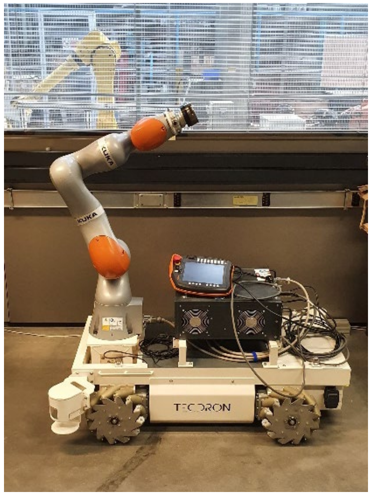
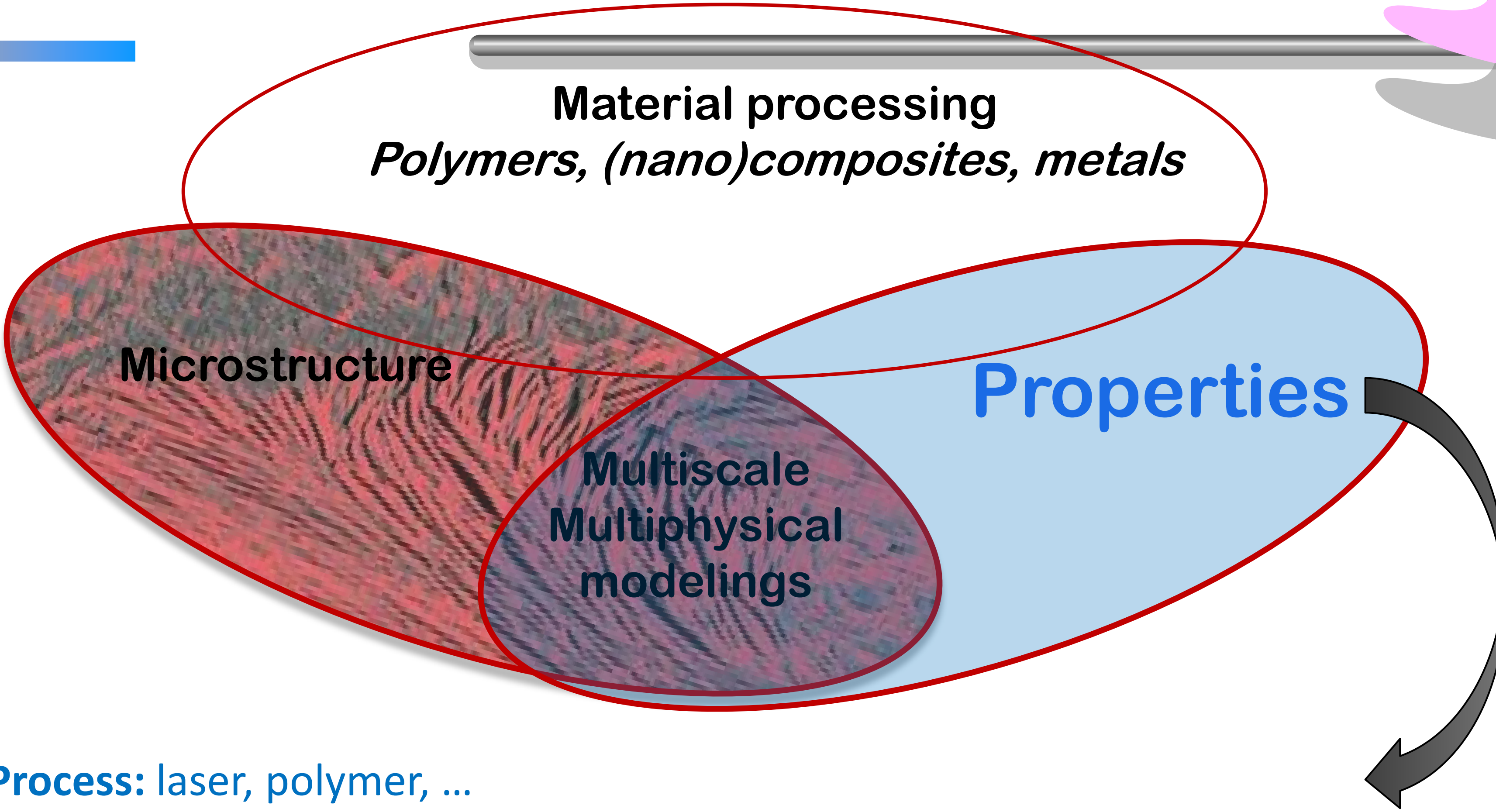
ESTIMATION THEORY AND PARAMETRIC IDENTIFICATION

Mechbal N. / Guskov M. / Rebillat M.

*Laboratoire **PIMM**, UMR-CNRS*
Arts et Métiers ParisTech (ENSAM), Paris



PIMM laboratory UMR – CNRS - Le Cnam



- **Process:** laser, polymer, …
- **Mechanical,** electrical, thermal, optical, dimensional properties
- **Long-term properties**: durability of polymers, composite, gigacycle fatigue of steel

Multidisciplinary competences

From chemistry and mechanics to and **control and IA**, from experimental to numerical simulations

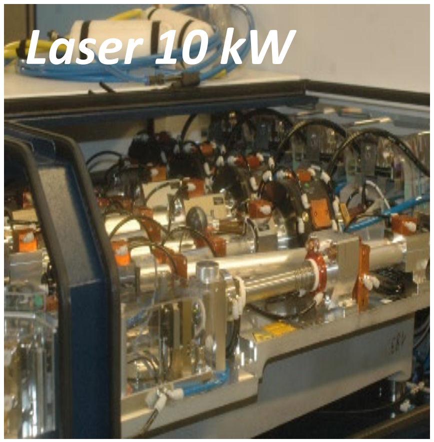
Mechbal N.

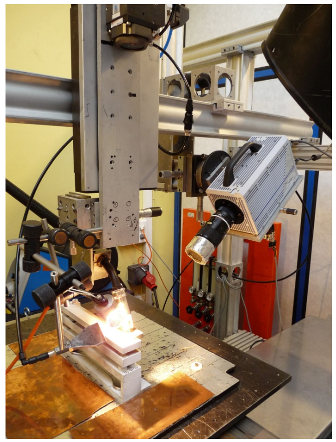
PIMM laboratory UMR – CNRS - Le Cnam

Laser processing

Welding, drilling, cutting, surface treatment

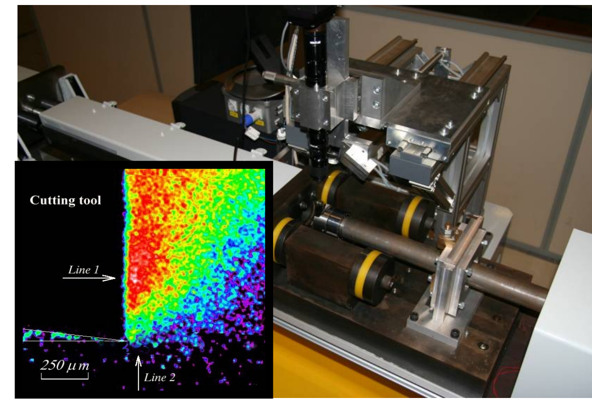
3D printing, heterogeneous assembling, laser chock





