Clustering Time Series

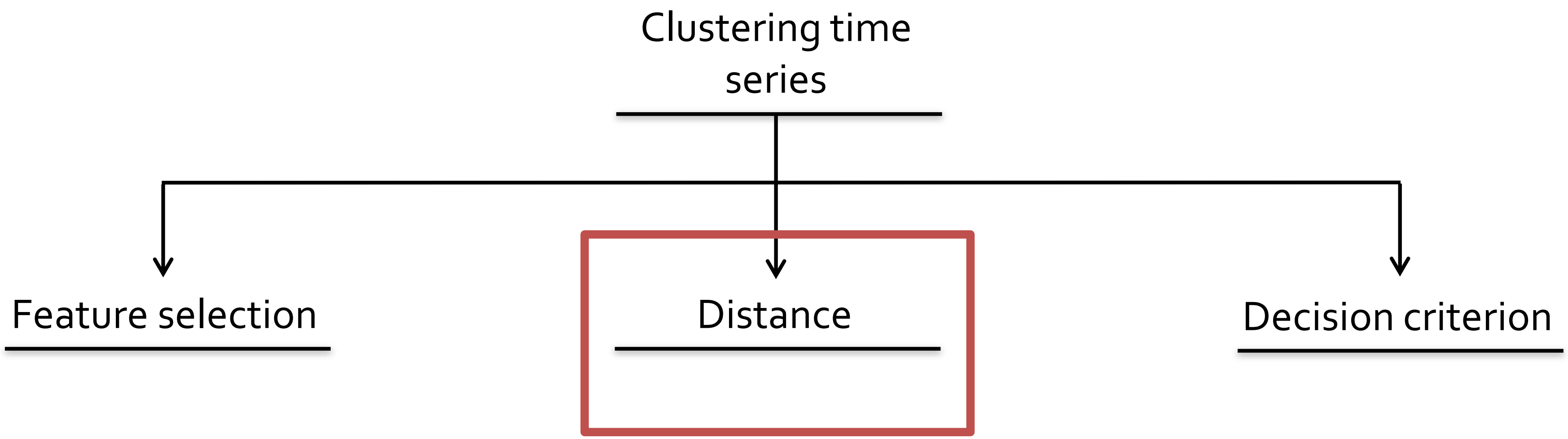
A homomorphic approach

Oliver Lauwers (oliver.lauwers@esat.kuleuven.be)
Bart De Moor (bart.demoor@esat.kuleuven.be)

Overview

Introduction

Problem



Unsupervised learning

Introduction

Problem

- 25000 smart meters
- Observe load data every 15 minutes
- Several years of data
- Seasonal effects obfuscate system dynamics

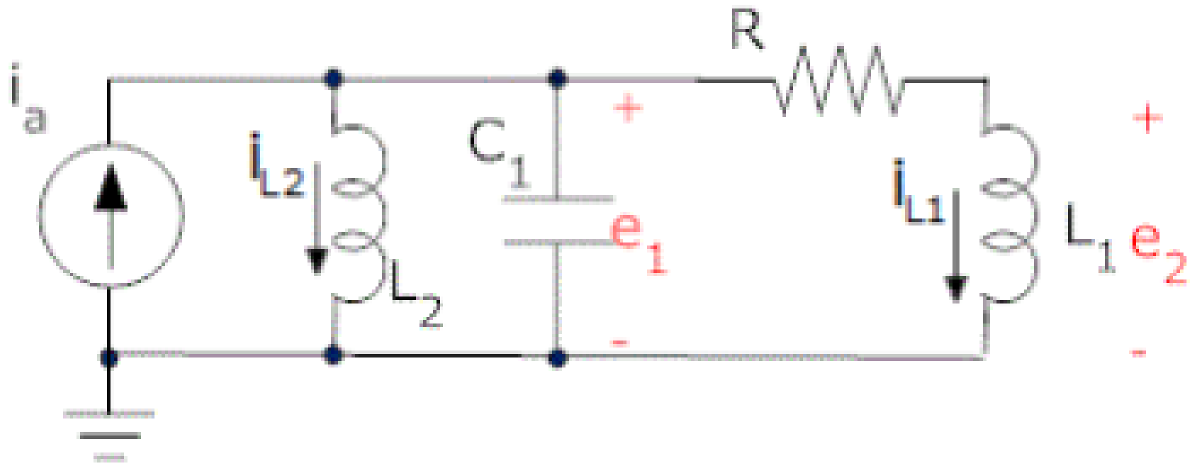
Introduction

Problem

Example

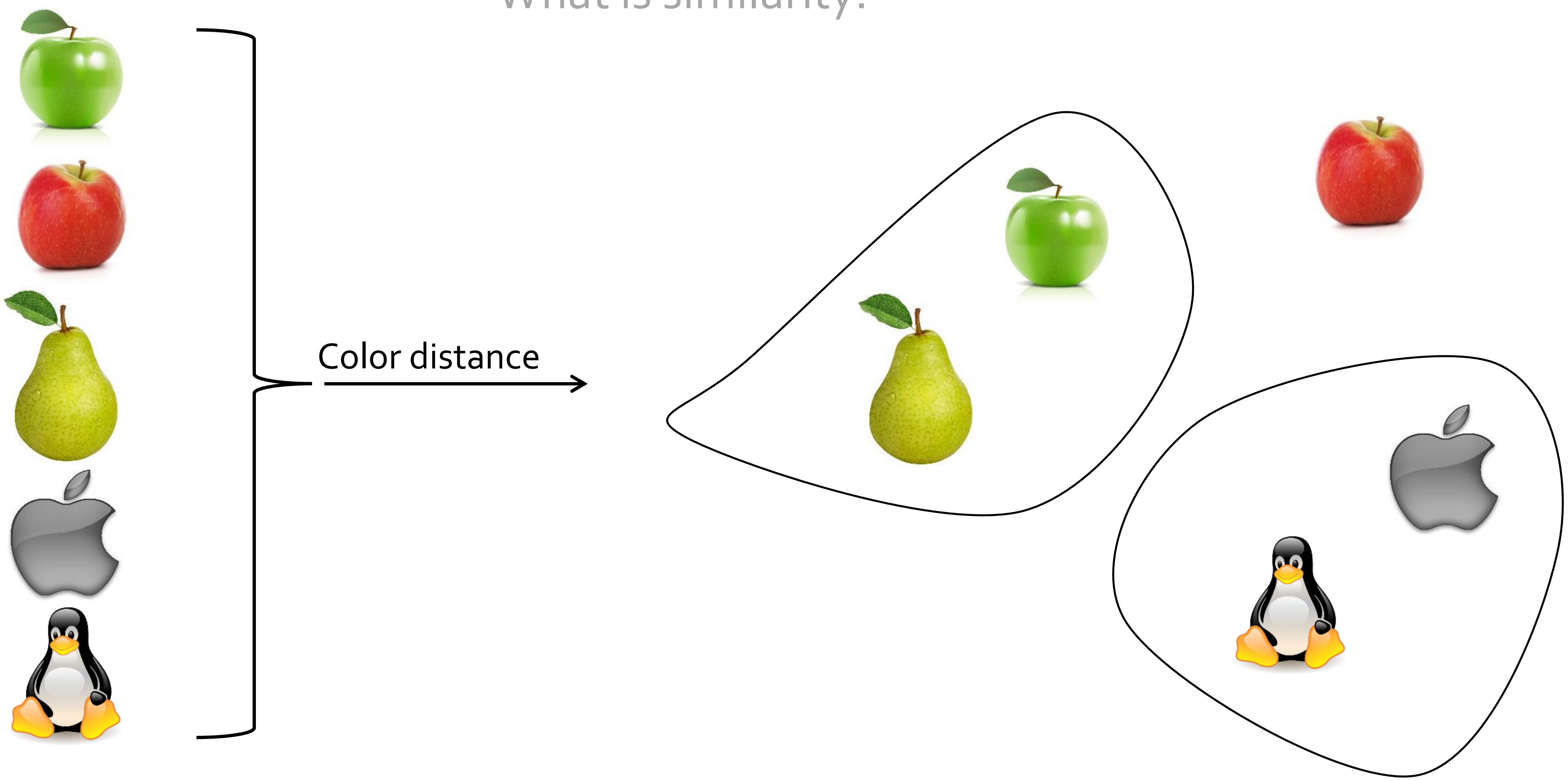
- 2 circuits
- 800 different inputs per circuit
- 2^16 timesteps

Objective:
Cluster input-output data coming from the same circuit together



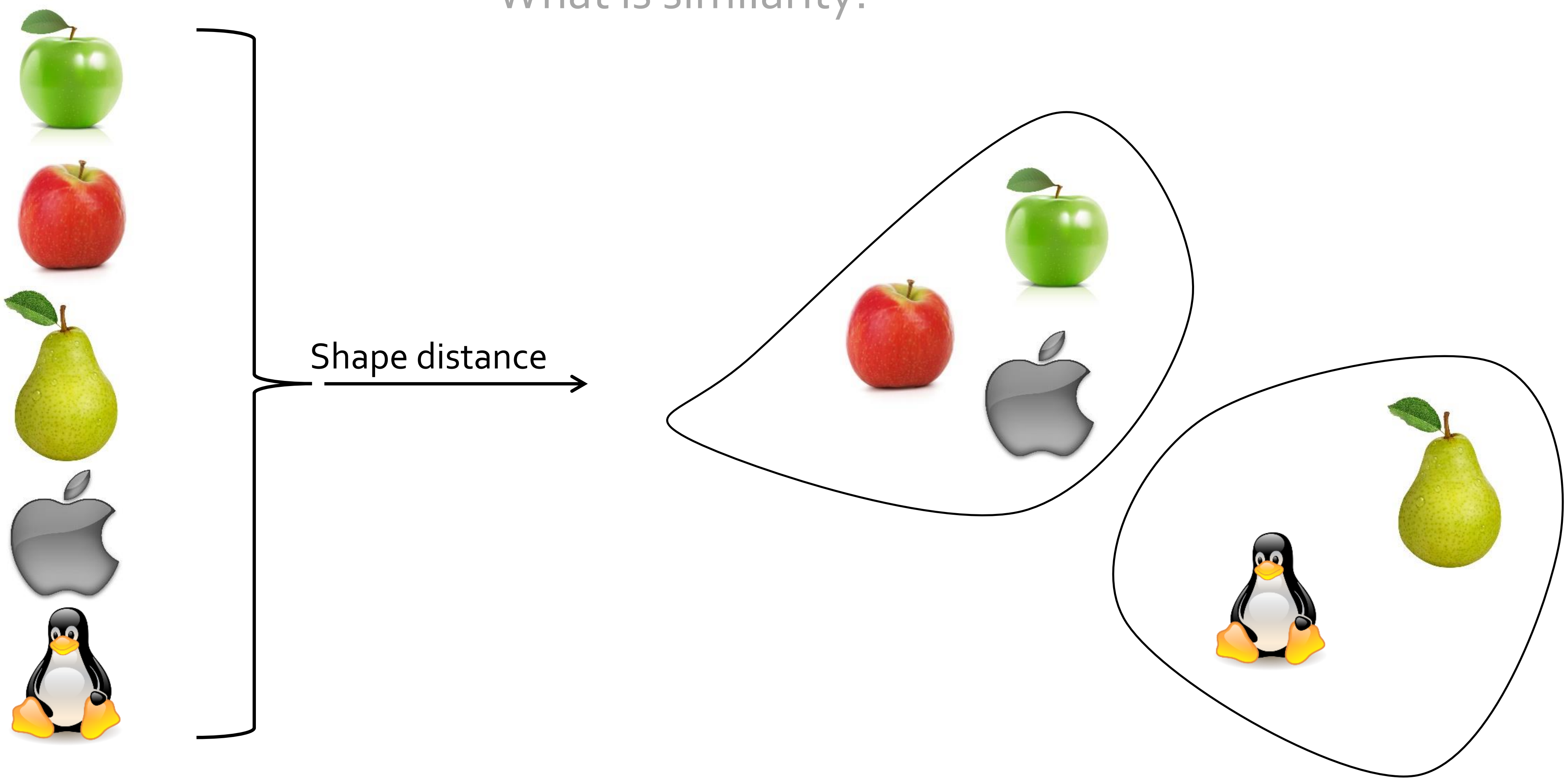
Introduction

What is similarity?



