Adaptive Signal Processing and Machine Intelligence

Course Introduction

Danilo Mandic

room 813, ext: 46271

Department of Electrical and Electronic Engineering

Imperial College London, UK

d.mandic@imperial.ac.uk, URL: www.commsp.ee.ic.ac.uk/∼mandic

Office hours: Tuesdays 14:00 - 15:00

Signal Processing at Imperial College

- □ **Dennis Gabor**, Nobel Prize laureate (1981 holography), nonlinear filters, Gabor filters (1950s-1980s), setting the scene for nonlinear adaptive filtering, information theory, and kernel architectures
- □ **Colin Cherry**, 'cocktail party effect', 1950s - 1970s, setting the scene for 'Blind Source Separation' (BSS)
- □ **Anthony Constantinides**, 'digital filters', 'Lagrange NNs', 1970s - now, setting the scene for the digital processing of real world signals

Currently, we specialise in many aspects of signal, image and speech processing, array signal processing, cognitive and neural architectures, sensor networks and wireless communications.

- 10 members of staff,
- 60 PhD students,
- 20 RAs,
- MSc course in Communications and Signal Processing

Our alumni are in leading technical positions around the world

Our "intellectual capital" in signal processing

The experimental setup for Gabor's Hologram
http://www.nobelprize.org/nobel_prizes/physics/articles/biedermann/

Spectral transformations for digital filters

A. G. Constantinides, B.Sc.(Eng.), Ph.D.

Indexing term: Digital filters

Abstract

The paper describes certain general transformations for digital filters in the frequency domain. The term digital filter is used to denote a processing unit operating on a sampled waveform, so that the input, output and intermediate signals are only defined at discrete intervals of time; the signals may be either p.a.m. or p.c.m. The transformations discussed operate on a lowpass-digital-filter prototype to give either another lowpass or a highpass, bandpass or band-elimination characteristic. The transformations are carried out by mapping the lowpass complex variable z^{-1} [where $z^{-1} = \exp(-j\omega T)$ and T is the time interval between samples] by functions of the form

$$e^{j\theta} \prod_{i=1}^{n} \frac{z^{-1} - \alpha_i}{1 - \alpha_i^* z^{-1}}$$

known as unit functions.

THEORY OF COMMUNICATION*

By D. GABOR, Dr. Ing., Associate Member.†

(The paper was first received 25th November, 1944, and in revised form 24th September, 1945

PREFACE

The purpose of these three studies is an inquiry into the essence of the "information" conveyed by channels of communication, and the application of the results of this inquiry to the practical problem of optimum utilization of frequency bands.

In Part 1, a new method of analysing signals is presented in which time and frequency play symmetrical parts, and which contains "time analysis" and "frequency analysis" as special cases. It is shown that the information conveyed by a frequency band in a given time-interval can be analysed in various ways into the same number of elementary "quanta of information," each quantum conveying one numerical datum.

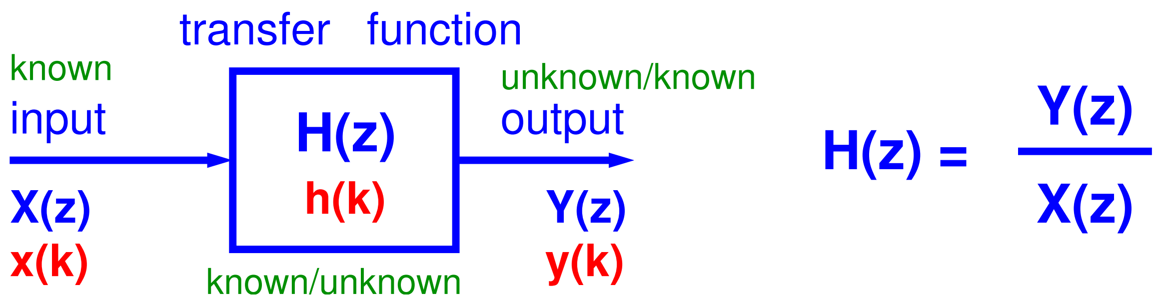
In Part 2, this method is applied to the analysis of hearing sensations. It is shown on the basis of existing experimental material that in the band between 60 and 1 000 c/s the human ear can discriminate very nearly every second datum of information, and that this efficiency of nearly 50% is independent of the duration of the signals in a remarkably wide interval. This fact, which cannot be explained by any mechanism in the inner ear, suggests a new phenomenon in nerve conduction. At frequencies above 1 000 c/s the efficiency of discrimination falls off sharply, proving that sound reproductions which are far from faithful may be perceived by the ear as perfect, and that "condensed" methods of transmission and reproduction with improved waveband economy are possible in principle.

In Part 3, suggestions are discussed for compressed transmission and reproduction of speech or music, and the first experimental results obtained with one of these methods are described.

The difference in this course

So far, you are familiar with problems for which:

- We have a **well defined transfer function** in the form



  which is **parametric** (model based), e.g. the stochastic AR model
  $\hat{x}(k) = a_1(k)x(k-1) + \cdots + a_p(k)x(k-p) + w(k), \qquad w \sim \mathcal{N}(0,1)$

- **Data are mostly stationary**

In this course we will introduce models which are:

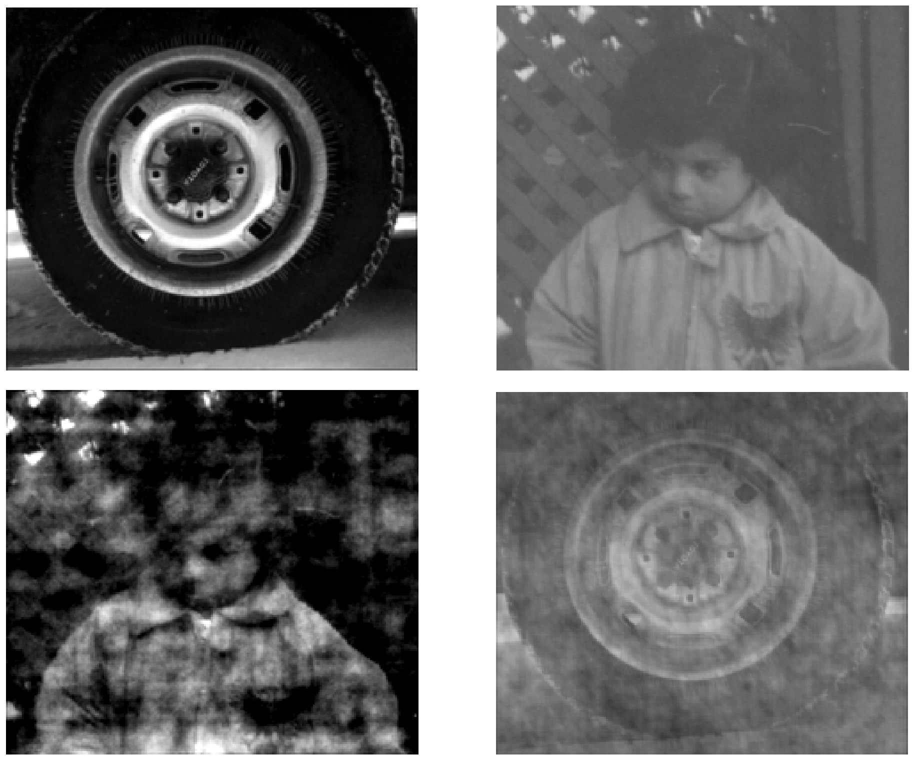
⊛ **Nonparametric**, that is, they do not assume any model *a priori*

⊛ **Adaptive**, capable of tracking the changes in the system/signal parameters in real time and in nonstationary environment

These are huge advantages that allow such models to operate in an online fashion and for real world data.