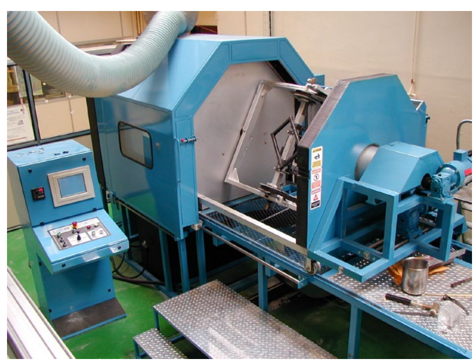
Machining

Metal thixoforming



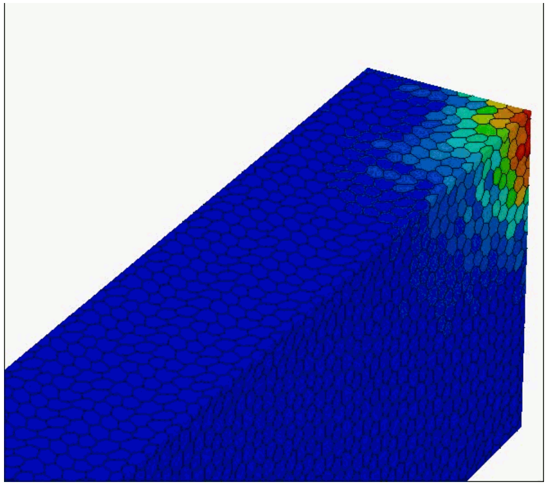
Polymer processing

(micro)injection, extrusion (multinanolayers), rotomolding, thermoforming, 3P printing

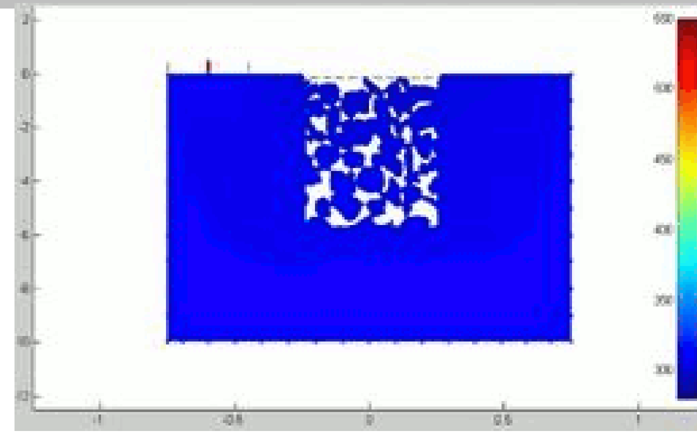


Mechbal N.

Multiscale and multiphysical modelings and simulations

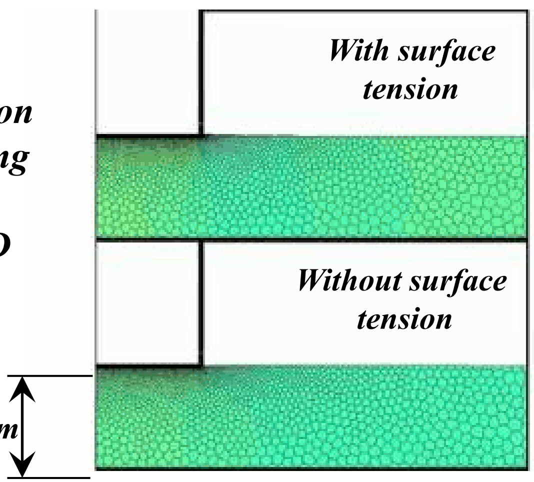


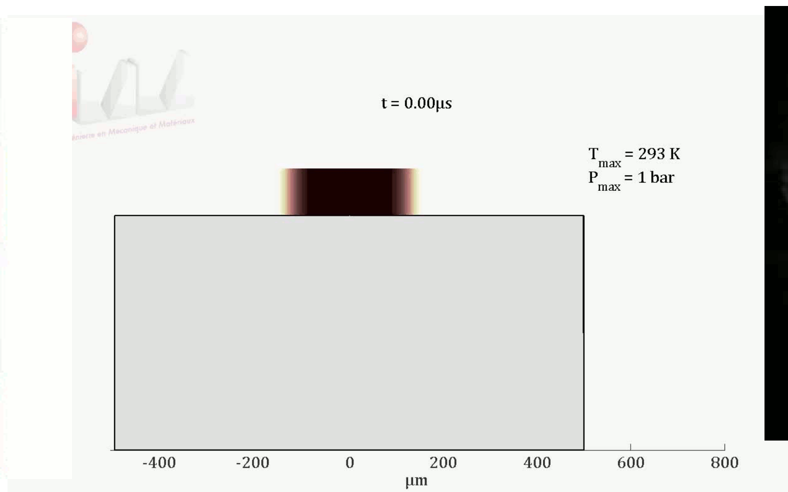
Simulation of steel machining
L.Illoul PhD (2009)



Simulation of laser fusion of a thermoplastic powder bed
Fadiplast project – Dassault A., D. Defauchy PhD (2013)

Simulation of silicon waffer nanoprinting
Silsef – CEA
H. Teyssèdre PhD (2013)

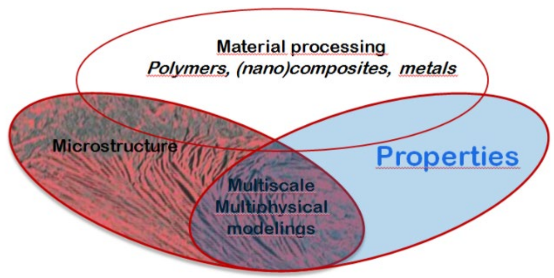




Simulation of laser drilling
DGA - J. Girardot PhD (2013)