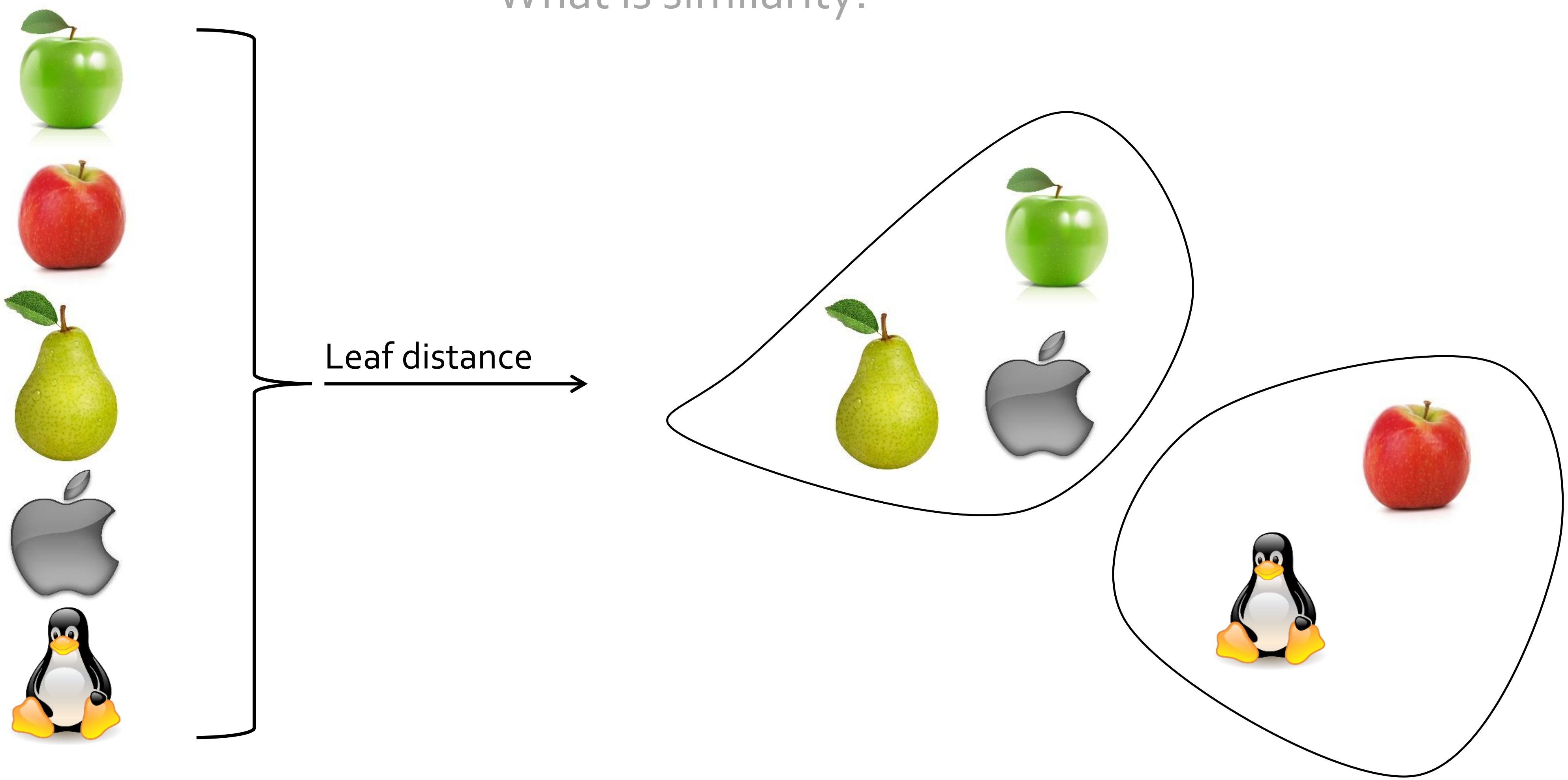
Introduction

What is similarity?



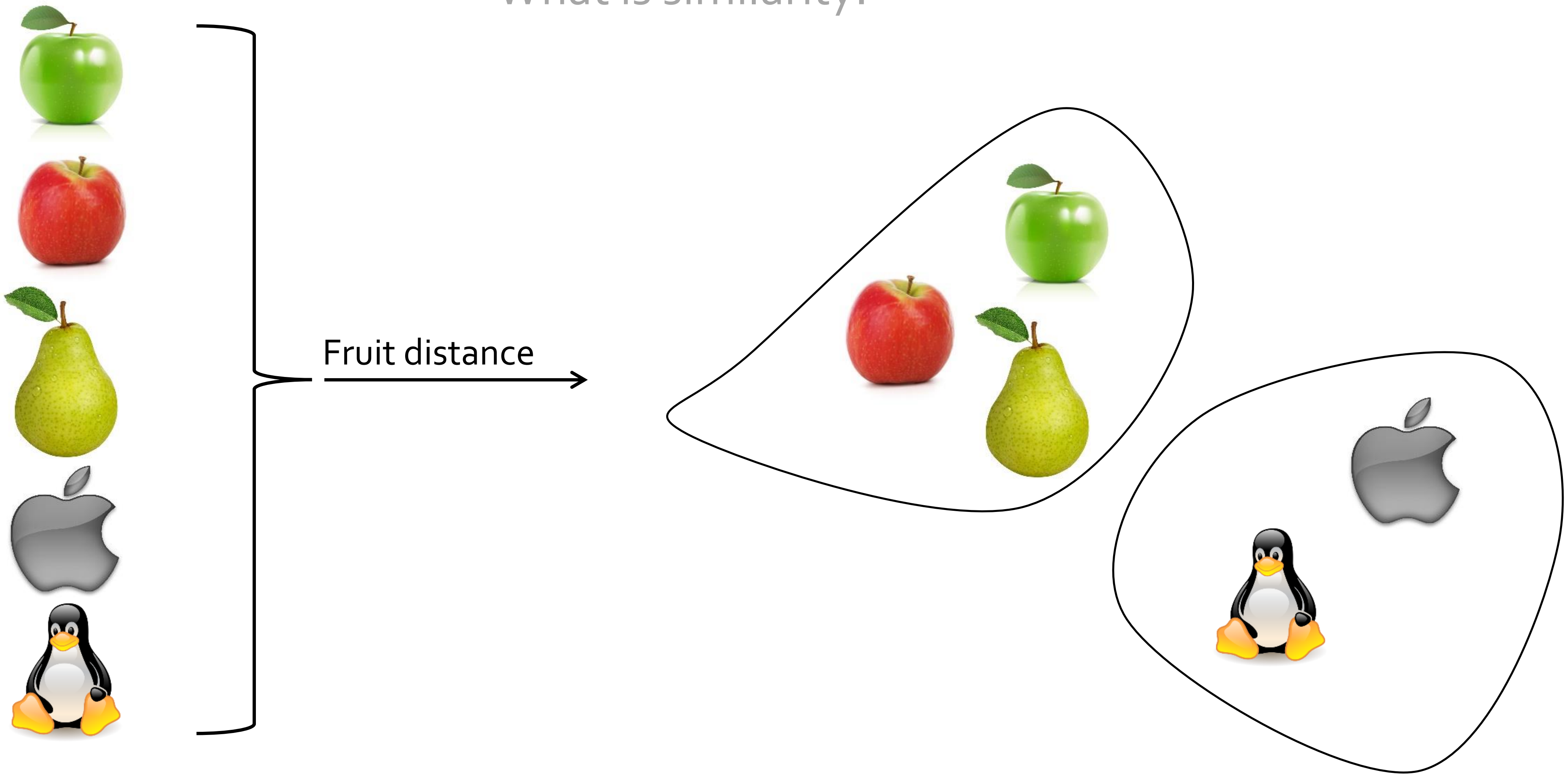
Introduction

What is similarity?



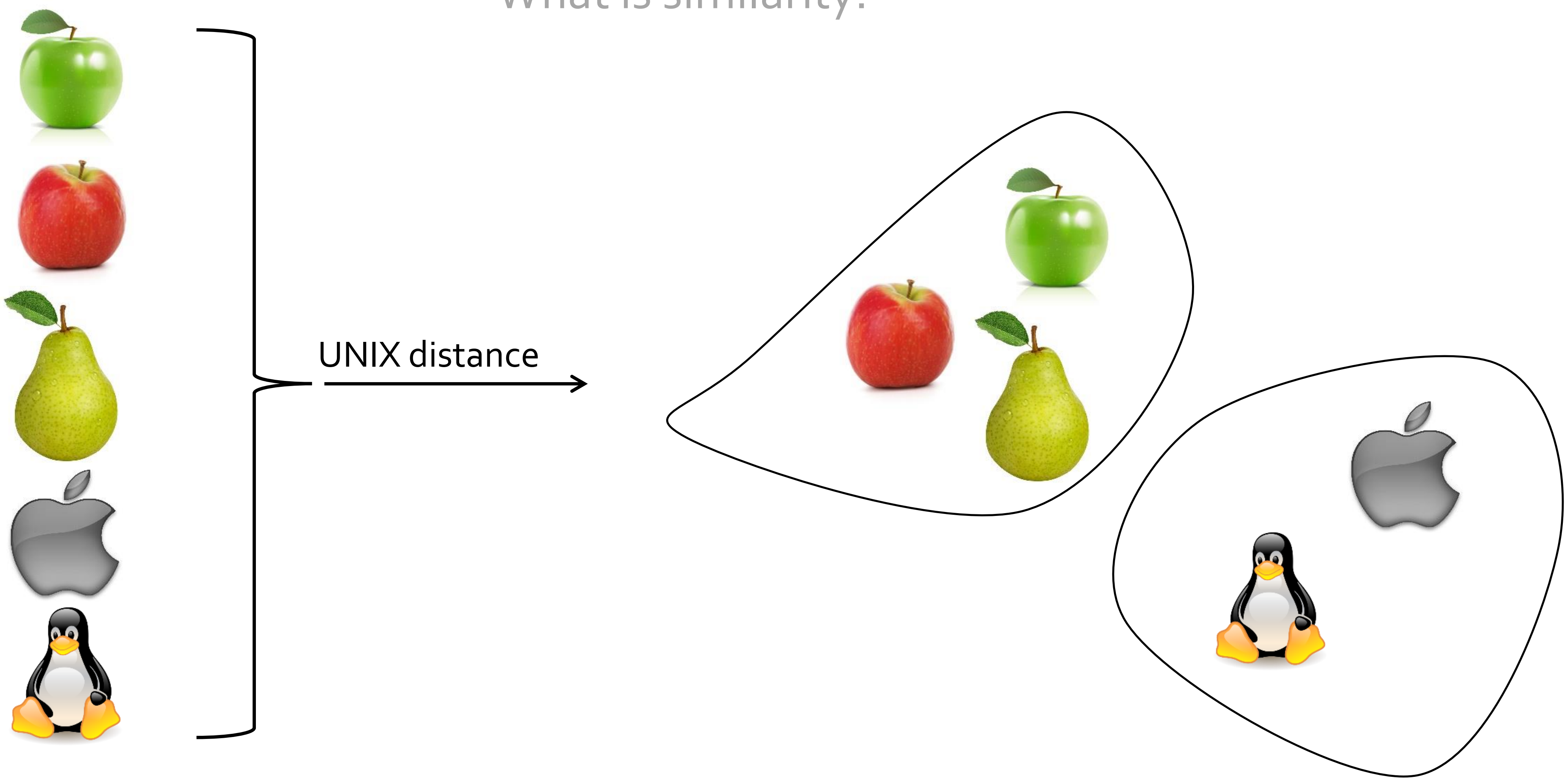
Introduction

What is similarity?



Introduction

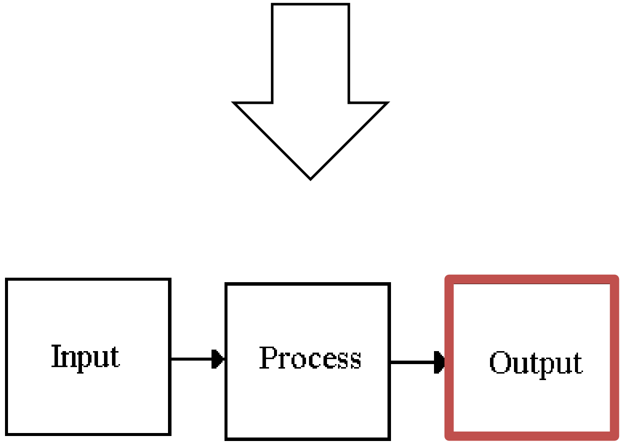
What is similarity?