Illustrative problem: "Quality of Experience" (QoE)

We will introduce conventional and model-based **spectral estimation** techniques, and show their **statistical** properties. How about the phase spectrum?

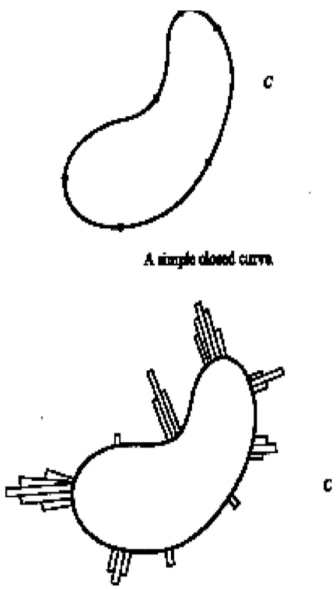


Surrogate images. *Top:* Original images I_1 and I_2; *Bottom:* Images $\hat{I}_1$ and $\hat{I}_2$ generated by exchanging the amplitude and phase spectra of the original images.

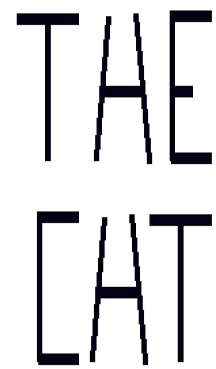
Complexity science $\looparrowright$ link with human perception

In the 50's phychologist Fred Attneave recorded eye dwellings on objects.

A closed curve (top) and the histogram of eye dwellings (bottom)

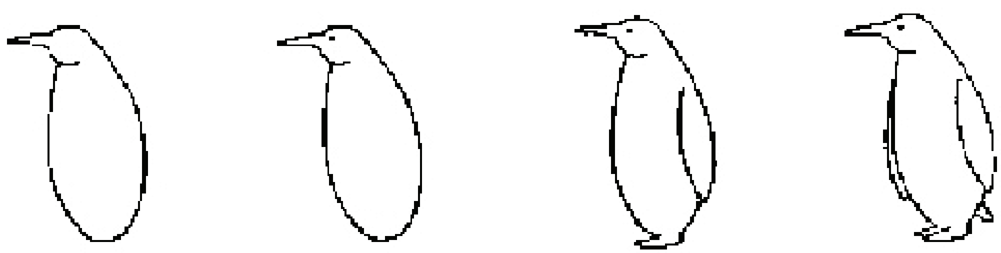


Beyond sparsity: What aspects of information can be compromised?



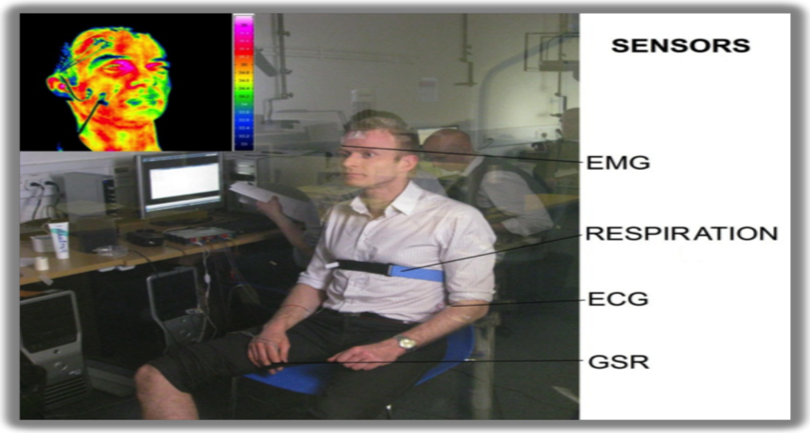
Marr & Hildreth (1980)

Is the drawing on the left still a penguin?



A challenge for DSP in the 21st century:
blurring the boundaries between ability and disability

Complexity in rehabilitation



Affective computing

Concierto for brain and chamber quartet

Spectral analysis

Latin Specter means a ghostly apparation, English spectre

- The word *spectrum* introduced by Newton in relation to his studies of the decomposition of white light into a band of light colours, when passed through a glass prism

- Spectral analysis as an established and ever expanding discipline $\looparrowright$ we are currently working on time-frequency estimation on nonlinear and nonstationary data

- Beginning about one century ago with the work by Schuster on detecting cyclic behaviour in time series

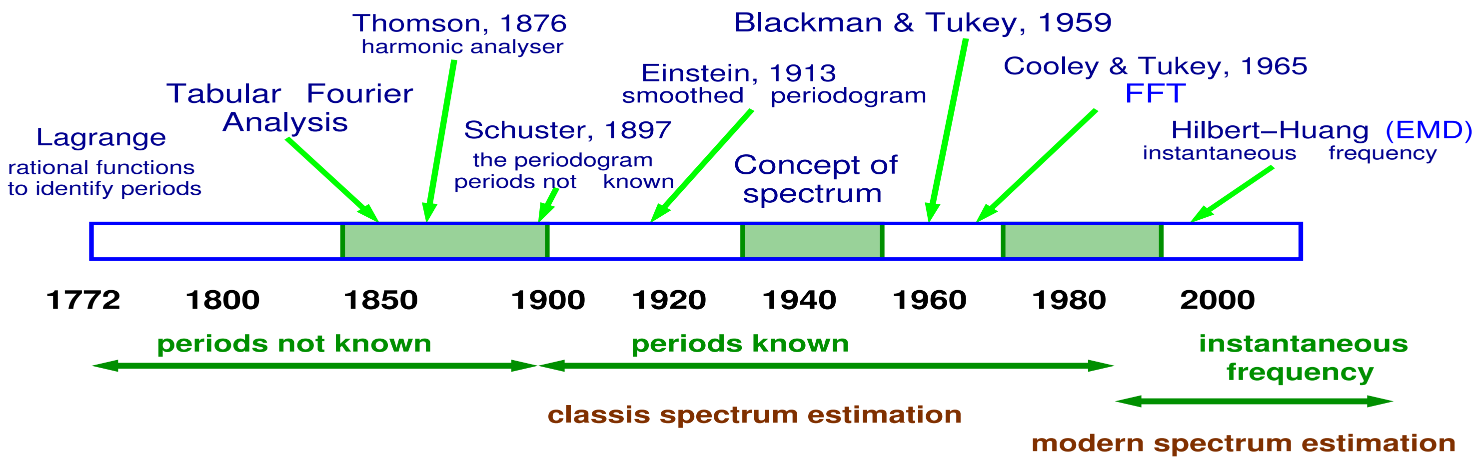
- Omni–present now (genomics, financial engineering, cognitive radio)

Despite the roots of the word spectrum, I hope the students will be a vivid presence in this course.