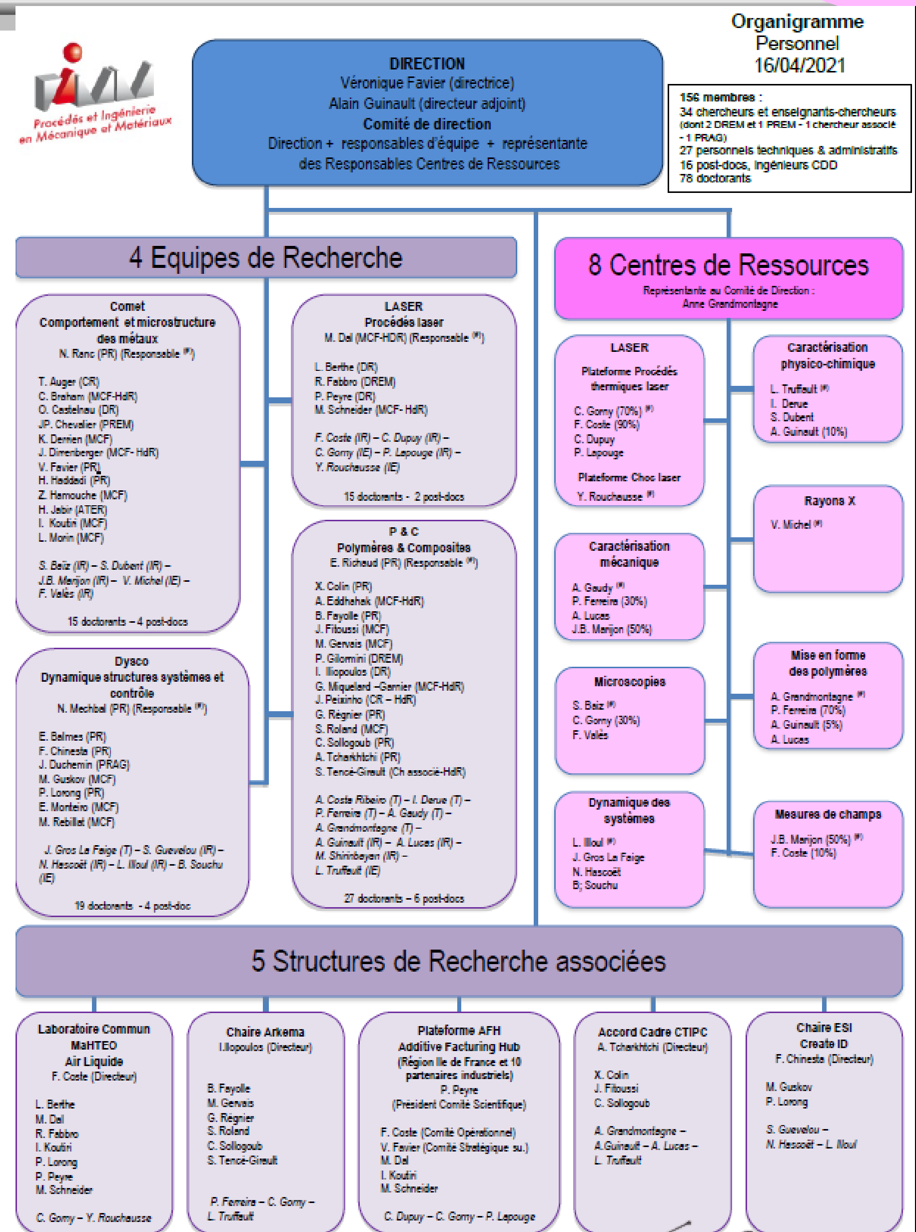
PIMM laboratory UMR - CNRS - Le Cnam

PIMM structuration



- **4 équipes**
  - Comet, P&C, Laser **& Dysco**

- **Equipe DYSCO : Dynamique Structures Systèmes et Contrôle**
  - ***Axe: Commande et surveillance des structures intelligentes***
    - Robotique & Structures intelligentes
    - Surveillance des structures
    - Contrôle Actif



Mechbal N.

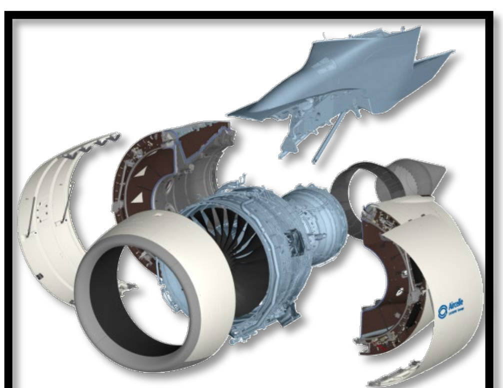
Projects examples

Intelligent Robot of Assistance

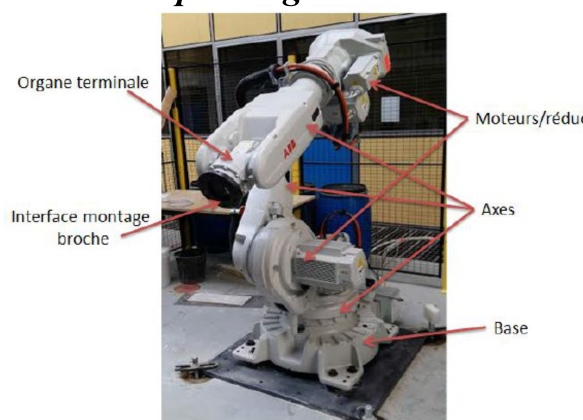
Vibration analysis

Thin Shell Tactile Sensing by Acoustic Wave Diffraction Patterns

SHM for plane engine and structures



Machining and 3D concrete printing robots



inflatable robot

"Tell me and I forget

Teach me and I remember

Involve me and I learn"

Benjamin Franklin

Course Organization

- **Lecture materials:**
  - Course handout
  - Some basic Matlab code
- **Organization:**
  - Course meeting times: every **Tuesday from 8h30 to 12h**
  - Homework: issued in lecture and solution is given in the following class
  - Matlab Class: Students are allowed to work together in designing algorithms, in interpreting error messages, and in discussing strategies for finding bugs, but NOT in writing code.
- **Evaluation**
  - Exams / Scientific articles review
  - Matlab project / Homeworks assignment
- **Contact:** nazih.mechbal@ensam.eu

There are 3 main Parts in this course:

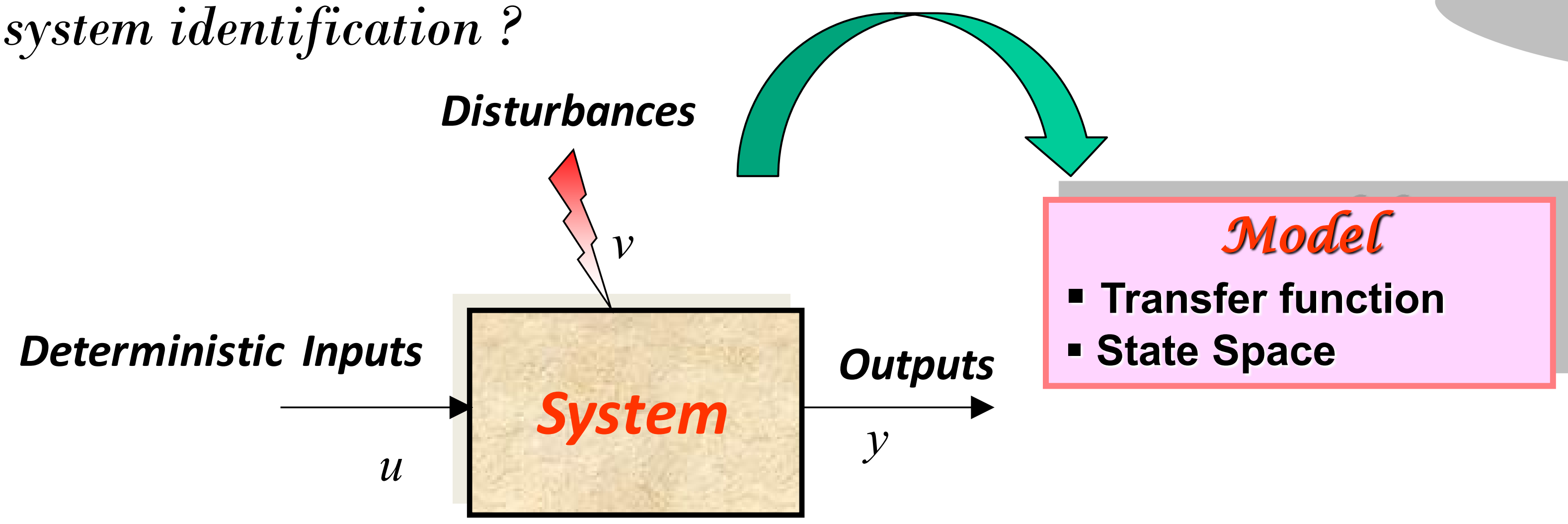
- **Part 1 Estimation Theory – N. Mechbal – 6 lectures**
  - Review of probability theory
  - Linear Estimation Theory – Deterministic and Bayesian methods
  - Nonlinear estimation – EKF, UKF and PF
- **Part 2 Parametric identification theory – N. Mechbal & M. Rebillat – 5 lectures**
  - Mathematical foundations of system identification
  - Parameter estimation
  - Sensors and Signal processing for identification – discrete time processing, frequency analysis, denoising
- **Part 3 Model identification of flexible manipulators – M. Guskov – 3 lectures**
  - Modeling of flexible manipulators
  - Modeling for vibration analysis
  - Application: Robotic machining

The derivation of a system description from observed data is called system identification and the resultant system description is a model.