Introduction

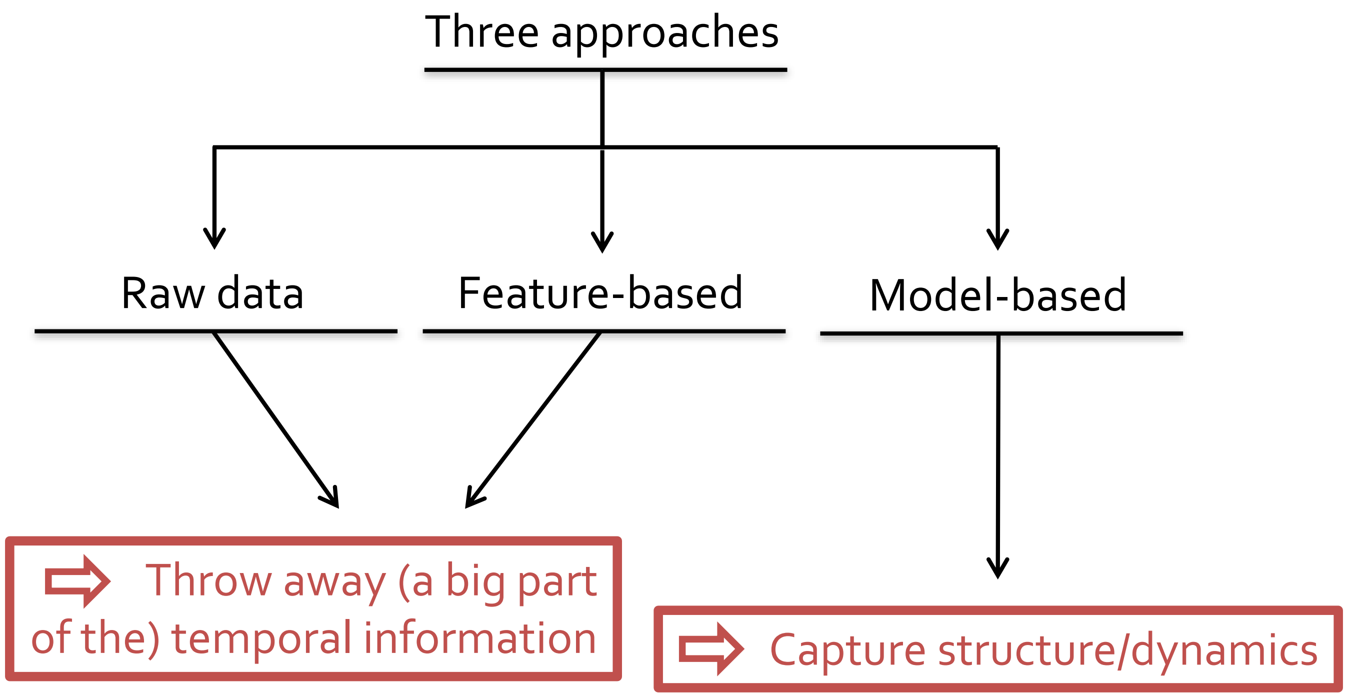
The time series state of the art

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{cases}$$



Introduction

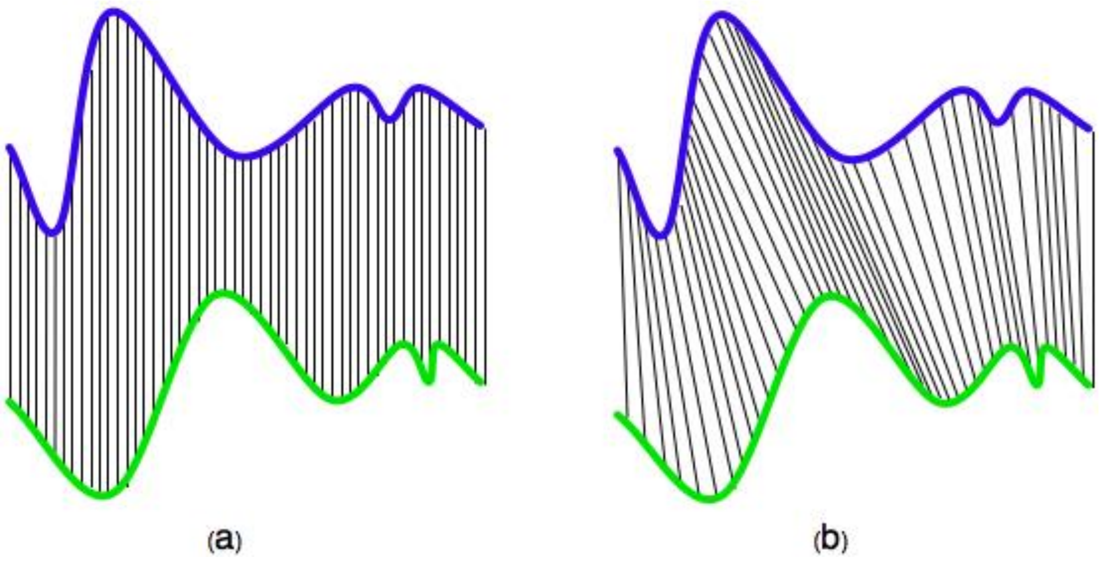
The time series state of the art



Liao, T. Warren. "Clustering of time series data—a survey." Pattern recognition 38.11 (2005): 1857-1874.

Introduction

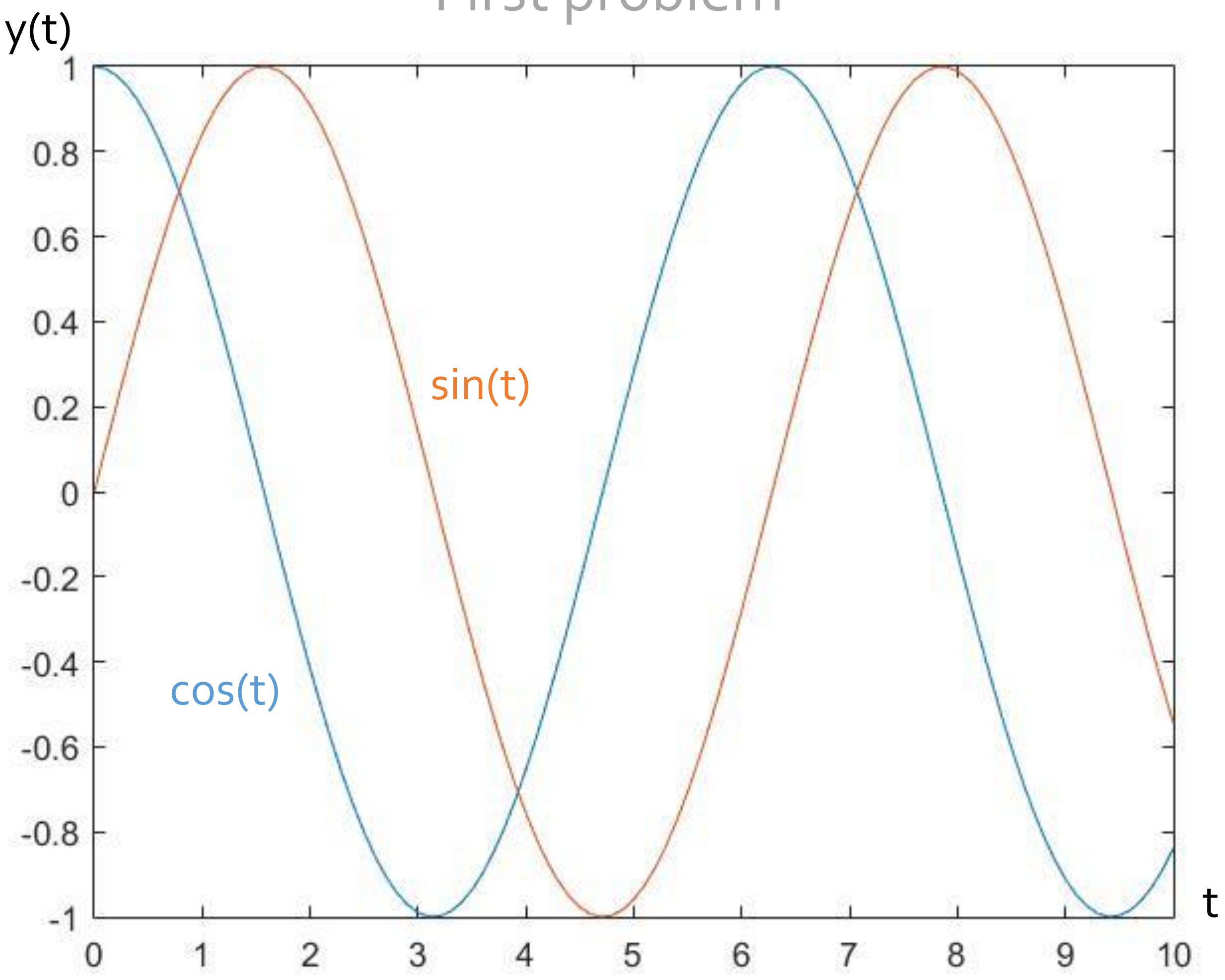
First problem



⇨ Temporal correlations?