SE and ASP: The beginnings

From observing periodic and planetary motions

- Everywhere around us: the concept of frequency and sinewave



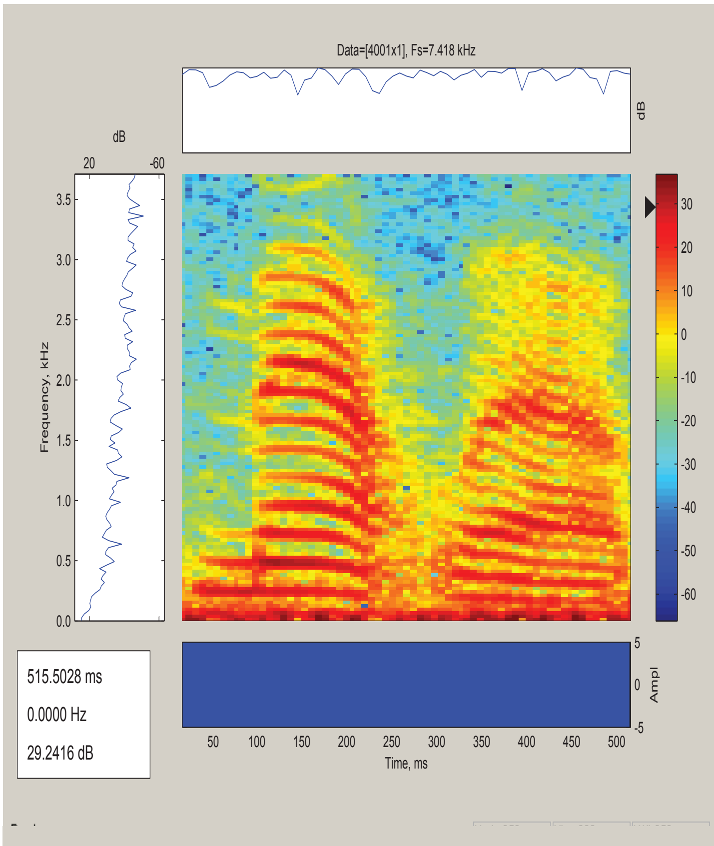
- Intimately related with the concept of complex numbers *(fundamental theorem of algebra)*
- Fourier's work from around $\sim$1800, FFT - mid 1960s, Lomb periodogram for irregularly sampled data (1972)

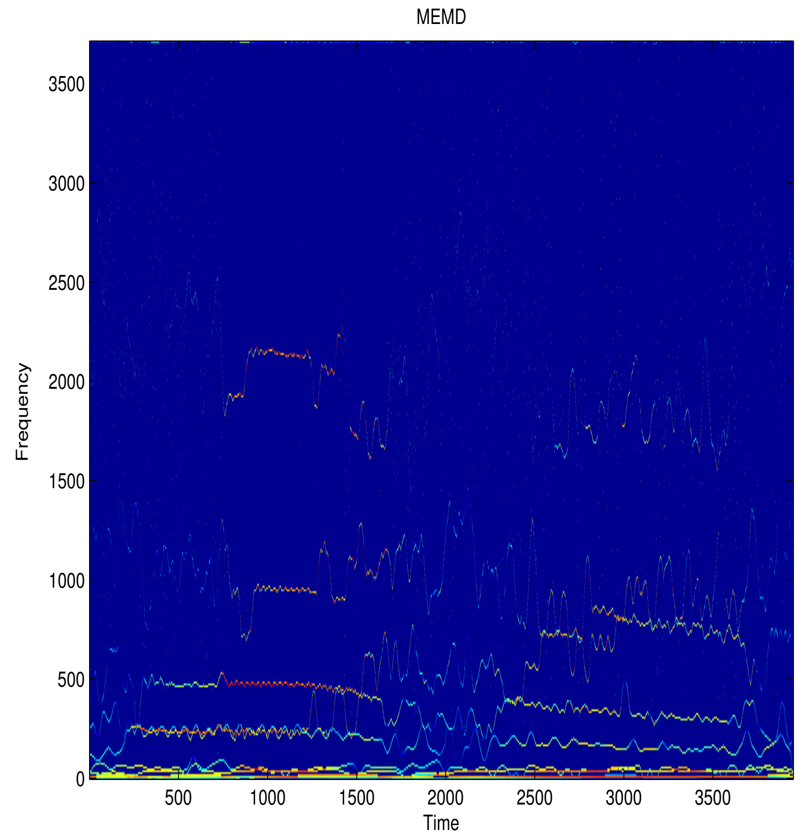
Recent advances in **Spectrum Estimation:** $\circledast$ irregularly sampled data, $\circledast$ very few data points (genomics and proteomics), $\circledast$concept of **instantaneous frequency**, $\circledast$ spectrum estimation of nonstationary data

In the quest for 'instantaneous anything'! A time-frequency representation of speech "Matlab"

(STFFT spectrogram)

(Hilbert-Huang spectra)