Description in the form of mathematical models is based on mathematical functions that themselves involve parameters which give model its scale, rates, … and that we wish to estimate

Motivations and Objectives

✷ *Why system identification ?*



✦ *Derivate a relevant <u>mathematical model</u> of a dynamical system based on input output data*

✷ *Why models ?*

✦ *Interpretation of observations*

✦ *Prediction and forecasting of states, outputs and behavior*

✦ *Necessary link between experiments and decision*

Mathematical model vs Physical model

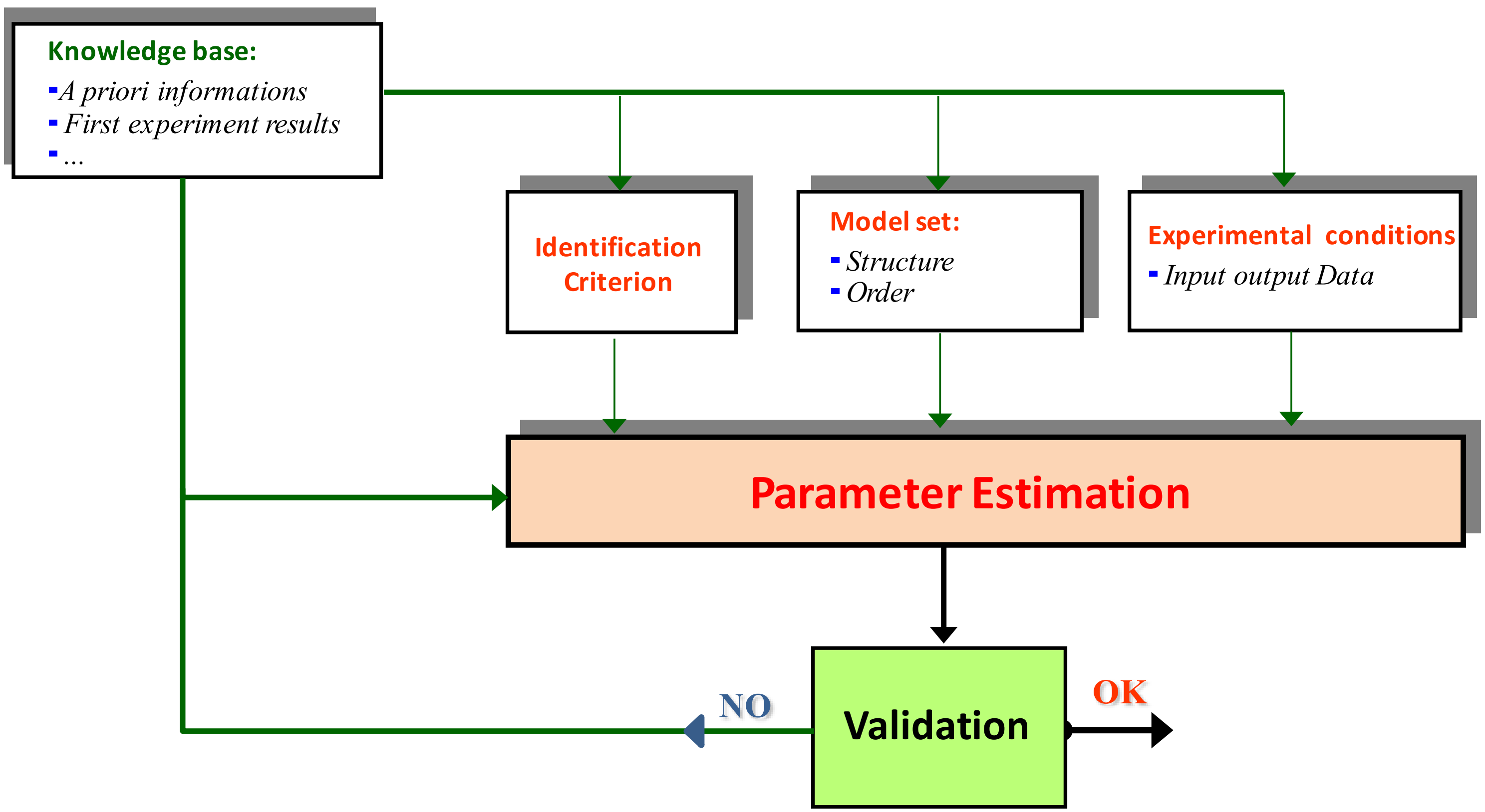
- Two main approaches to construct a mathematical model of a dynamic system:

  1. **Physical:** use physical laws (First principles, conversion laws, ..) that carefully explain the underlying essential mechanisms of observed phenomena.
     - Are not falsified by the available experiments.
     - Nonlinear partial differential equations
     - ∞-dimensional
     - Used for simulation and conception
  2. **Mathematical:** experimental identified model that look for relations in the data. *It quality is dictated by the ultimate goal it serves (control, estimation or supervision purpose).*
     - *Depends on experimental conditions: valid under prescribed conditions*
     - *Ordinary differential equations*
     - *Finite dimension*
     - *No physical meaning*

- ***Engineers have to deal with a trade-off between model complexity and accuracy***

System identification procedure

- The identification is an experimental approach for determining the parametric dynamic model of a process (system) which will act at best as the actual process



- It is important to realize that building *models from first principles only and system identification based only on measurements* are two extreme cases. Quite often a combination of the two is encountered.
- *The message is that system identification provides a meaningful engineering alternative to physical modeling. Compared to models obtained from physics, system identification models have a limited validity and working range and in some cases have no direct physical meaning.*

- **Interrogations ?**
  - How to choose the model structure ?
  - How to choose inputs ?
  - How to integrate model *a priori* knowledge ?
  - How to choose the adequate identification method ?
  - How to evaluate (validation) the quality of the obtained model?

✷ Model structure and characterization

- Modeling a process: "*Predicting it behavior under the influence of a known stimulus* "

- This step, essentially qualitative, is informed and validated by the rest of the procedure. This is often the most difficult.

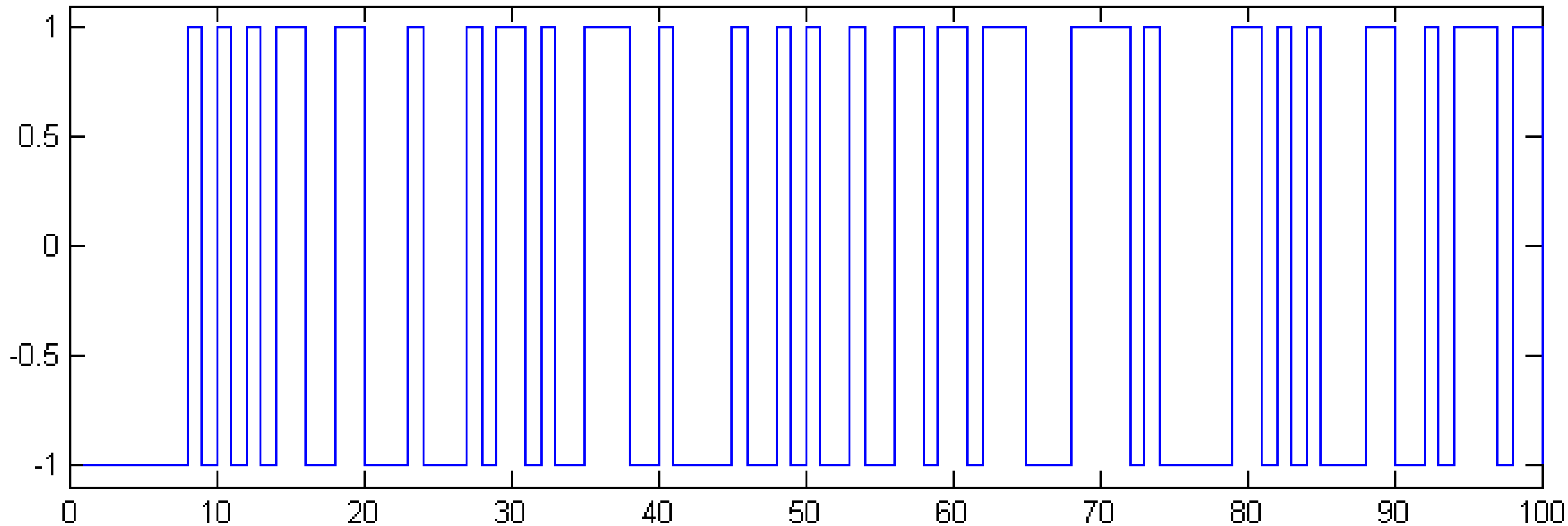
it appeals as much to the experience as to the methodology.