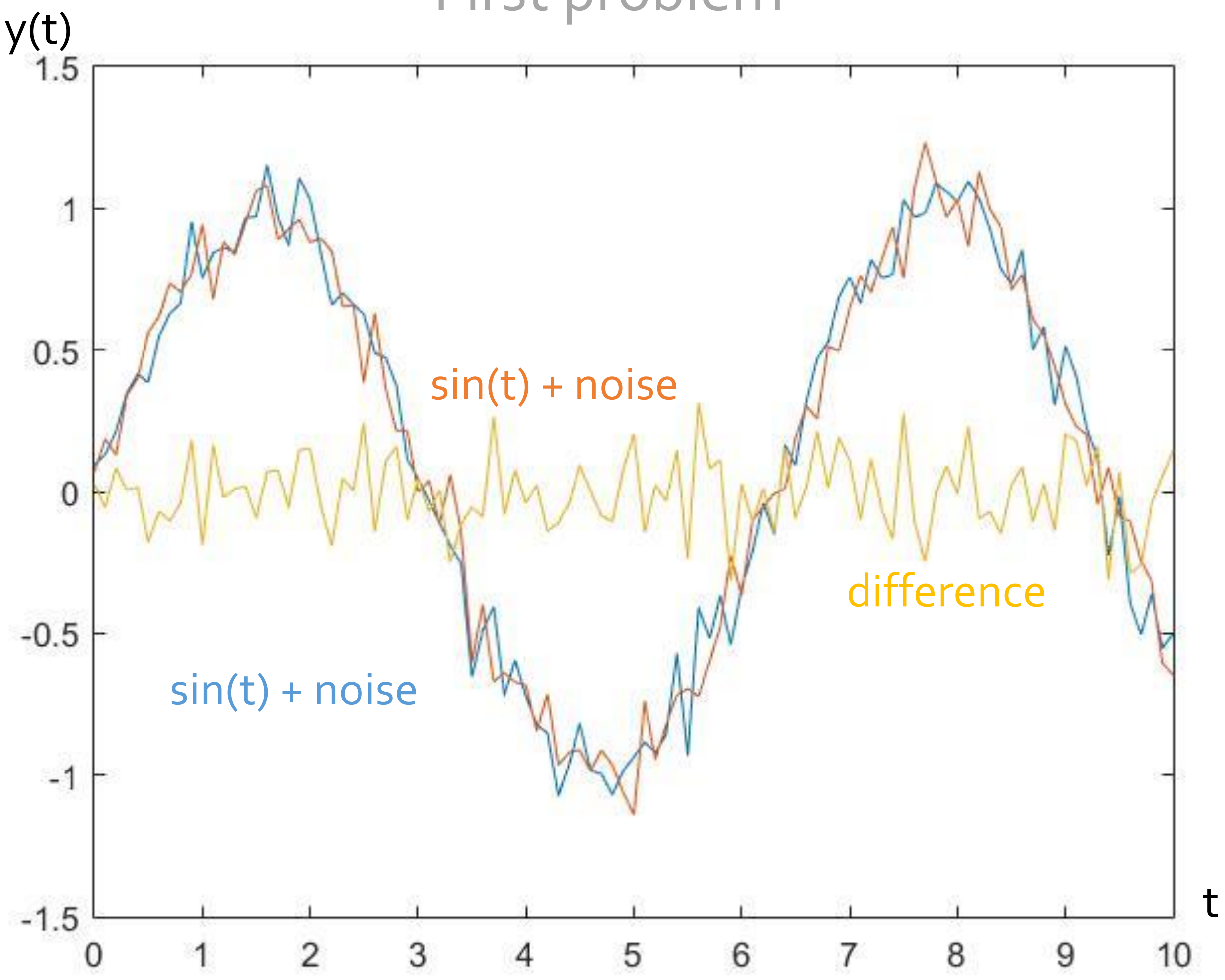
Introduction

First problem



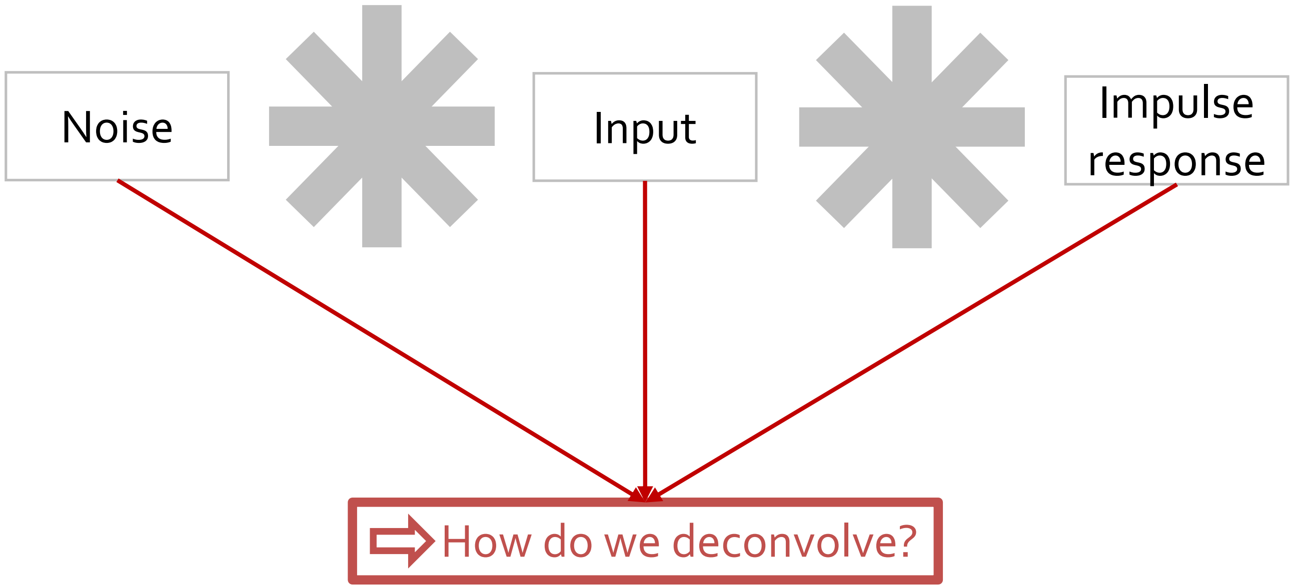
Introduction

First problem



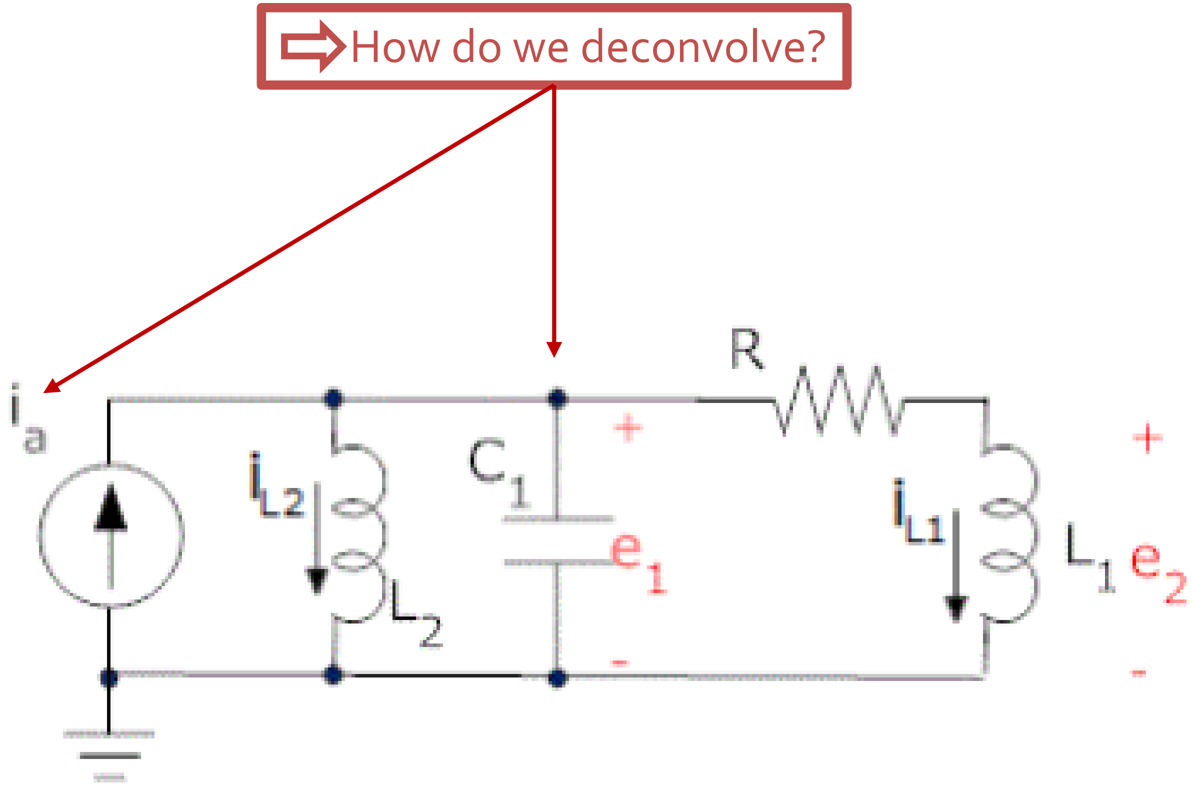
Introduction

Second problem



Introduction

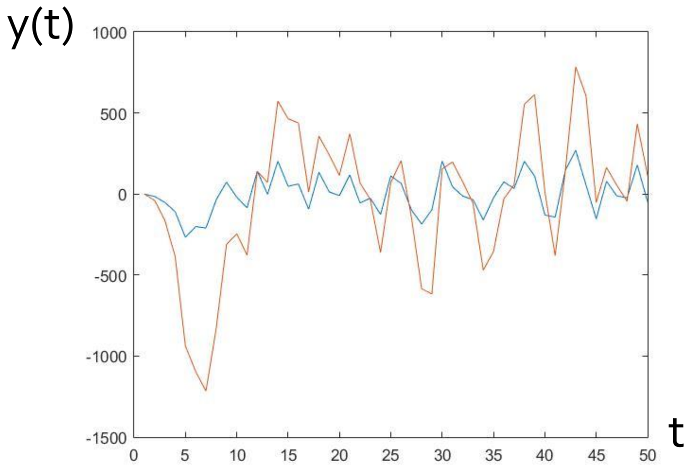
Second problem



http://lpsa.swarthmore.edu/Representations/SysRepSS.html

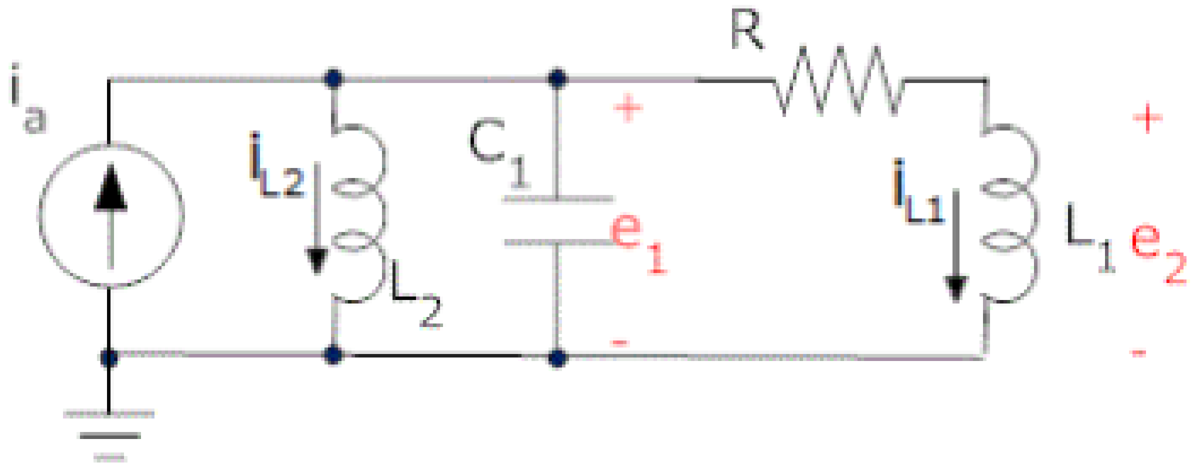
Introduction

Second problem



Same input,
different circuits

⇨ Input can dominate dynamics!

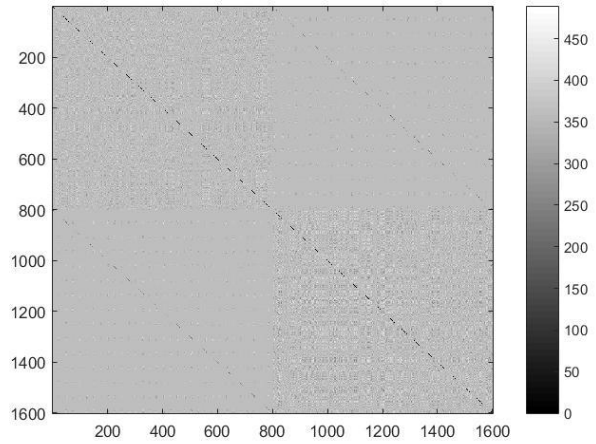


http://lpsa.swarthmore.edu/Representations/SysRepSS.html

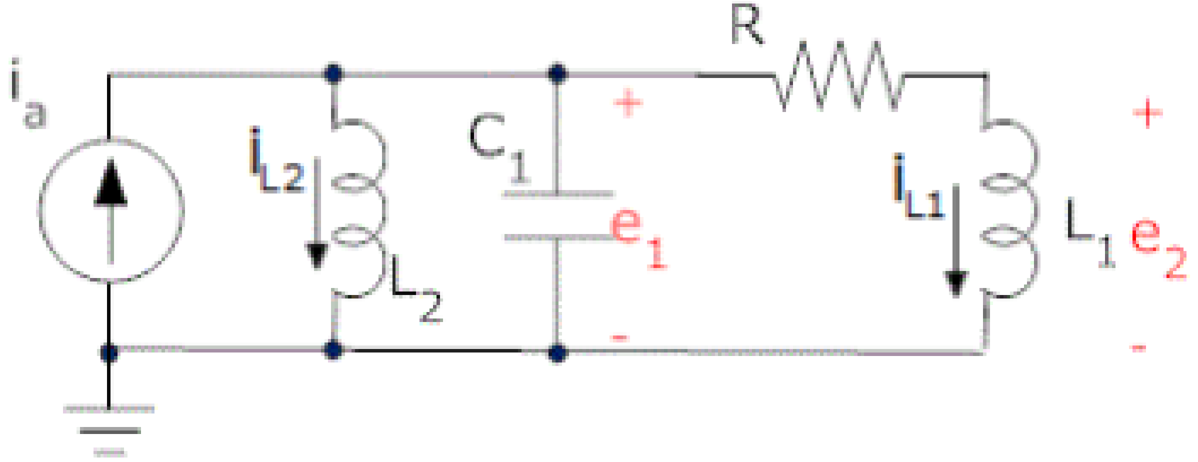
- 2 circuits
- 800 different inputs per circuit
- 2^16 timesteps

Introduction

The time series state of the art



⇨ Traditional techniques don't distinguish the models!