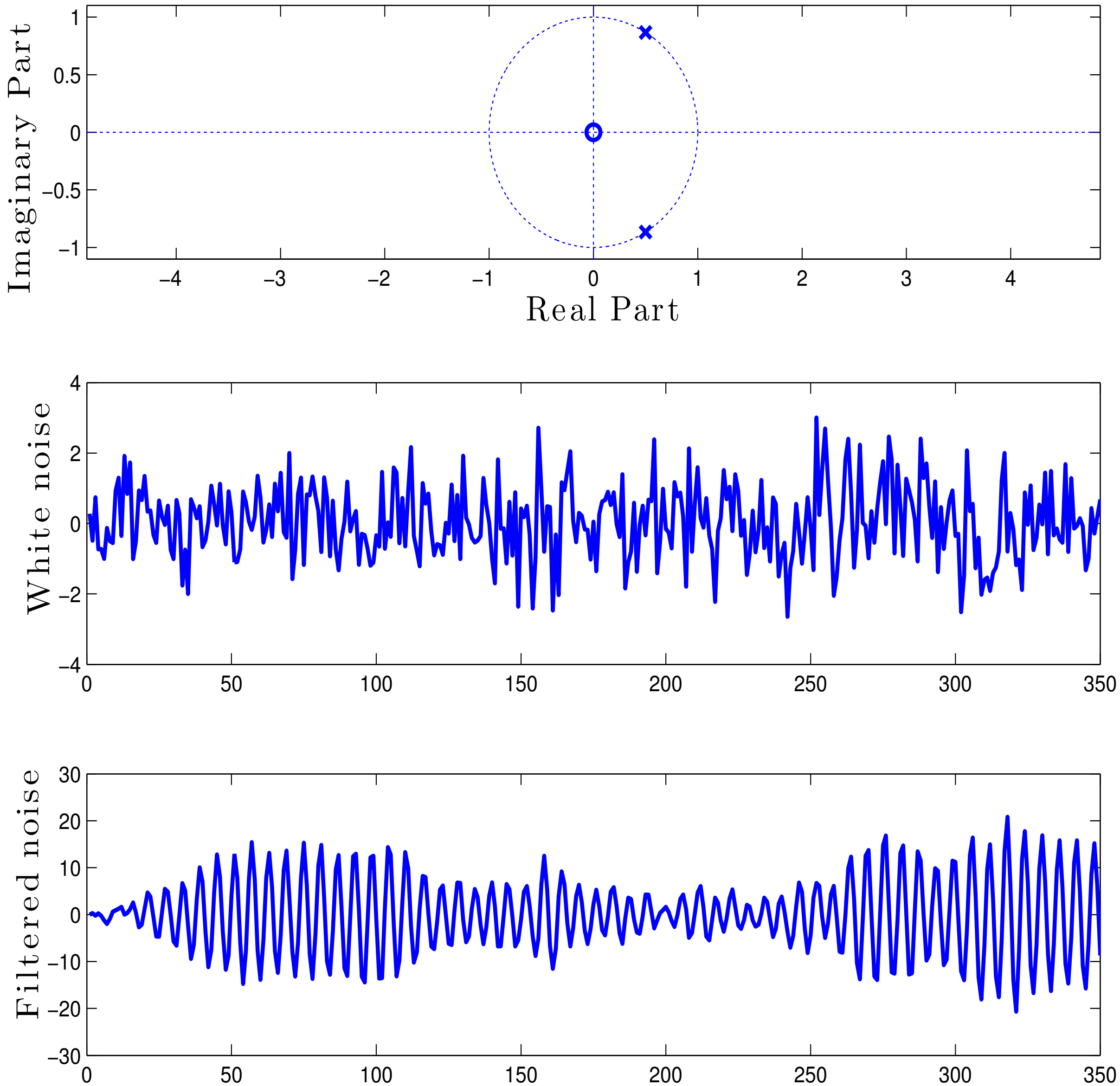


(win-len=256, overlap=200, ftt-len=256)

(no. of directions=64)

Spectral vs. adaptive processing: A real-world sinewave?

Shall we use spectral estimation or adaptive filtering to estimate it?



Matlab code:

```
z1=0;
p1=[0.5+0.866i,0.5-0.866i];
[num1,den1]=zp2tf(z1,p1,1);
zplane(num1,den1);
s=randn(1,1000);
s1=filter(num1,den1,s);
figure;
subplot(311),plot(s),
subplot(313),plot(s1),
subplot(312),;
zplane(num1,den1)
```

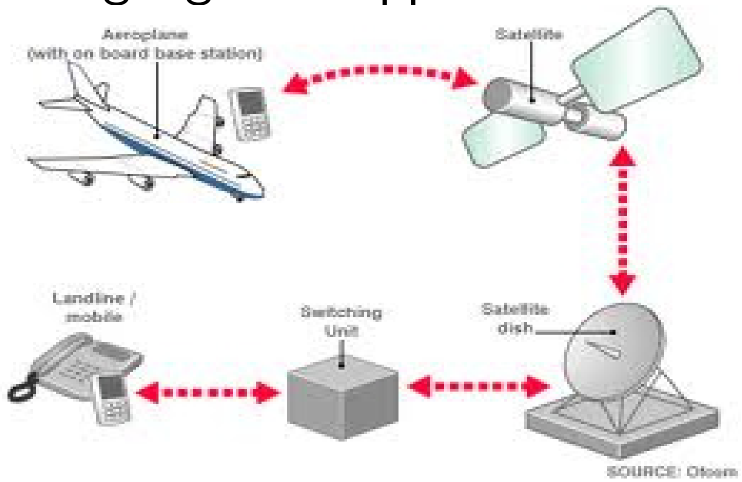
The AR model of a sinewave

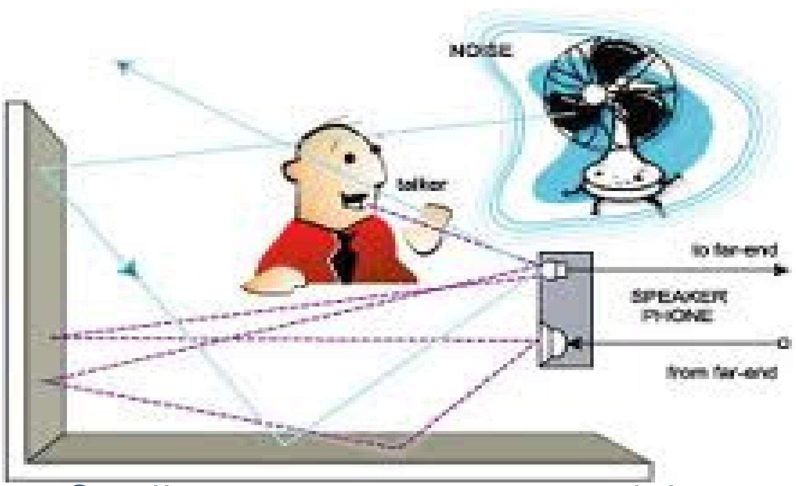
```
x(k)=a1*x(k-1)+a2*x(k-2)+w(k)
a1=-1, a2=0.98, w~N(0,1)
```

Adaptive signal processing

(also adaptive learning systems and the concept of neural networks)

To motivate the need for **adaptive signal processing** when processing real signals, and highlight its applications

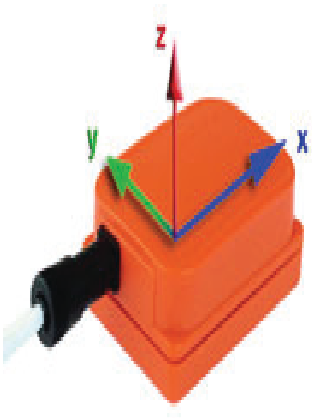




Applications. Satellite communications, mobile communications, finance, audio.