Dynamical Model

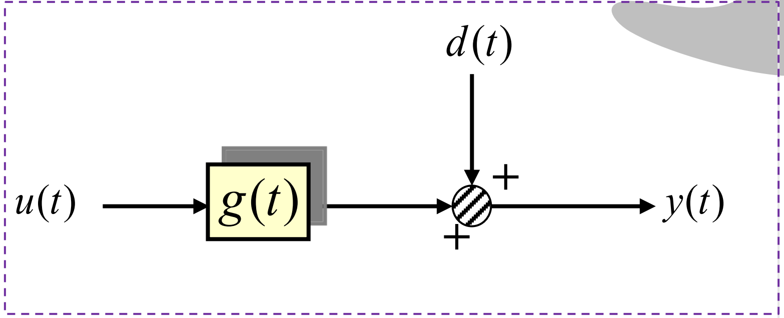
- ***Non parametric***
  (ex. correlation analysis, frequency response, …)
- ***Parametric***
  (ex. transfer function, …)
  - ***Continuous***
  - ***Discrete***

Choice of stimulus

- Excitation signal must be finite energy and amplitude.
- Quite often a broadband signals are desirable: it sufficiently exciting and the date contains information of the system in a large range frequencies.
- For example, a PRBS (Pseudo-Random Binary Signal). It is called "pseudo-random" as it exhibits some properties that approximate a random signal. It is a "***staircase***" signal that is suitable for actuators than a random signal



Non-parametric identification



Correlation technique

$$y(t) = g(t) * u(t) + d(t) \qquad *: \text{convolution}$$

$$y(t) = \underbrace{\sum_{k=0}^{\infty} g(k)u(t-k)}_{\text{Processus}} + \underbrace{d(t)}_{\text{perturbation}}$$

- *We use the Wiener-Hopf equation:* $R_{yu}(\tau) = E\{y(t)u(t-\tau)\} = \sum_{k=1}^{\infty} g(k)R_{uu}(k-\tau)$

- *If now the input is white noise of variance α then we have :* $\boxed{\hat{g}(\tau) = \dfrac{R_{yu}(\tau)}{\alpha}}$

Spectral technique

$$y(k) = \underbrace{\sum_{i=0}^{\infty} g(i)u(k-i)}_{\text{Processus}} + \underbrace{d(k)}_{\text{perturbation}} = G(z)u(k) + d(k)$$

- *Using the PSD*

$$\hat{G}(e^{j\omega T}) = \frac{\hat{S}_{yu}(\omega)}{\hat{S}_{uu}(\omega)} \quad \text{avec} \quad \begin{cases} \hat{S}_{yu}(\omega) = \sum_{\tau=-N}^{N} \hat{R}_{yu}(\tau)e^{-i\tau\omega} \\ \hat{S}_{uu}(\omega) = \sum_{\tau=-N}^{N} \hat{R}_{uu}(\tau)e^{-i\tau\omega} \end{cases}$$

- *We can also use an Empirical transfer Function Estimation (ETFE) by performing DFT of u and y:*

$$\boxed{\hat{G}(e^{j\omega T}) = \frac{Y_N(\omega)}{U_N(\omega)}}$$

Mechbal N.

Non-parametric identification methods

- Simple Example: Suppose that the system we want to identify is given by:

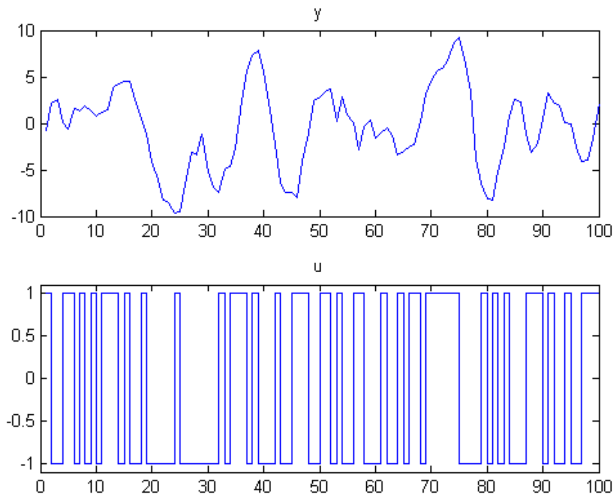
$$y(k)A(z) = B(z)u(k) + d(k)$$

with

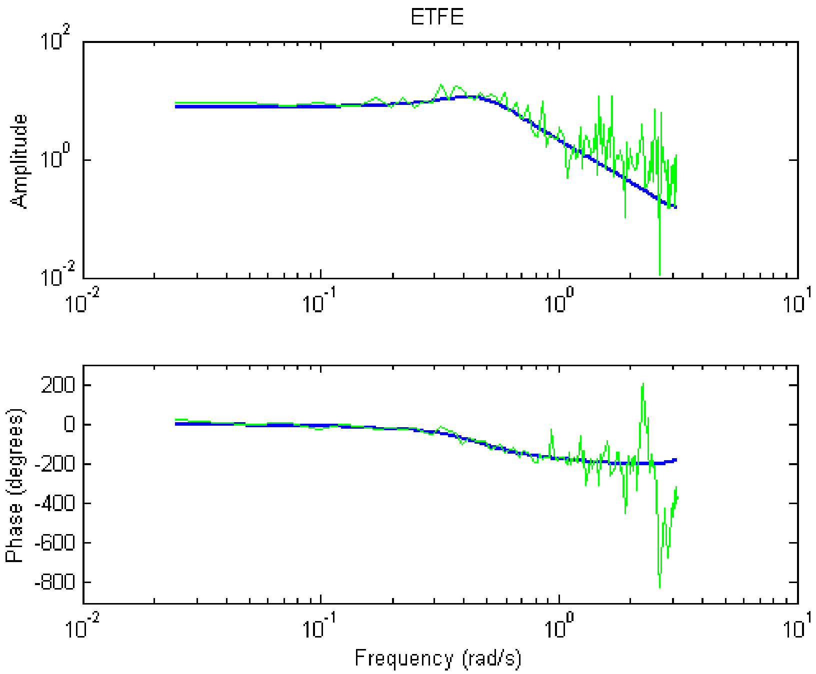
$$A(z) = 1 - 1.5z^{-1} - 1 + 0.7z^{-2}$$