http://lpsa.swarthmore.edu/Representations/SysRepSS.html

Overview

A homomorphic approach

Homomorphic signal processing

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{cases}$$

$u_1(k) \blacksquare u_2(k)$

$c\colon u_1(k)$

$H(u_1(k)) \bullet H(u_2(k))$

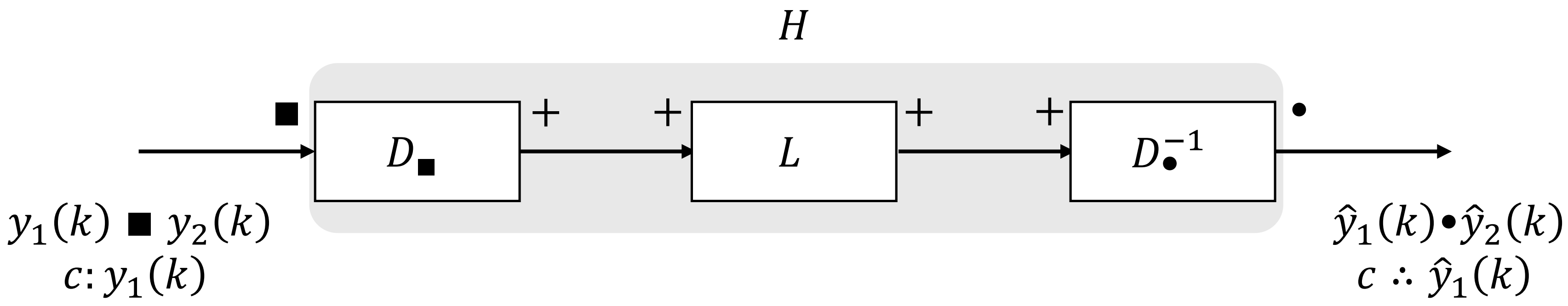
$c \therefore H(u_1(k))$

$y_1(k) \bullet y_2(k)$

$c \therefore y_1(k)$

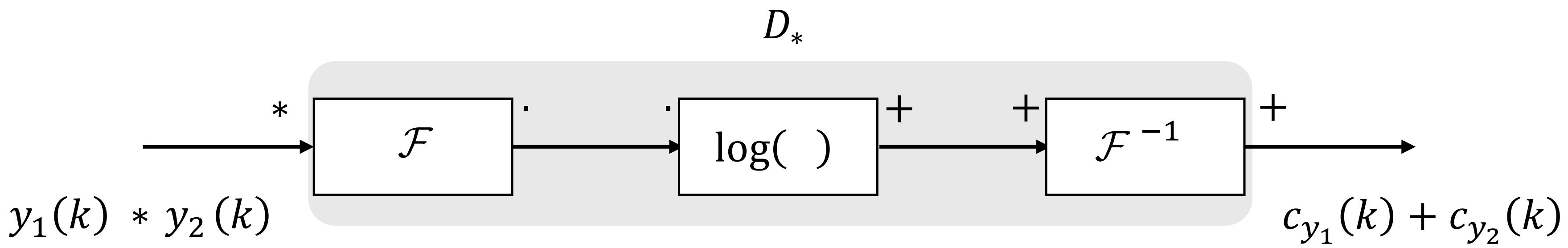
A homomorphic approach

Homomorphic signal processing



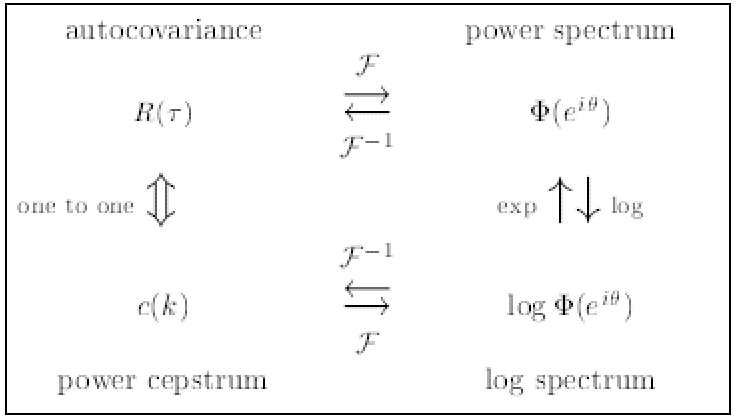
A homomorphic approach

The cepstrum



A homomorphic approach

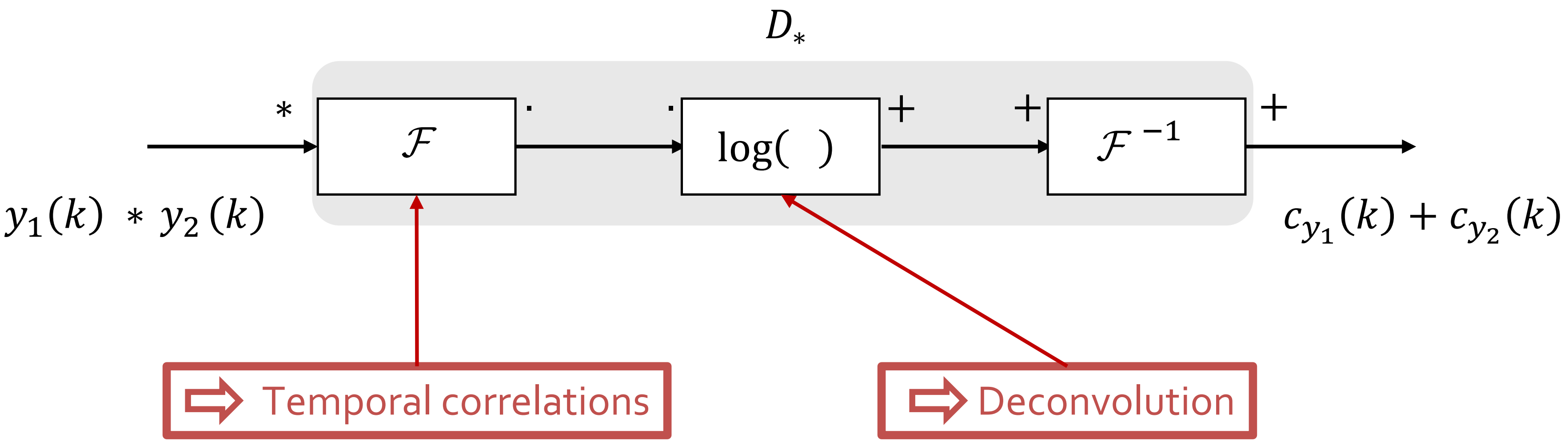
The cepstrum



De Cock, Katrien, and Bart De Moor. "Principal angles in system theory, information theory and signal processing." PhD thesis. 2002.

A homomorphic approach

The cepstrum



Overview

Introduction
- Problem
- What is similarity?
- The time series state of the art
- First problem
- Second problem

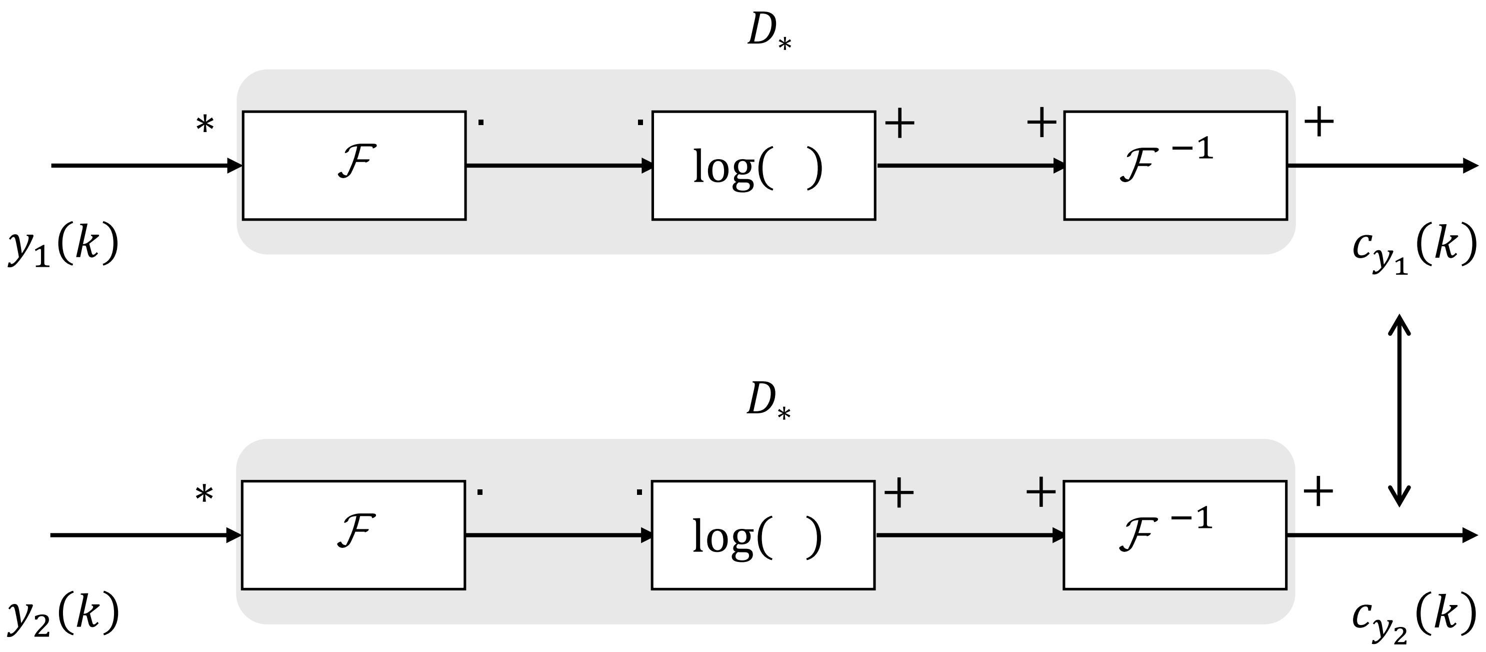
A homomorphic approach
- Homomorphic Signal Processing
- The cepstrum

Cepstrum distance
- Martin distance
- Stochastic models
- Deterministic models
- Deterministic extension option 1
- Deterministic extension option 2

Conclusion
- Summary of the extended distance
- Future research

Cepstrum distance

Martin distance



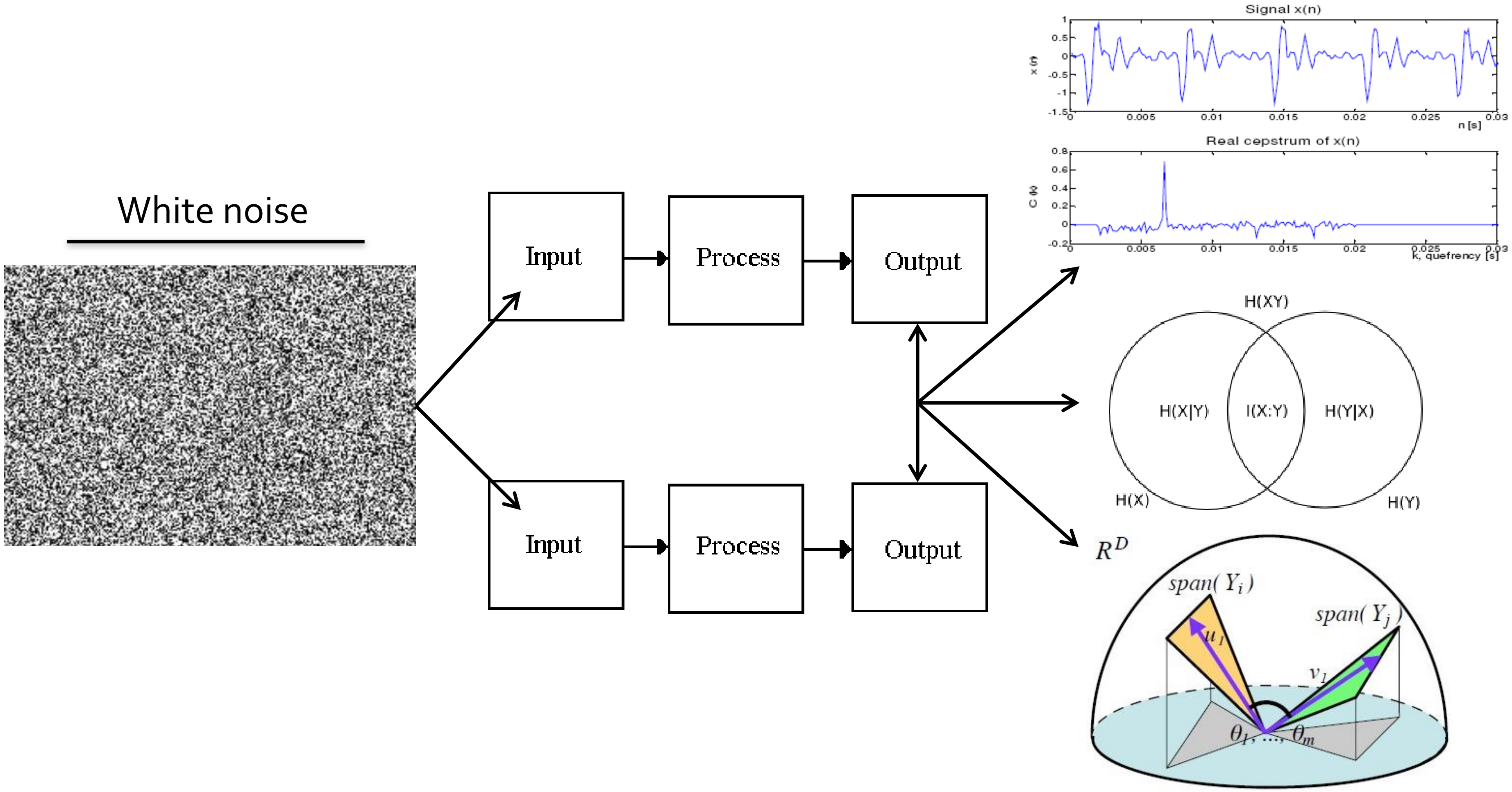
$$d(y_1, y_2) = \sum_{k=1}^{\infty} k\left(c_{y_1}(k) - c_{y_2}(k)\right)^2$$

Martin, Richard J. "A metric for ARMA processes." IEEE Transactions on Signal Processing 48.4 (2000): 1164-1170.

Cepstrum distance

Stochastic models

White noise



Cepstrum

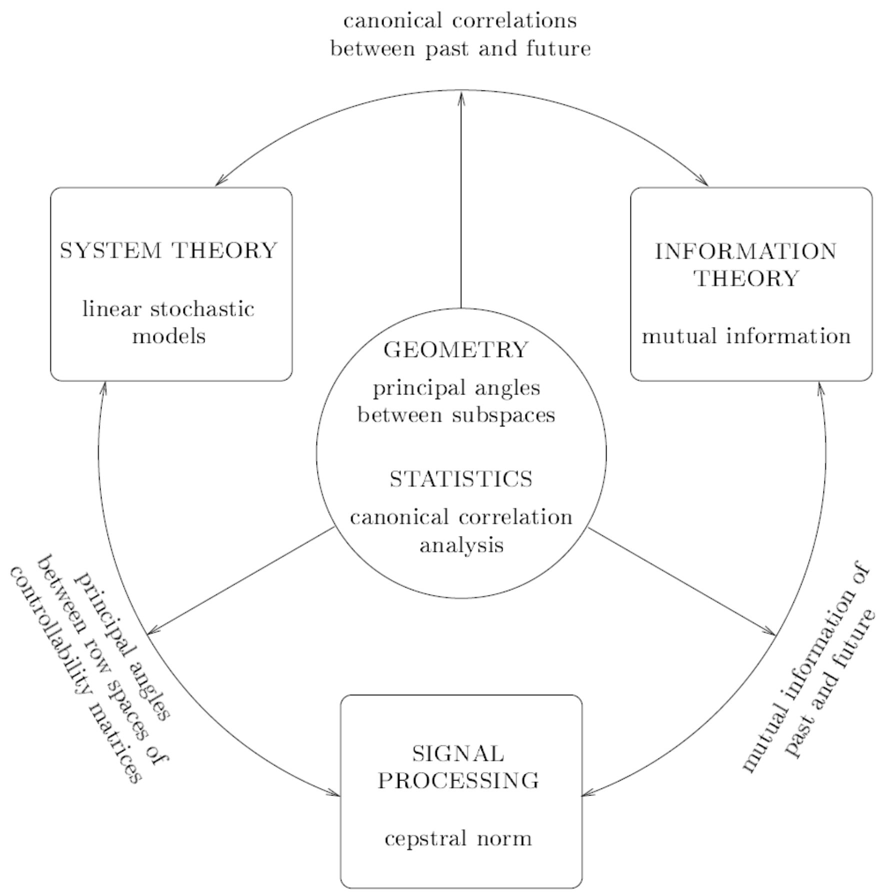
Mutual information

Subspace angles

De Cock, Katrien, and Bart De Moor. "Subspace angles and distances between ARMA models."
Proc. of the Intl. Symp. of Math. Theory of networks and systems. Vol. 1. 2000.