Adaptive signal processing: Applications

DSP is benefitting enormously from the progress in sensor technology



Renewable Energy

2D and 3D anemometers

control of wind turbine

Body motion sensor

3D - position, gyroscope, speed

gait, biometrics

Wearable technologies

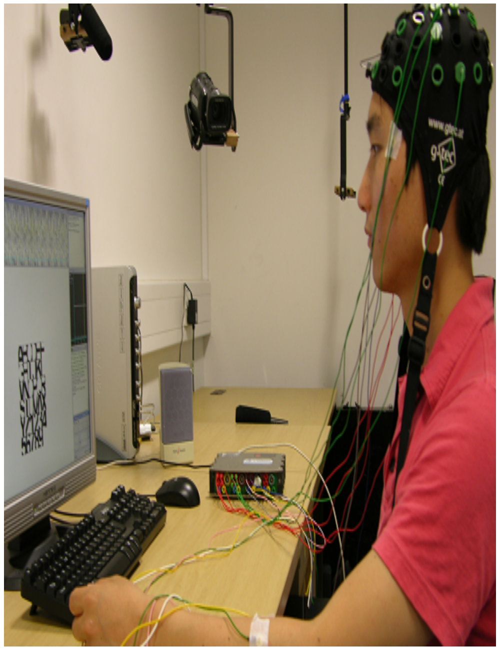
Biomechanics

virtual reality

Wind sensors - 3D anemometer

challenges for multidimensional DSP (e.g. division algebras)

Machine Intelligence applications



Brain Computer Interface

Decoding brain activity

to control computers

Spect. Est., ASP

Medical Applications

3D time-space

2D and 3D electromagnetic field

ASP, MI,

Avionics

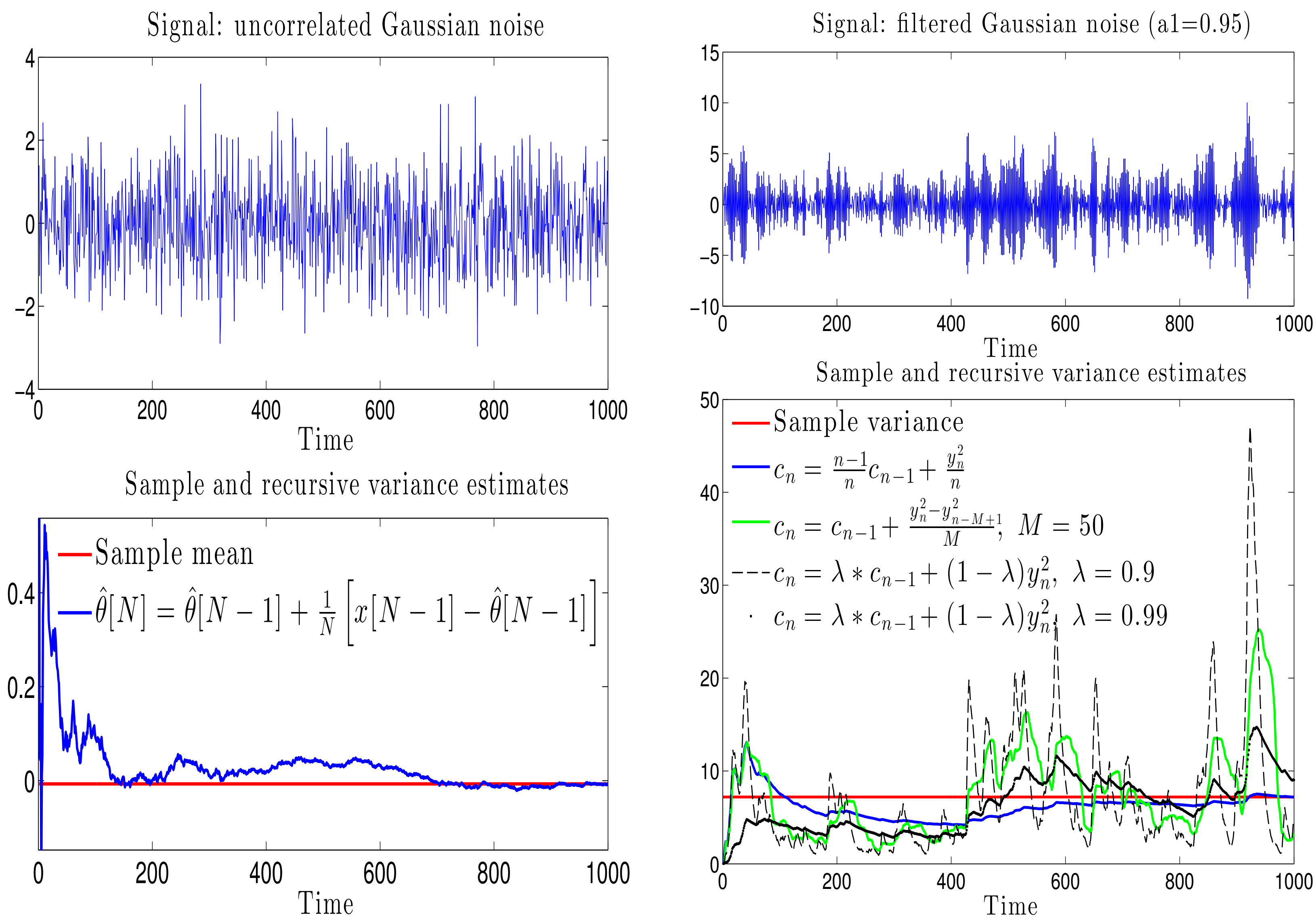
Trajectory tracking

Radar: Manoeuvre prediction

ASP, MI

Example: Practical estimators for real-time processing

Top left: uncorrelated Gaussian Top right: correlated Gaussian Bottom: corresponding variance estimates



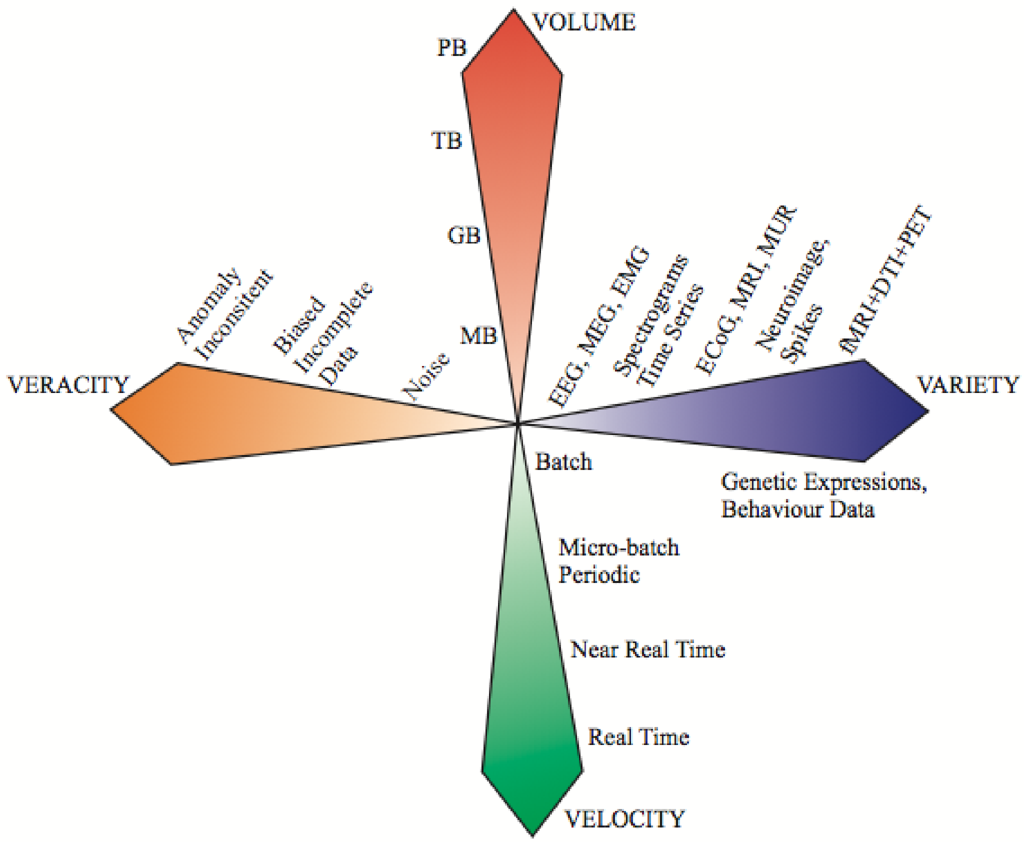
ASP & MI: The beginnings

The creation of the LMS algorithm (Widrow 1960)