$$B(z) = z^{-1} + 0.5z^{-2}$$

Input/output using a PRBS

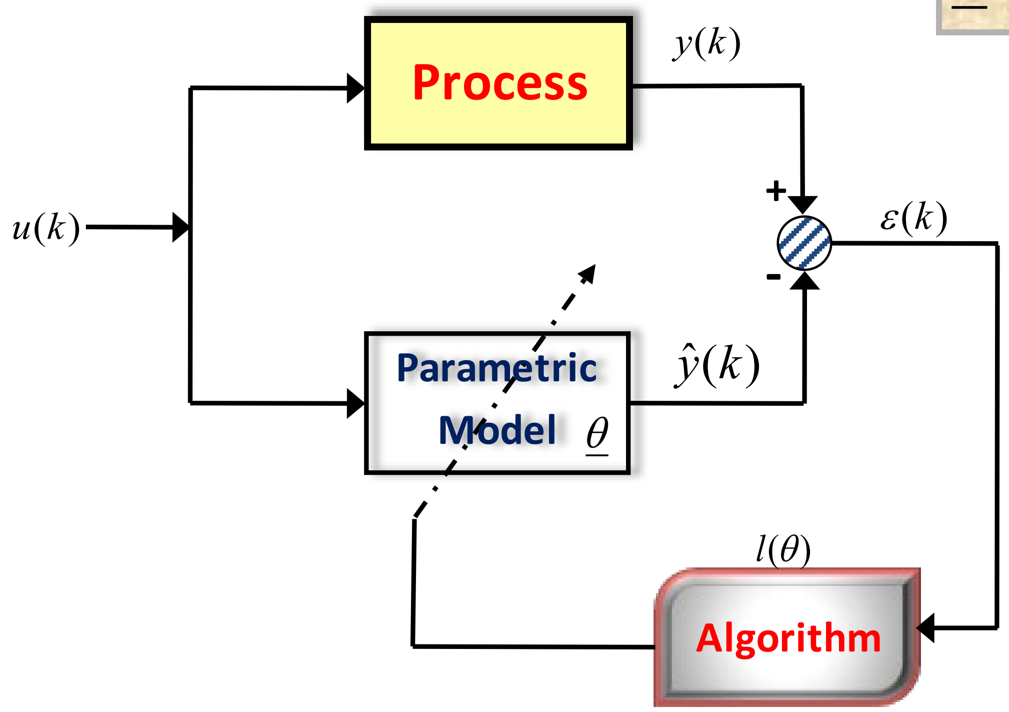


Spectral Analysis



Parametric identification

$$\underline{\hat{\theta}} = \arg\min f_N(\underline{\theta}, \Phi)$$

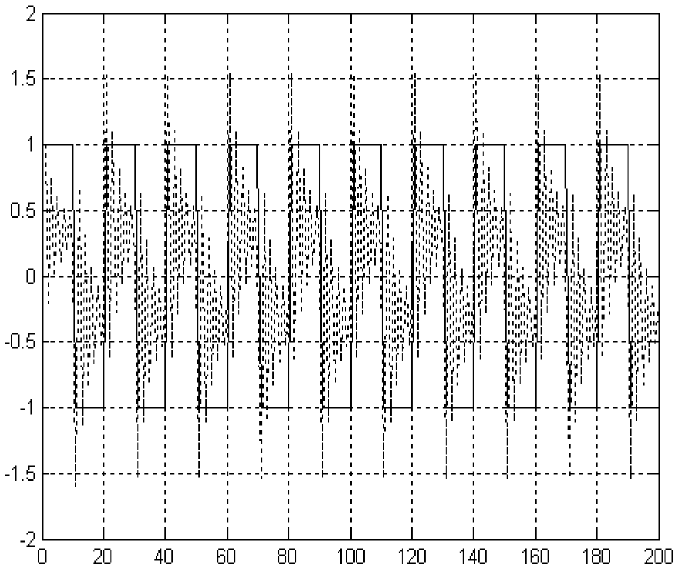


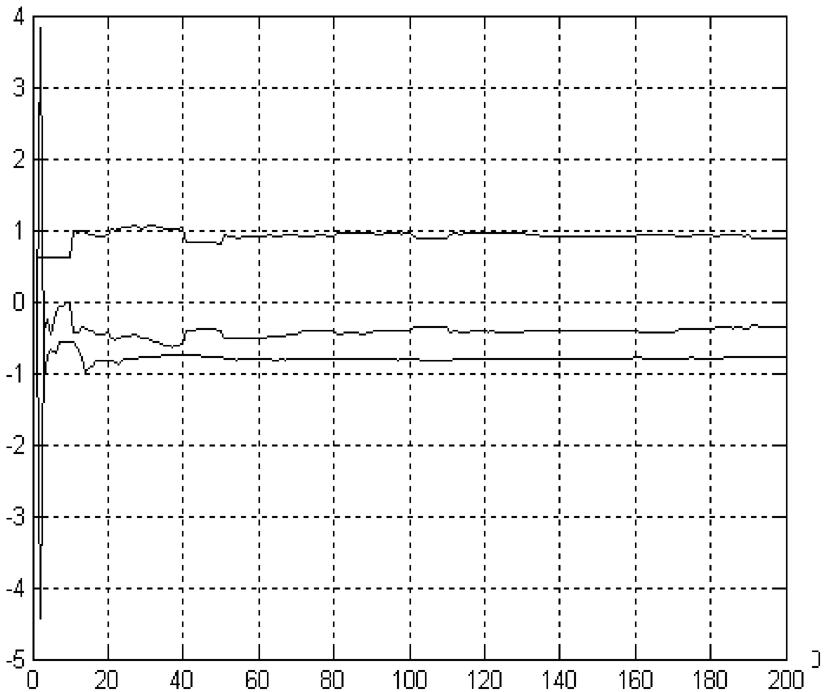
In general, the different identification algorithms are based on the minimization of a ***gap or distance*** *between the system and model.*

(Ljung, 1993)

Simple example: Recursive Least square (RLS) method

- Function : $J = \sum_{i=1}^{N} \left(z_i - f(x_i, \theta)\right)^2$
- The system to be identified is given by: $G(z) = \dfrac{1 - 0.4z^{-1}}{1 + 0.8z^{-1}}$
- The parameters to be estimated are : -0.4, 0.8 and 1
- Input/output and parameter estimation



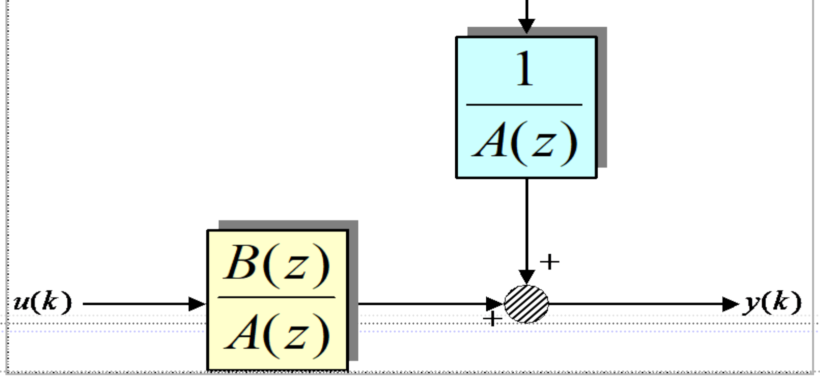


Mechbal N.

Model structure

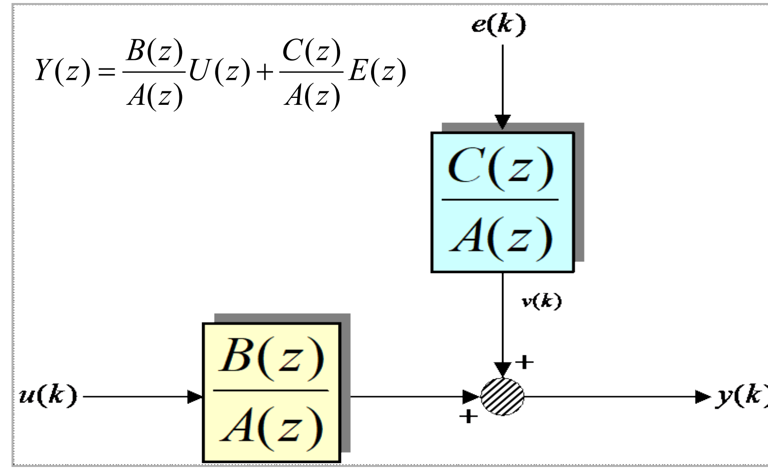
ARX : "AutoRegressive eXogenous"