Cepstrum distance

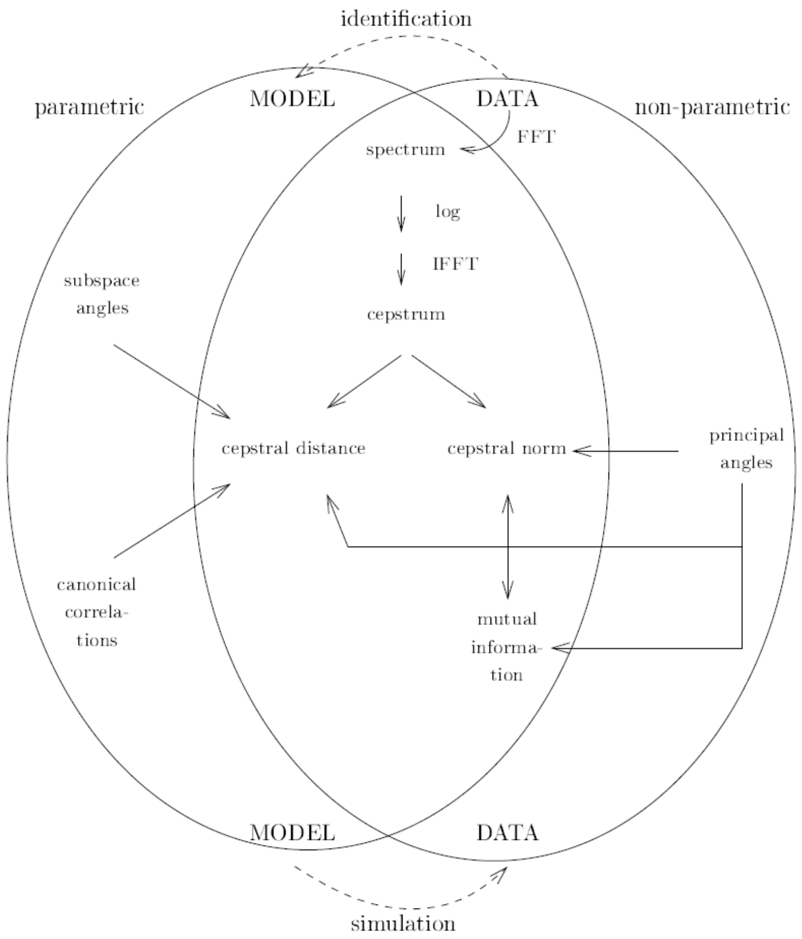
Stochastic models



De Cock, Katrien, and Bart De Moor. "Principal angles in system theory, information theory and signal processing." PhD thesis. 2002.

Cepstrum distance

Stochastic models



De Cock, Katrien, and Bart De Moor. "Principal angles in system theory, information theory and signal processing." PhD thesis. 2002.

Cepstrum distance

Stochastic models

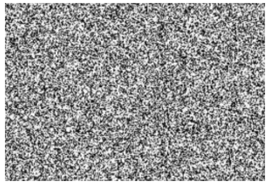
We can calculate distances

- Between time series
  - Calculate distance matrix explicitly
- Between models
  - Model and calculate distances based on poles/zeros
- Between time series and models
  - Put time series through inverse model

De Cock, Katrien, and Bart De Moor. "Principal angles in system theory, information theory and signal processing." PhD thesis. 2002.

Cepstrum distance

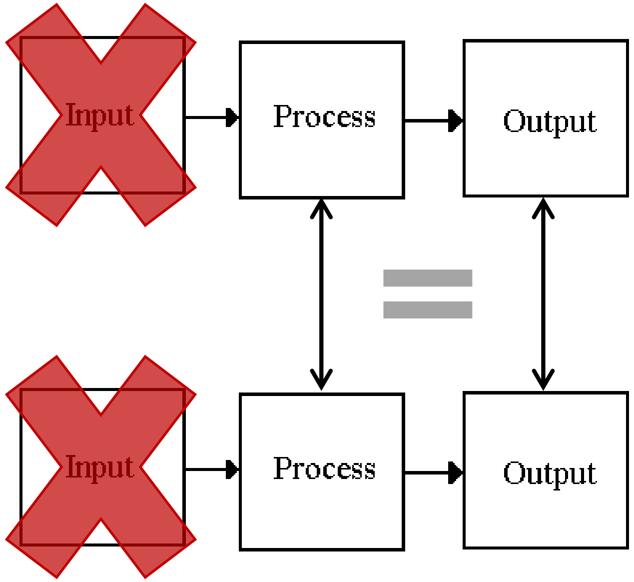
Stochastic models

$E($  $, \text{State}) = 0$

⇨ Contains all the relevant information!

Cepstrum distance

Stochastic models



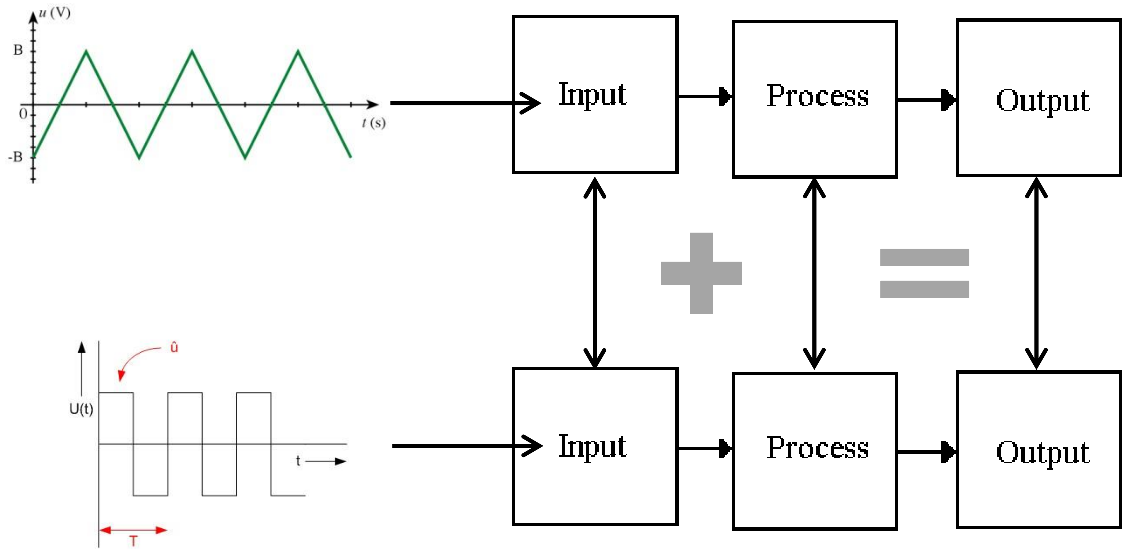
Cepstrum of output equals (for $k > 0$) the cepstrum of the transfer function
The cepstrum can be written in terms of poles (α_i) and zeros (β_i):

$$c_y(k) = \sum_{i=1}^{p} \frac{\alpha_i^k}{k} - \sum_{i=1}^{q} \frac{\beta_i^k}{k} \qquad (\forall k > 0).$$

De Cock, Katrien, and Bart De Moor. "Subspace angles and distances between ARMA models."
Proc. of the Intl. Symp. of Math. Theory of networks and systems. Vol. 1. 2000.

Cepstrum distance

Deterministic models

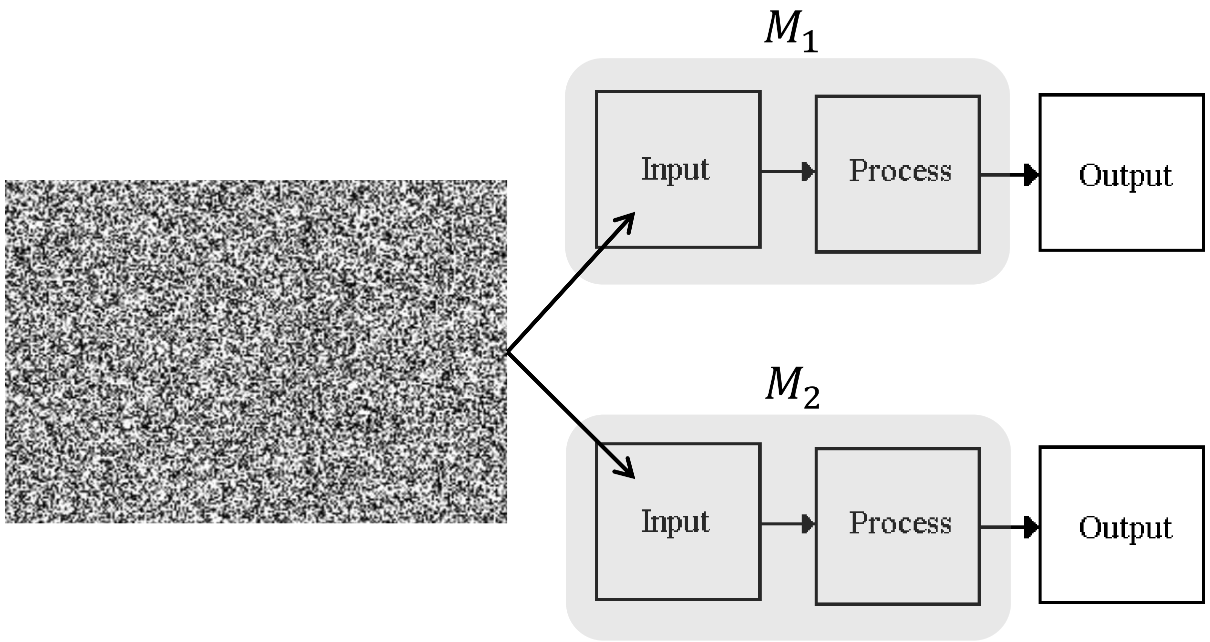


⇨ Convolution becomes addition

Cepstrum distance

Deterministic distance option 1

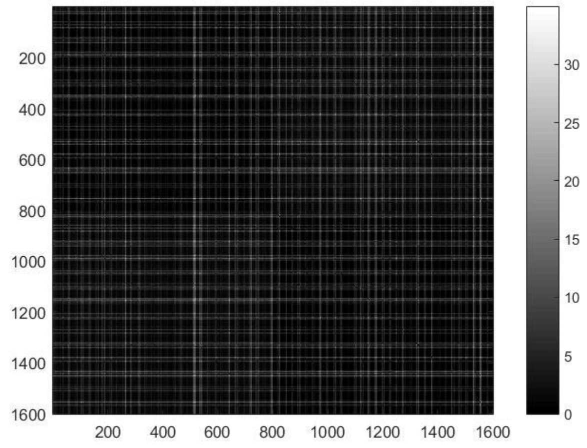
Model the input as part of the dynamics



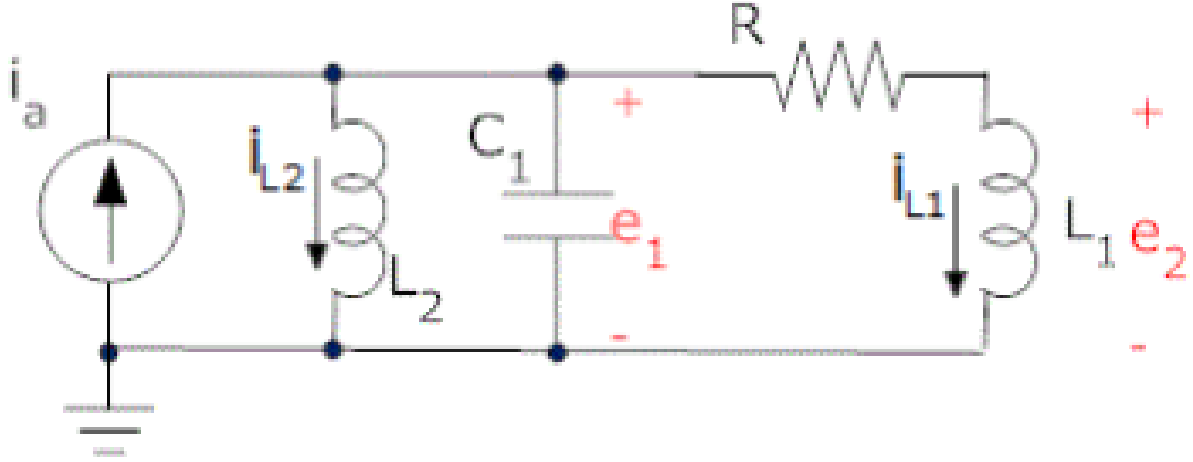
- 2 circuits
- 800 different inputs per circuit
- 2^16 timesteps

Deterministic distance

Deterministic distance option 1



⇨ Input can dominate dynamics!