- interference cancellation in telephone lines (in maths NLMS from 1927 Kaczmarz)
- complex least mean square (CLMS) 1975
- affine projection and proportionate NLMS (revolution in acoustics) - 1990s-2000s
- magnitude-only LMS, phase-only LMS (late 1990s onwards)
- widely linear (augmented) CLMS in the 2000s
- cooperative estimation over sensor networks, mid-2000s onwards
- quaternion LMS (QLMS) for 3D and 4D data (e.g. wind) - 2009
- Kalman filter: early 1960s until now (extended, unscented, particle)
- blind source separation early 1990s onwards
- neural networks: 1947, connectionsm 1986, reservoir computing "Echo State Networks" 2004 onwards, deep learning 2010-, Big Data 2012-

Setting the scene

Current trends: The 'Four Vs' in big data analysis



Does this remind you of the Olympic motto:

Citius – Altius – Fortius (faster, higher, stronger)

Course structure

1. **Spectral estimation.** Classical spectral estimation, signal modelling, autoregressive and subspace based spectral estimation

2. **Adaptive filtering.** Gradient based adaptive filters, complex-valued and multidimensional adaptive filters, recursive least squares

3. **Machine intelligence.** Links with state-space models, concept of an artificial neuron, neural networks, deep learning, links with tensors and Big Data

4. **Case studies.** Practical examples in communications, acoustics, biomedical engineering, renewable energy, finance

5. **Lecture course with problem sets**

   - *Modern spectral estimation: 5 lectures*
   - *Adaptive signal processing: 9 lectures*
   - *Machine intelligence: 6 lectures*

6. **Assessment:** 100% Coursework,   Feedback after Assignment 1

Course structure

Special emphasis will be on demonstrating the close links between spectral estimation, adaptive signal processing and machine intelligence

Throughout the course we will discuss Matlab implementation

Portable data acquisition

Comparison of features of bio-amplifier

Imperial Amplifier (I-Amp)

Dimensions: ~40 x 15 x 15 mm
Weight: ~ 30 g
Power supply: one coin baterry
Recording time: 48+ hrs

g.USBamp

Avatar EEG

Dimensions: 197 x 155 x 40 mm
Weight: 1000 g
Channel no: 16
Power supply: mains
Price: 13,000 Euro (~10,000 GBP)

Dimensions: 76 x 53 x 38 mm
Weight: 60 g
Channel no: 8
Power supply: 2 AA baterries
Recording time: 24 Hrs
Price: >3,000 GBP

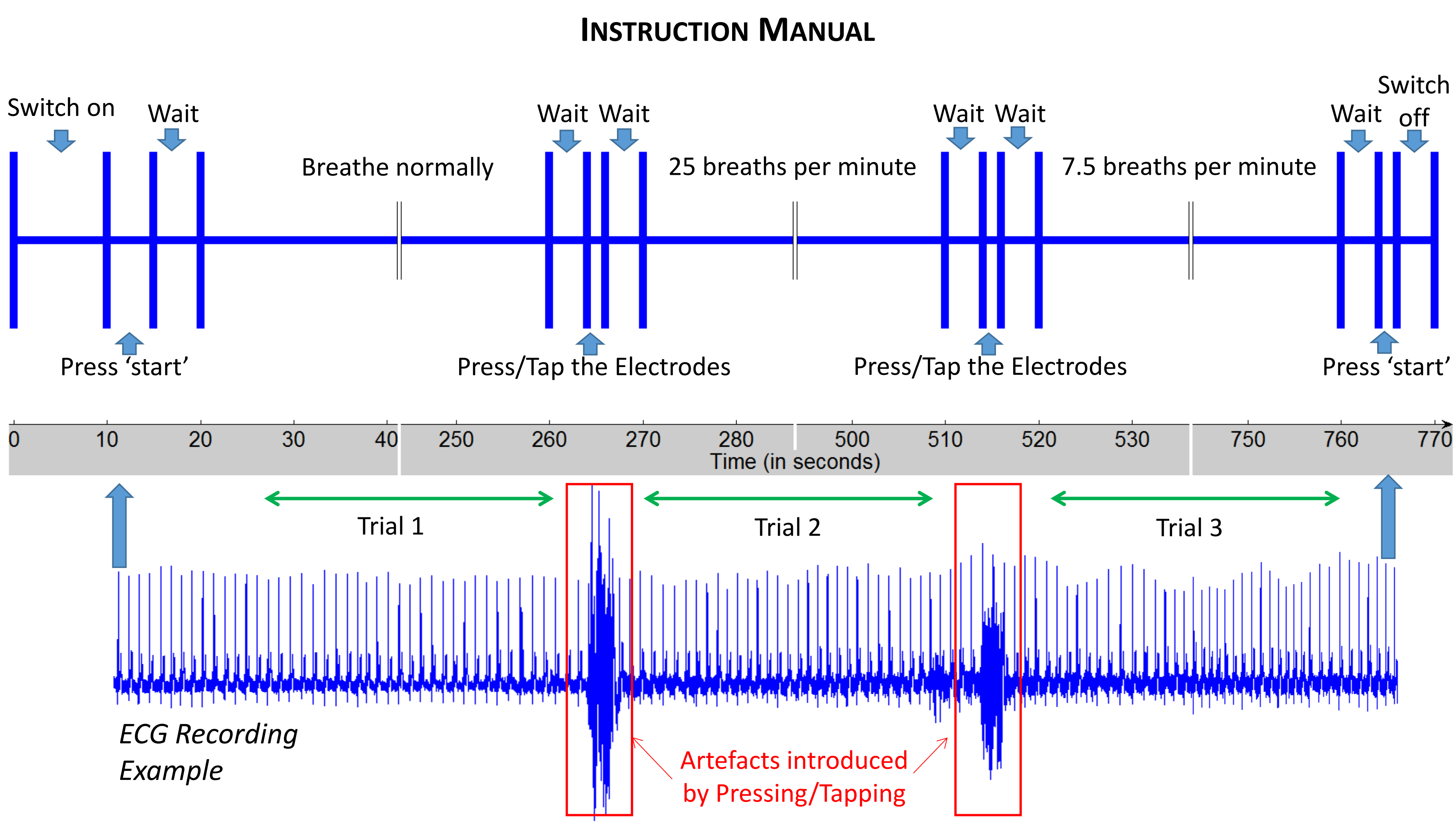
Coursework $\looparrowright$ partly based on students' own data

Experiment summary: Students recorded their ECG under varying respiration patterns: (1) unconstrained breathing; (2) controlled breathing at a fast pace (50 breaths/min) and slow pace (15 breaths/min).