$$Y(z)=\frac{B(z)}{A(z)}U(z)+\frac{1}{A(z)}E(z)$$



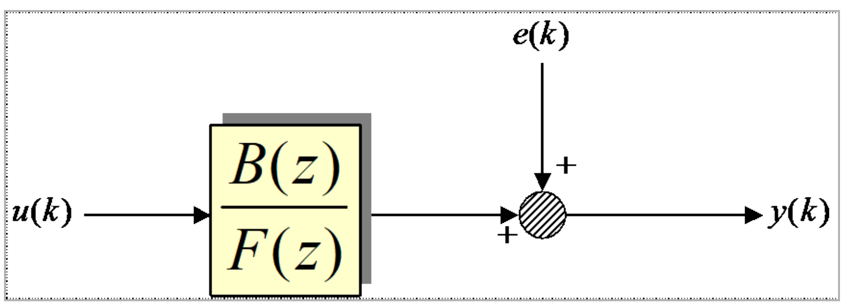
■ ***ARMAX : "Moving Average"***

$$Y(z)=\frac{B(z)}{A(z)}U(z)+\frac{C(z)}{A(z)}E(z)$$



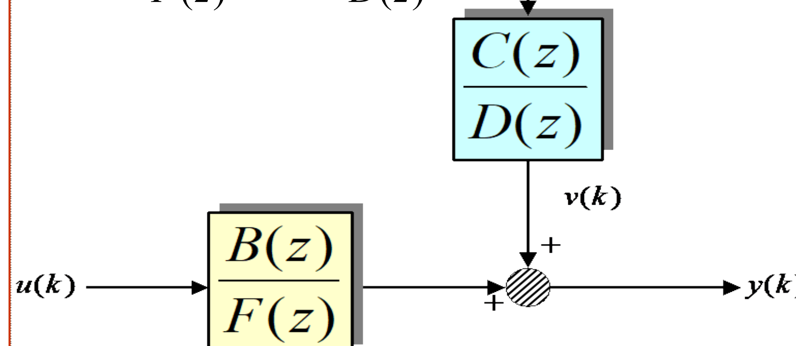
■ ***OE : "Output Error "***

$$Y(z)=\frac{B(z)}{A(z)}U(z)+E(z)$$



■ ***BJ : "Box et Jenkins "***

$$Y(z)=\frac{B(z)}{F(z)}U(z)+\frac{C(z)}{D(z)}E(z)$$



Mechbal N.

State Estimation -- System Identification

- **Discrete nonlinear, non-Gaussian multivariable state space model**

state *Process noise*

$$x_{k+1} = f(x_k, u_k, v_k, \theta_k)$$

Model Parameters

$$y_k = g(x_k, u_k, \omega_k, \theta_k)$$

observation *Known input* *Measurement noise*

- **Problems:**
  - Estimation of the state
  - Estimation of static parameters present in dynamic models. This is accomplished using the information available in measured input and output signals from the underlying system.
  - Presence of non-Gaussian noise
  - MIMO system difficult to handle with classical approach

Mechbal N.

State Estimation -- System Identification

- **Linear state space model with Gaussian additive noise**

state *Process noise*

$$x_{k+1} = A(\theta_k)x_k + B(\theta_k)u_k + v_k$$
$$y_k = C(\theta_k)x_k + D(\theta_k)u_k + \omega_k$$

observation *Known input* *Measurement noise*

- **Mixed Linear/Nonlinear state space model:** Rao-Blackwellization

Nonlinear state *Process noise*

$$x^n_{k+1} = A^n(\theta_k, x^n_{k+1})x_k + B^n(\theta_k, x^n_{k+1})u^n_k + v^n_k$$

Linear state

$$x^l_{k+1} = A^l(\theta_k, x^n_{k+1})x_k + B^l(\theta_k, x^n_{k+1})u^l_k + v^l_k$$

Known input

$$y_k = C(\theta_k, x^n_{k+1})x_{k\,k} + \omega_k$$

observation *Measurement noise*

Conditioned on the nonlinear states, the model described above is linear

Mechbal N.

How ?

- Need specific techniques

- The difficulty of estimating nonlinear systems are widely recognized (Ljung, 2008; Ljung & Vicino, 2005; Ljungbode, 2003).

- Intensive research, generally focuses on specific system classes such as those described by Volterra kernels (Bendat, 1990), neural networks (Narendra & Parthasarathy, 1990), nonlinear ARMAX (NARMAX) (Leontaritis & Billings, 1985), and State space structures (Van der Merwe 2004, Doucet and Tadic, 2003), ….

In this course

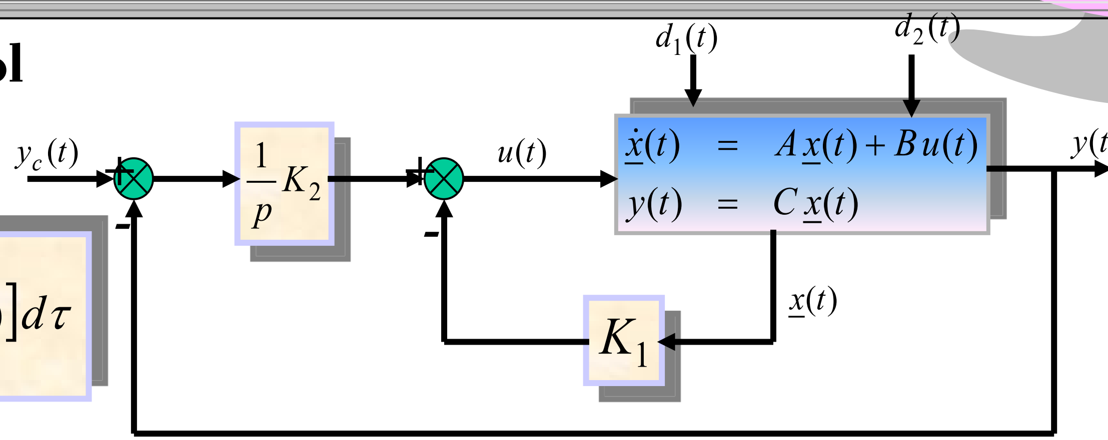
- Linear case: **LS, Maximum Likelihood, Bayesian KF estimators … and Subspace method for MIMO system** **(depending on the progressing)**
- Nonlinear case: **Bayesian estimator** (EKF, UKF) and **Particle methods** (SMC)

- **Subspace methods (if we have time !!)**
  - Subspace based identification methods estimate the state-space sequence using input/output signals. (Van Overchee & De Moor, 1996)
  - Types:
    - Output Only
    - Input/Output
  - Advantage
    - Possibility of real-time implementation
    - Applied to MIMO system
    - No parametrization is needed.
    - No model order is needed in advance.
    - The optimization problem is a convex LS problem.
  - Obstacles
    - Singular Value Decomposition (SVD) is computationally complex
    - Spurious modes
  - Existing Solutions
    - Recursive subspace identification methods (Delgado and Dos Santos, 2004; Mercère et al, 2004; Pongpairoj and Pourboghrat, 2006)
    - Stabilization Diagrams
    - Reinforced stabilization diagrams (Cauberghe et al, 2005; Fan et al, 2007)