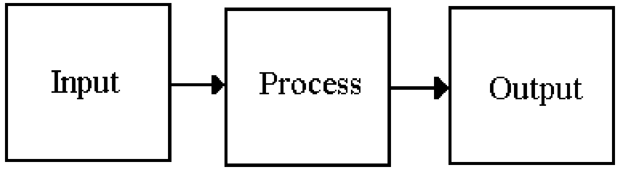


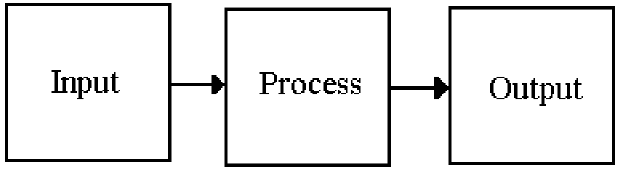
http://lpsa.swarthmore.edu/Representations/SysRepSS.html

Cepstrum distance

Deterministic distance option 2

Look only at the process

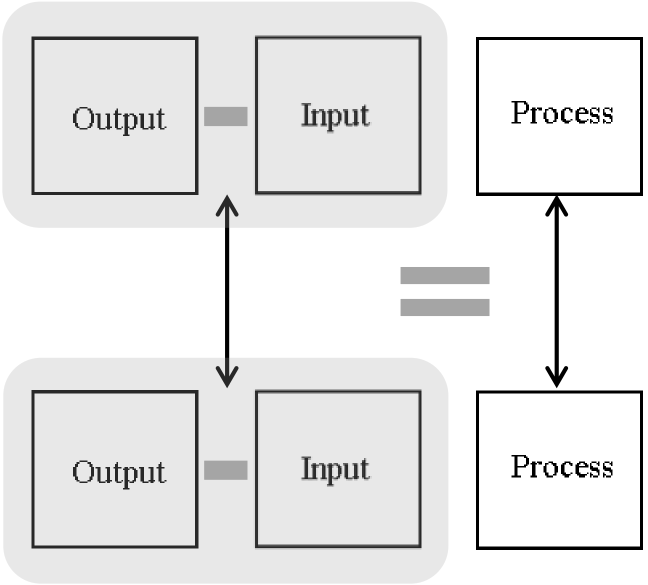




Cepstrum distance

Deterministic distance option 2

Look only at the process

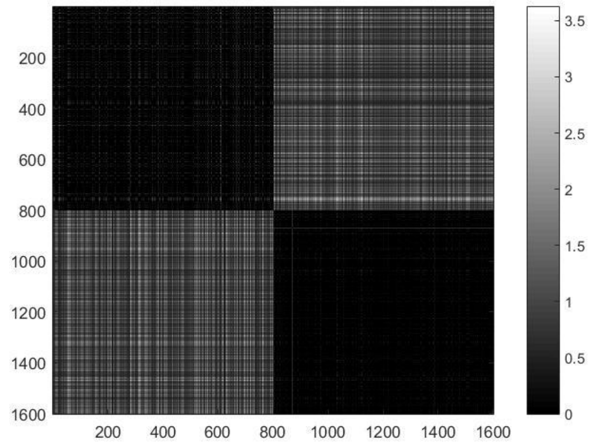


⇨ The transfer function cepstrum is obtained by simply subtracting the input cepstrum from the output cepstrum

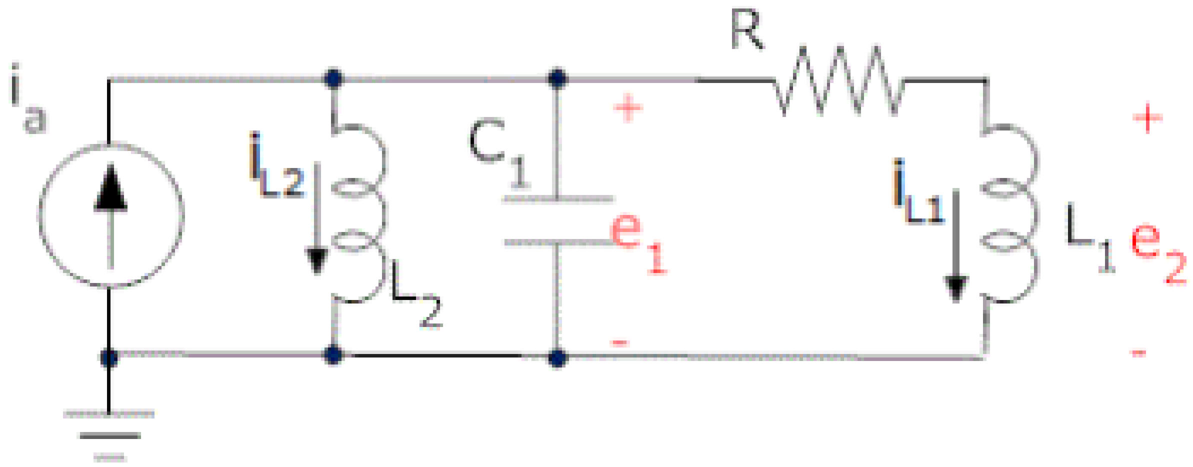
- 2 circuits
- 800 different inputs per circuit
- 2^16 timesteps

Cepstrum distance

Deterministic distance option 2



⇨ This extension works in practice

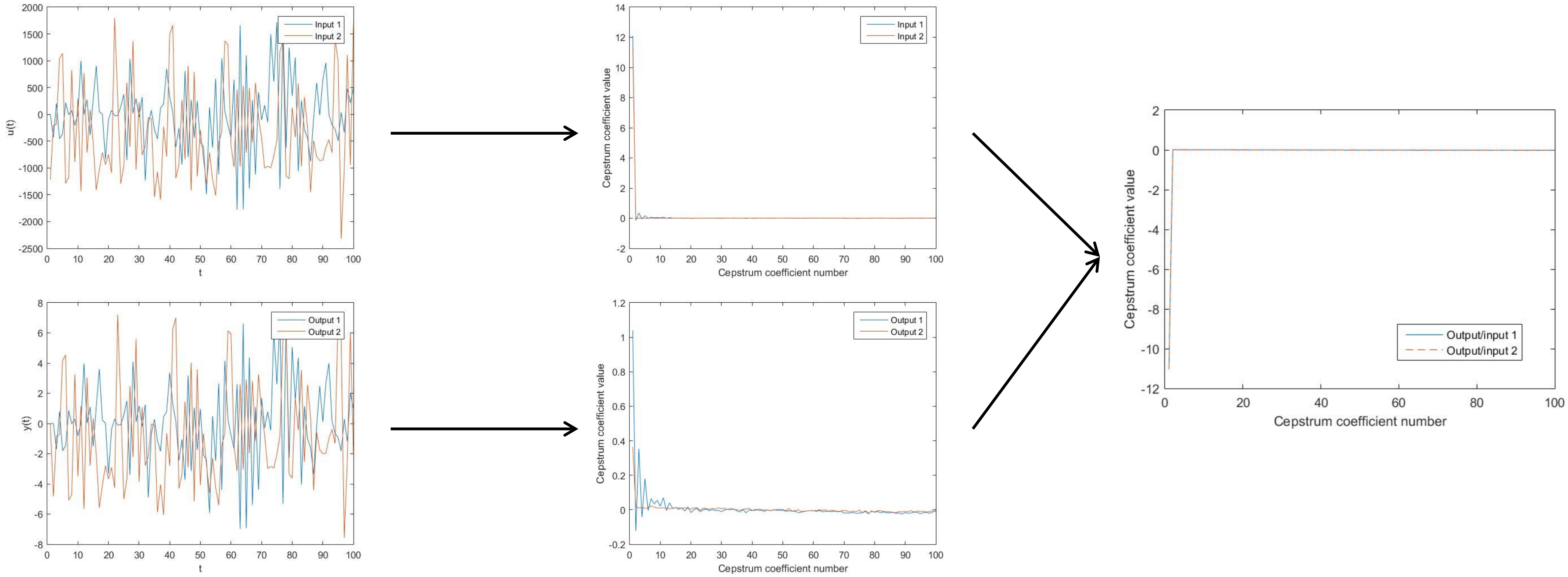


http://lpsa.swarthmore.edu/Representations/SysRepSS.html

Overview

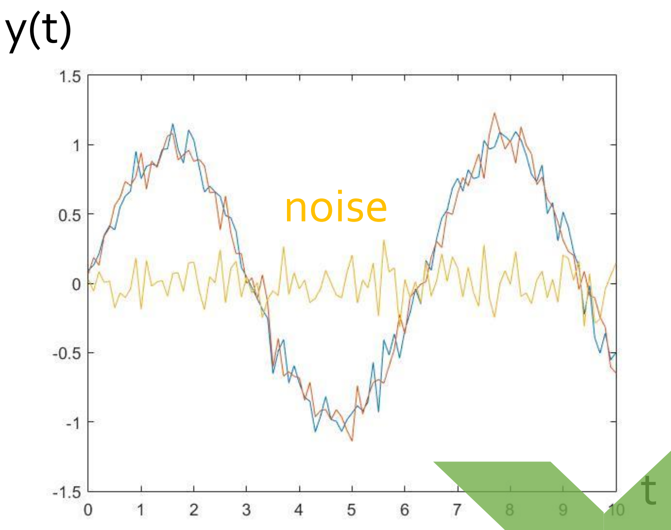
Conclusion

Summary of the extended distance

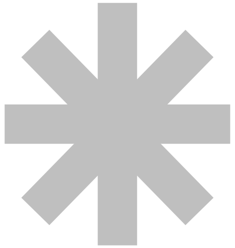


Conclusion

Summary of the extended distance



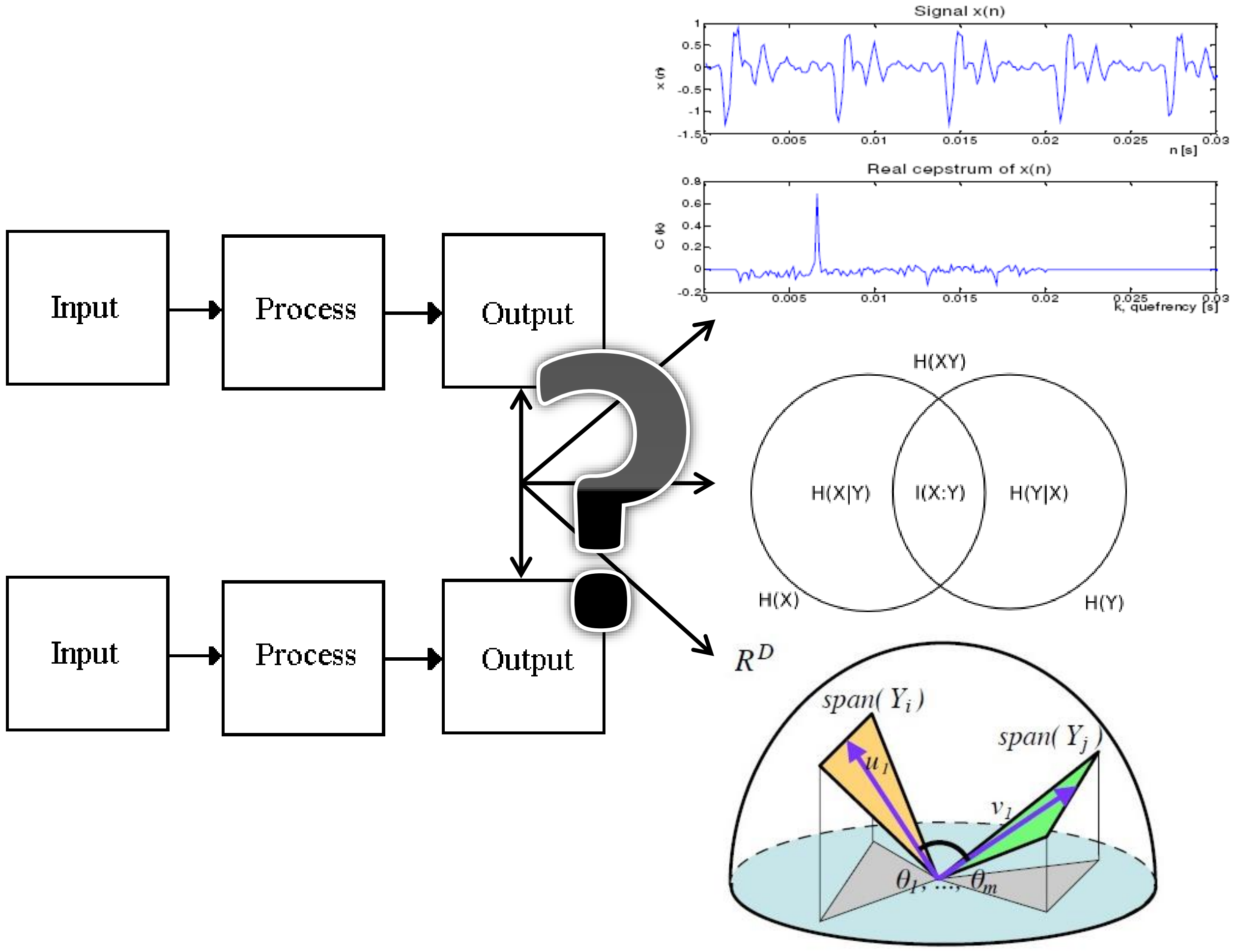
⇨ Temporal correlations?