Students were provided with easy to follow on-screen instructions.

Student placing electrodes on a team-mate.

ECG recording instruction



Prerequisites and reference material

- The course is largely self-contained, there are no prerequisites
- Having attended Digital Signal Processing, Advanced Signal Processing, and basic Probability Theory and Statistics would be useful
- Knowledge of some form of programming language is essential for the understanding of the implementation issues of the algorithms covered
- My preference is Matlab, the *de facto* language of scientific computing (at least in technology), especially for non-hardware implementation

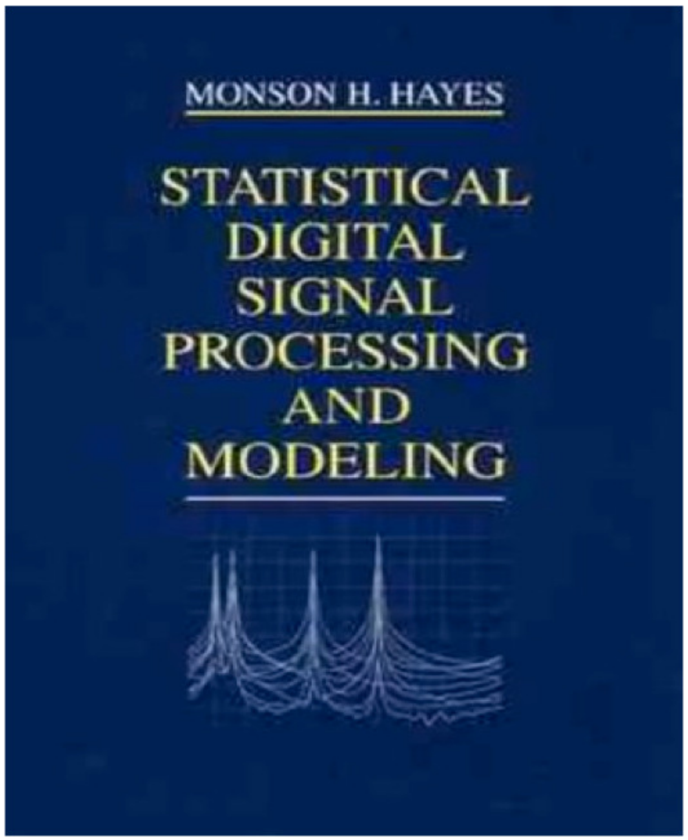
Literature:

1. **Course notes and problem/answer sets: by Dr Mandic**
2. **The book by M. Hayes for Spectral Estimation**
3. **Hayes' book and D. Mandic's book for Adaptive Filtering**

   Any other bits and pieces will be in course notes

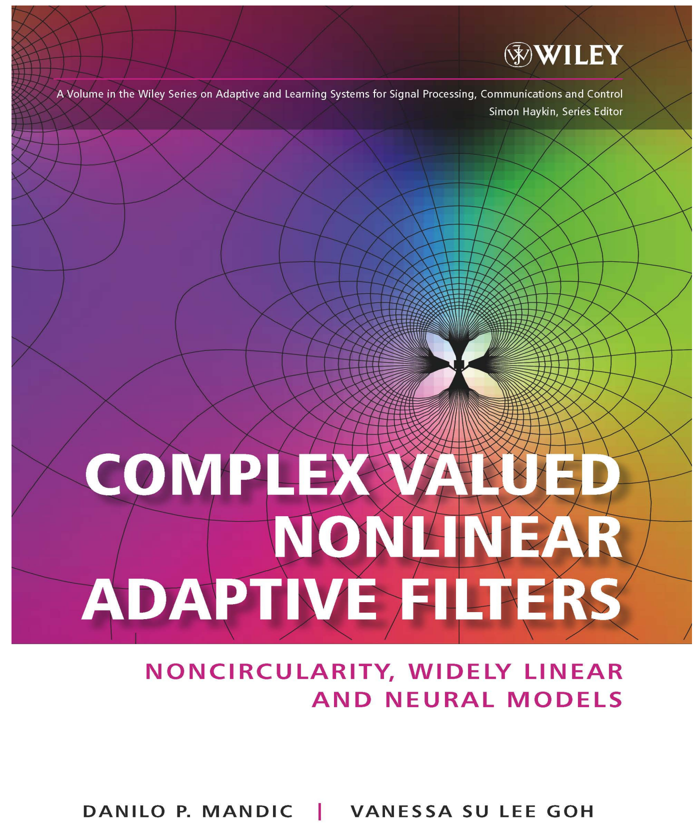
Textbooks: Recommended

M. Hayes (*Statistical Signal Processing*, several editions)



spectral estimation and part of adaptive filters

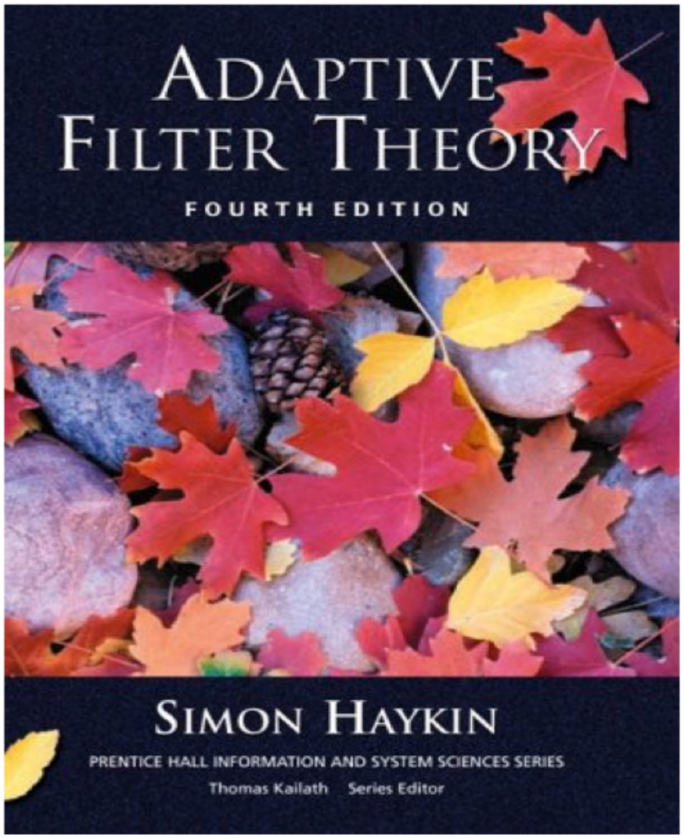
D. Mandic and S. Goh (*Complex Adaptive Filters*, Wiley 2009).



