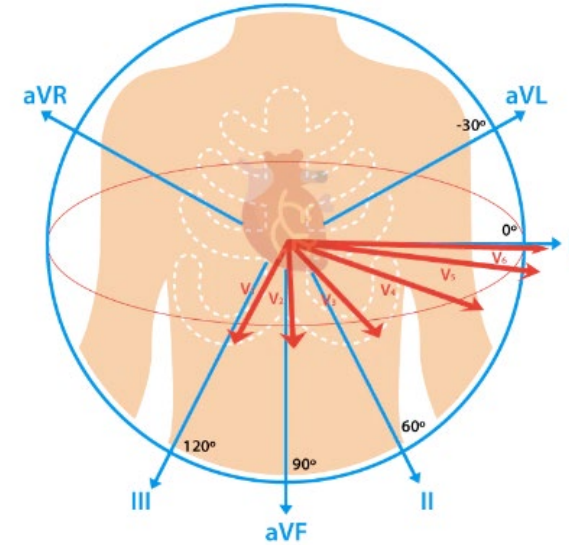
- G. H. Golub and C. F. Van Loan, Matrix Computations, John Hopkins Univ. Press, 3rd Ed., 1996.
- T.K. Moon and W.C. Stirling, Mathematical Methods and Algorithms for Signal Processing, Prentice Hall, 2000.
- N. Golyandina and A. Zhigljavski, Singular Spectrum Analysis for Time Series, Springer, 2013.
- Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control. Steven L. Brunton, J. Nathan Kutz
(https://www.youtube.com/watch?v=gXbThCXjZFM&list=PLMrJAKhleNNSVjnsvigIFoY2nXildDCcv&ab_channel=SteveBrunton, <http://databookuw.com/databook.pdf>)

Lab experiment 1: 12-leads ECG

10 electrodes placement of a 12-leads ECG



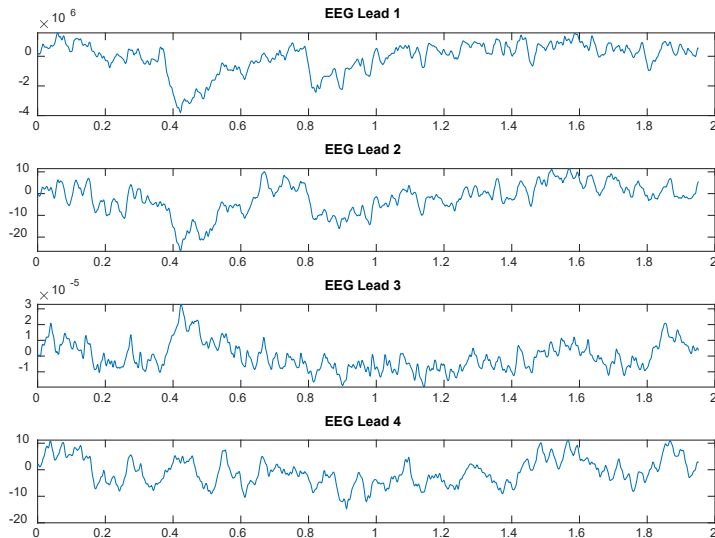
12-leads ECG



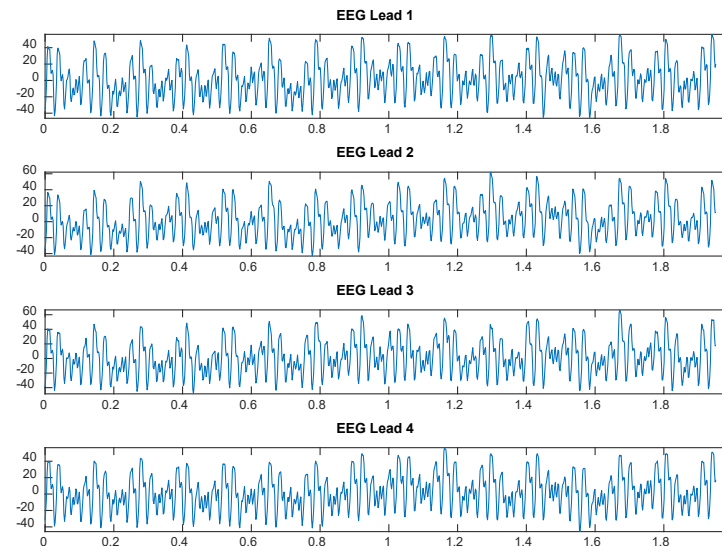
Lab experiment 2: Singular values and process complexity

EEG signals during brain stimulation for parkinson disease.

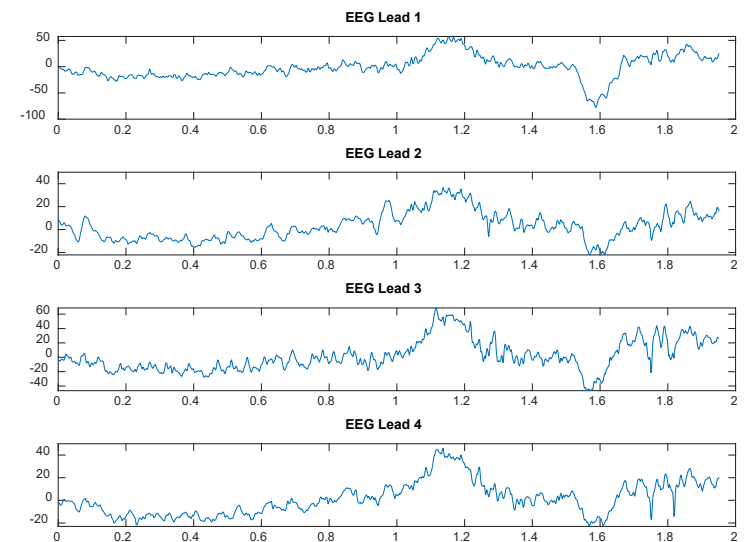
EEG before stimulation



EEG during stimulation



EEG after stimulation



Lab experiment 3: Drift cancellation and frequency component extraction

PPG and accelerometer signals from a running subject.