⇨ How do we deconvolve?

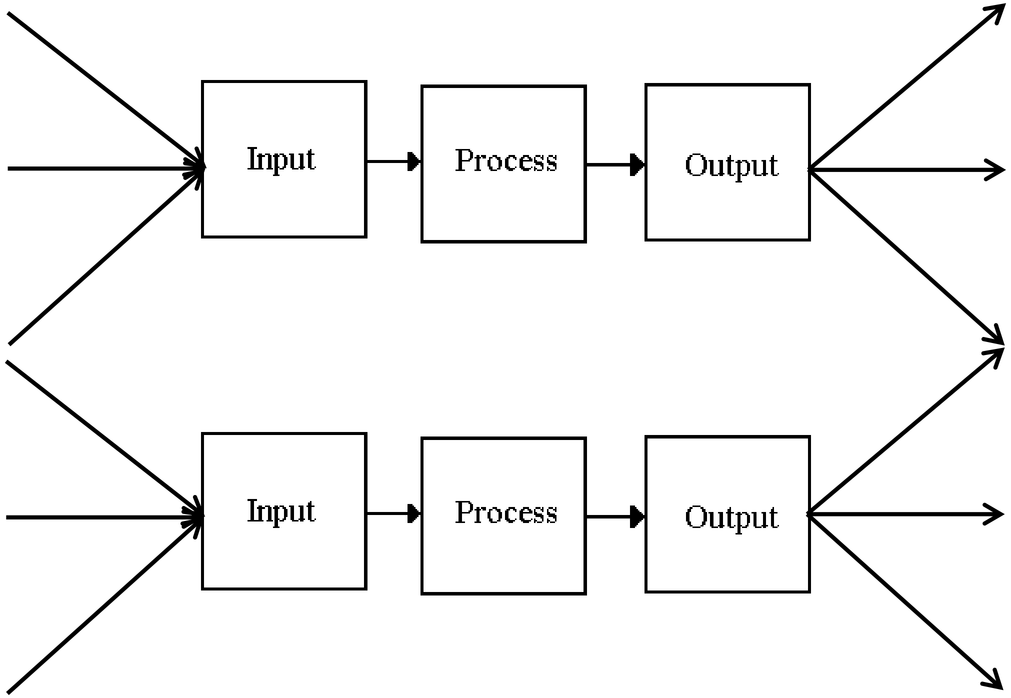
Conclusion

Further research?

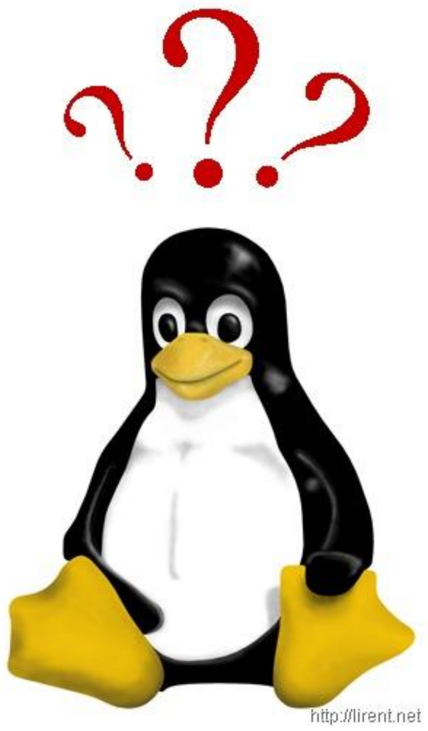


Conclusion

Further research?



Thank y🍎u f🍎r y🍎ur attenti🍎n



Extra Slides

Cepstrum of output equals (for $k > 0$) the cepstrum of the transfer function.
The cepstrum can be written in terms of poles (α_i) and zeros (β_i):

$$c_y(k) = \sum_{i=1}^{p} \frac{\alpha_i^k}{k} - \sum_{i=1}^{q} \frac{\beta_i^k}{k} \qquad (\forall k > 0).$$

We can then define a norm:

$$||\log H||^2 = \sum_{k=1}^{\infty} k c_y(k)^2$$

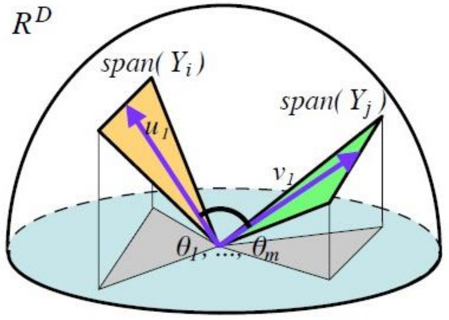
We can fill in:

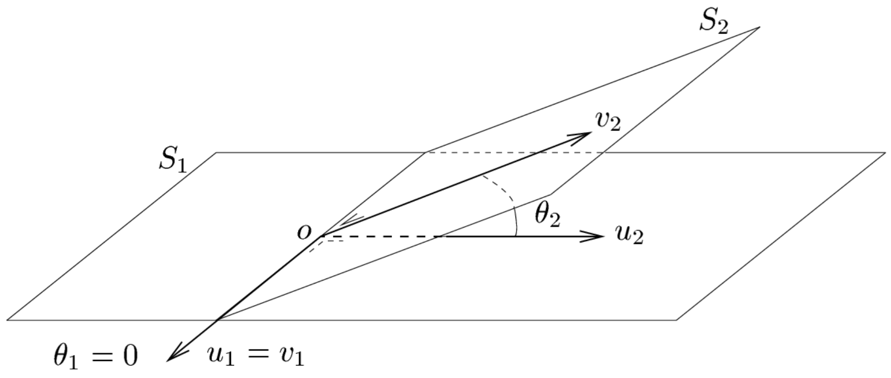
$$||\log H||^2 = \sum_{k=1}^{\infty} k \left(\sum_{i=1}^{p} \frac{\alpha_i^k}{k} - \sum_{i=1}^{q} \frac{\beta_i^k}{k} \right)^2$$

$$\sum_{k=1}^{\infty} \frac{x^k}{k} = -\log(1-x) \qquad (\forall\ |x| > 0)$$

$$||\log H||^2 = \frac{\prod_{i=1}^{p} \prod_{j=1}^{q} |1-\alpha_i \overline{\beta}_i|^2}{\prod_{i,j=1}^{p} (1-\alpha_i \overline{\alpha}_j)^2 \prod_{i,j=1}^{q} (1-\beta_i \overline{\beta}_j)^2}$$

Extra Slides





$$\cos[u_i < v_i] = \frac{|u_i v_i^T|}{\|u_i\| \|v_i\|}$$

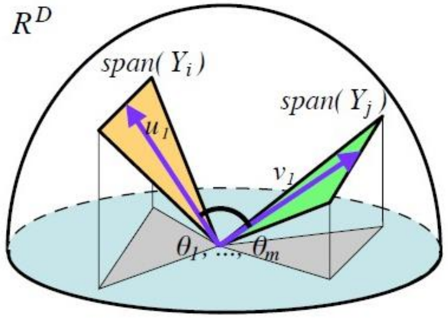
$$\rho_{xy} = \frac{E\{XY\}}{\sqrt{E\{X^2\}}\sqrt{E\{Y^2\}}} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \lim_{j \to \infty} \cos\left[x^{(j)} < y^{(j)}\right]$$

De Cock, Katrien, and De Moor, Bart. "Subspace angles and distances between ARMA models."
Proc. of the Intl. Symp. of Math. Theory of networks and systems. Vol. 1. 2000.

Extra Slides

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{cases}$$

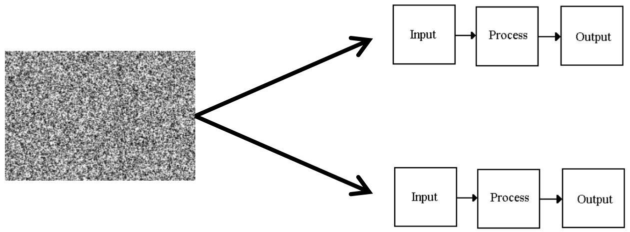
$$\Downarrow$$

$\cos^2\Big($ 

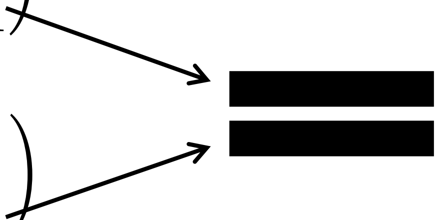
$\Big)$

$$= \lambda\left(\left(Y_{0|i-1}^{(1)} Y_{0|i-1}^{(1)^T}\right)^{-1} Y_{0|i-1}^{(1)} Y_{0|i-1}^{(2)^T} \left(Y_{0|i-1}^{(2)} Y_{0|i-1}^{(2)^T}\right)^{-1} Y_{0|i-1}^{(2)} Y_{0|i-1}^{(1)^T}\right)$$

Extra Slides

$U_{n|i-1} =$ 

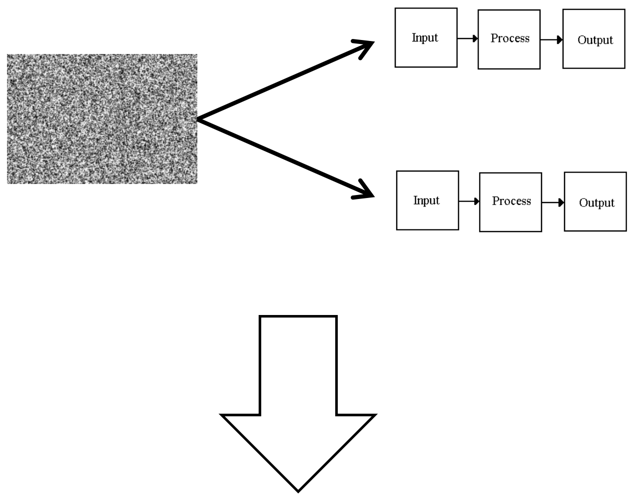
$$row(Y^{(1)}_{0|i-1}) = row\begin{pmatrix} X^{(1)}_n \\ U_{n|i-1} \end{pmatrix}$$



$$row(Y^{(2)}_{0|i-1}) = row\begin{pmatrix} X^{(1)}_n \\ U_{n|i-1} \end{pmatrix}$$

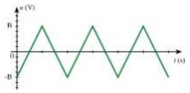
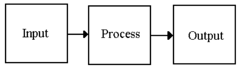
Extra Slides

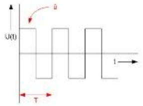
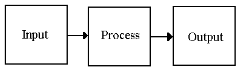
$U_{n|i-1} =$



$$\lambda\left(\left(X_n^{(1)}{X_n^{(1)}}^T\right)^{-1} X_n^{(1)}{X_n^{(2)}}^T \left(X_n^{(2)}{X_n^{(2)}}^T\right)^{-1} X_n^{(2)}{X_n^{(1)}}^T\right)$$

Extra Slides

$U_{0|i-1}^{(1)} =$  $\longrightarrow$ 

$U_{0|i-1}^{(2)} =$  $\longrightarrow$ 

$$\lambda\begin{pmatrix} \left(X_n^{(1)}X_n^{(1)^T}\right)^{-1} X_n^{(1)}\left(\prod_{row(X_n^{(2)})}+\prod_{row(U_{n|i-1}^{(2)})}\right)X_n^{(1)^T} & \left(X_n^{(1)}X_n^{(1)^T}\right)^{-1} X_n^{(1)}\left(\prod_{row(X_n^{(2)})}+\prod_{row(U_{n|i-1}^{(2)})}\right)U_{n|i-1}^{(1)^T} \\ \left(U_{n|i-1}^{(1)}U_{n|i-1}^{(1)^T}\right)^{-1} U_{n|i-1}^{(1)}\left(\prod_{row(X_n^{(2)})}+\prod_{row(U_{n|i-1}^{(2)})}\right)X_n^{(1)^T} & \left(U_{n|i-1}^{(1)}U_{n|i-1}^{(1)^T}\right)^{-1} U_{n|i-1}^{(1)}\left(\prod_{row(X_n^{(2)})}+\prod_{row(U_{n|i-1}^{(2)})}\right)U_{n|i-1}^{(1)^T} \end{pmatrix}$$

⇨ Consistent with stochastic case