real, complex, and neural adaptive filters

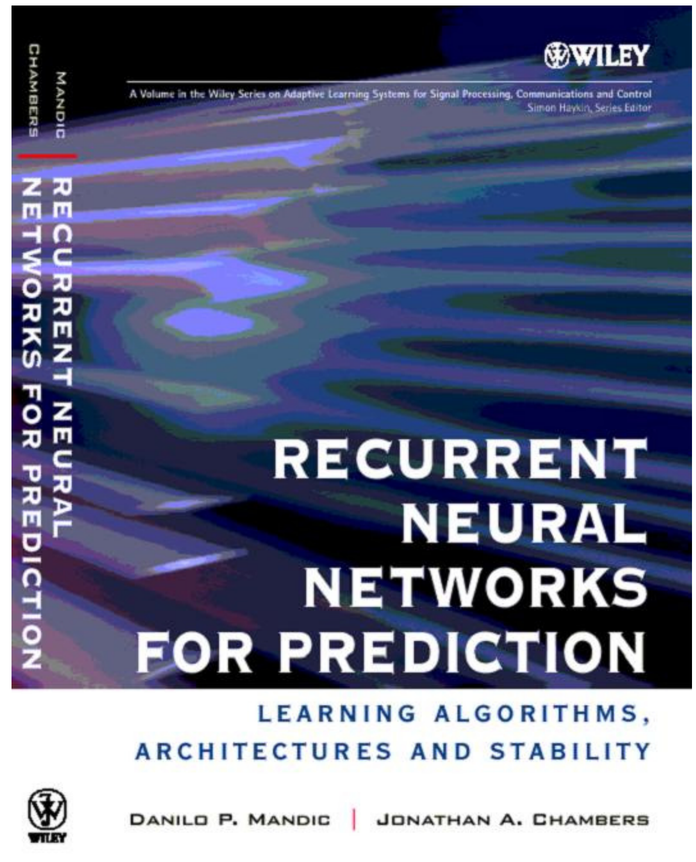
Textbooks: Additional reading

S. Haykin (*Adaptive Filtering Theory*, several editions)



a comprehensive account of adaptive filters

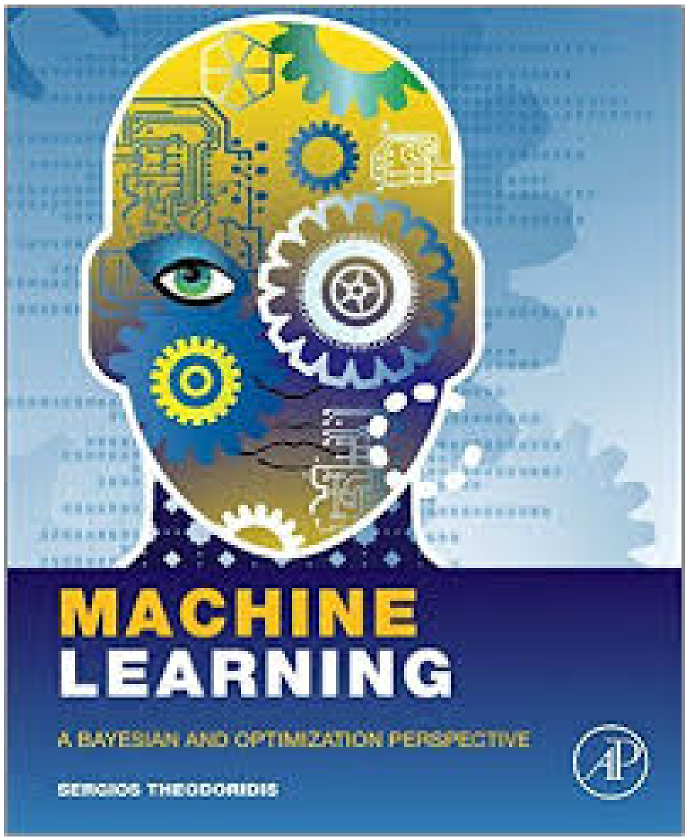
D. Mandic & J. Chambers (*RNNs for Prediction*, Wiley 2001).



feedback and neural network architectures

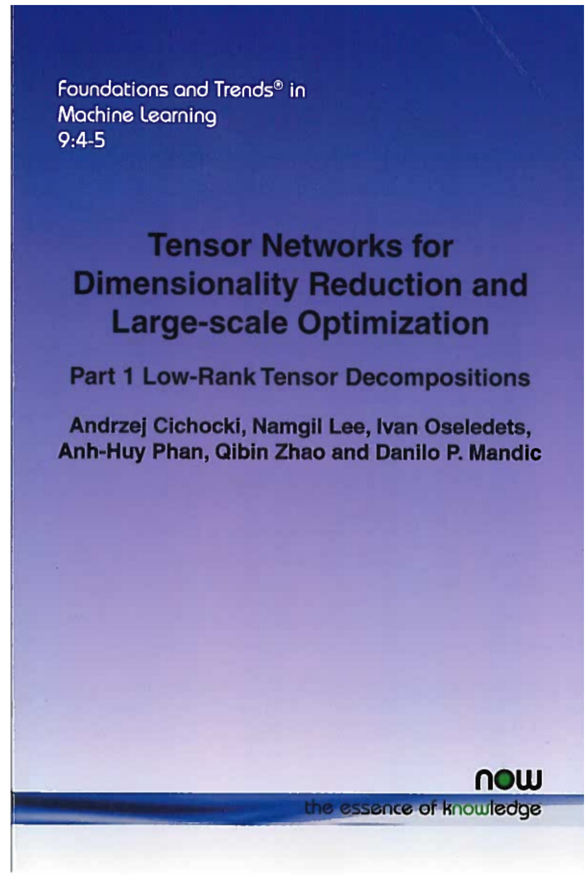
Useful supporting material

S. Theodoridis (*Machine Learning*, 2015)



a Bayesian and optimisation perspective to machine learning

A. Cichocki, D. Mandic, *et al.* (*Tensor Networks*, 2016).



Big Data, tensors, dimensionality reduction

Course plan

1 Lect: Week 2, Course introduction and motivation

2 Lect: Week 2-3, Classical spectral estimation

4 Lect: Week 3-4, Modern spectral estimation

6 Lect: Week 4-6, Stochastic gradient based adaptive filters

6 Lect: Week 6-8, Complex, feedback, and fully adaptive filters

3 Lect: Week 9-10, Nonlinear and neural filters and application case studies

Course web page: www.commsp.ee.ic.ac.uk/∼mandic/Teaching

Lectures, additional reading, homework, problem sets, and other material will be put on course webpage

To conclude:

ASP & MI is a combination of two very important areas in modern signal processing

- In this way, we are able to move beyond the concept of transfer function and to relate signals in both spectral and temporal domains
- Adaptive filters are an enabling technology for many real world applications in nonstationary environments (acoustics, speech, communications, teleconferencing, biomedicine, renewable energy, genomics and proteomics)
- Machine intelligence algorithms can operate in a real-time, and their coefficients are adapted online (in a data driven manner)
- The nonparametric, data-adaptive, mode of operation allows us not to assume any model imposed on the data

The material in this course - statistical signal processing - is generic and is applicable across the areas of electronics and engineering

Notes:

○

Notes:

◦

Notes:

○

Notes:

∘