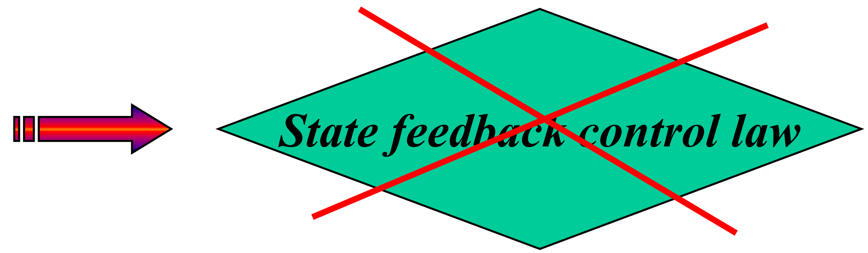
STATE ESTIMATION

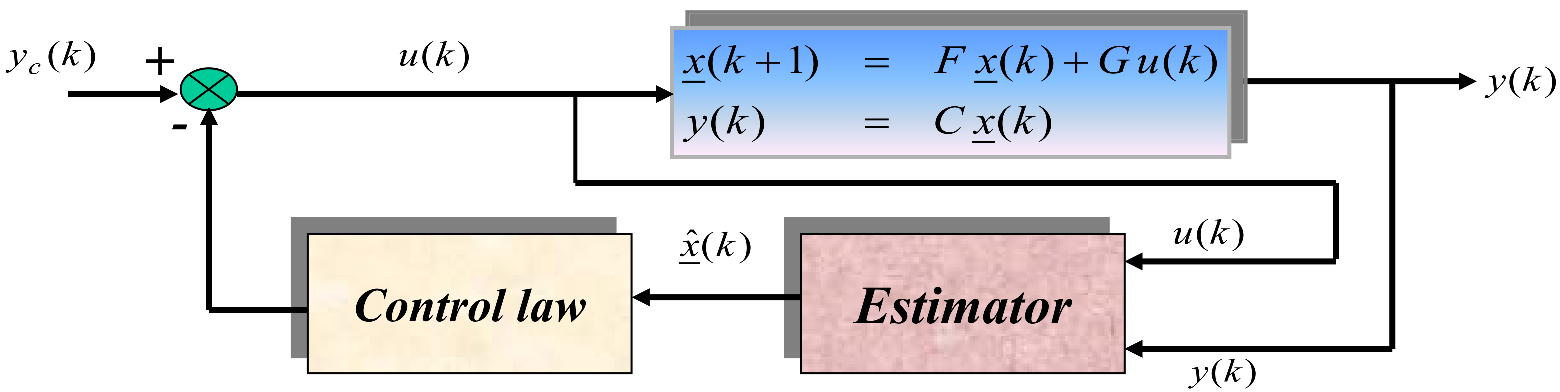
✷ State estimation for control

$$u(t) = -K_1\, \underline{x}(t) - K_2 \int_0^t \left[y(\tau) - y_c(\tau) \right] d\tau$$



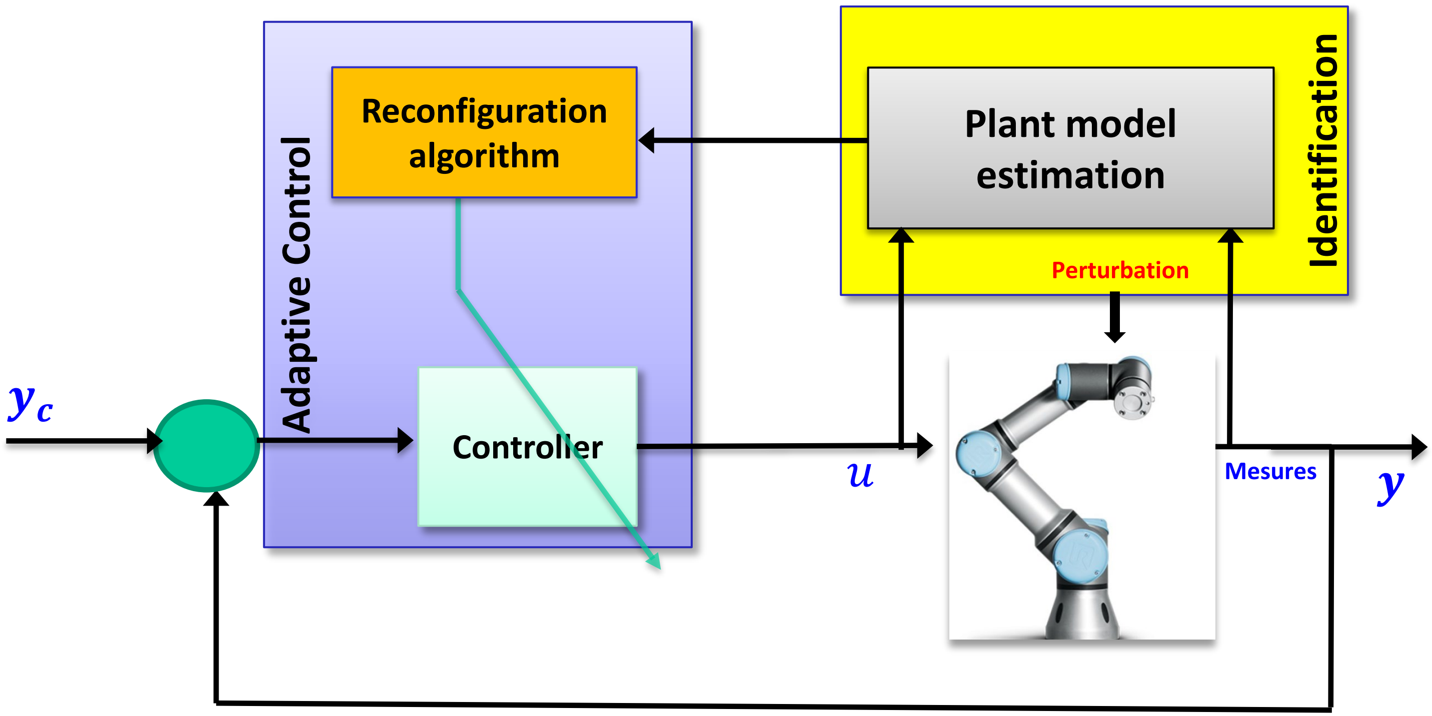
Some measurements are missing





Mechbal N.

✷ **Parameter/State identification → Adaptive control**



Mechbal N.

State Estimation -- System Identification

- **Nonlinear identification towards Nonlinear State estimation**
  - The strategy employed is rather well-known. The idea is ***to augment the states with the parameters into a new state vector.***
  - By assuming a **random walk** variation for the parameters, the system identification problem can now be cast as a nonlinear state estimation problem, which opens up for possible use of all algorithms available for this problem.
  - The new state and model are:

$$\underline{x}^a(t) = \begin{bmatrix} \underline{x}(t) \\ \underline{\theta}(t) \end{bmatrix}$$

Random walk

$$\underline{x}^a_{k+1} = \begin{bmatrix} \underline{x}_{k+1} \\ \underline{\theta}_{k+1} \end{bmatrix} = \begin{bmatrix} F(\underline{x}_k, \underline{\theta}_k, \underline{u}_k) \\ \underline{\theta}_k \end{bmatrix} + \begin{bmatrix} 0 \\ \underline{q}_k \end{bmatrix}$$

The parameter is static, but we suppose that it is corrupted by a Gaussian noise in order to enhance the identification:

A target must move so that the hunter can spot it !!!

Mechbal N.

State Estimation -- System Identification

Nonlinear identification towards Nonlinear estimation

- ***Dual Estimation* problem:** Separate state-space representation is used for the state and the parameters. Two estimators are run simultaneously for $\underline{x}$ and $\underline{\theta}$.

  Iteratively, *the current estimate of $\underline{\theta}$ is used in the state filter as a given input and likewise the current estimate of the state is used in the parameter filter. The two coupled state representation used re given by:*

$$x_{k+1} = f(x_k, u_k, v_k, \hat{\theta}_k)$$
$$y_k = h(x_k, u_k, \omega_k, \hat{\theta}_k)$$

Estimated parameters

$$\theta_{k+1} = \theta_k + q_k$$
$$y_k = h(\hat{x}_k, u_k, \omega_k, \theta_k)$$

Estimated state

- ***Joint estimation problem:*** The strategy employed is rather well-known. The idea is ***to augment the states with the parameters into a new state vector.*** The new state to be estimated and model are:

$$\underline{x}^a(t) = \begin{bmatrix} \underline{x}(t) \\ \underline{\theta}(t) \end{bmatrix}$$

$$\underline{x}^a_{k+1} = \begin{bmatrix} \underline{x}_{k+1} \\ \underline{\theta}_{k+1} \end{bmatrix} = \begin{bmatrix} F(\underline{x}_k, \underline{\theta}_k, \underline{u}_k) \\ \underline{\theta}_k \end{bmatrix} + \begin{bmatrix} 0 \\ \underline{q}_k \end{bmatrix}$$

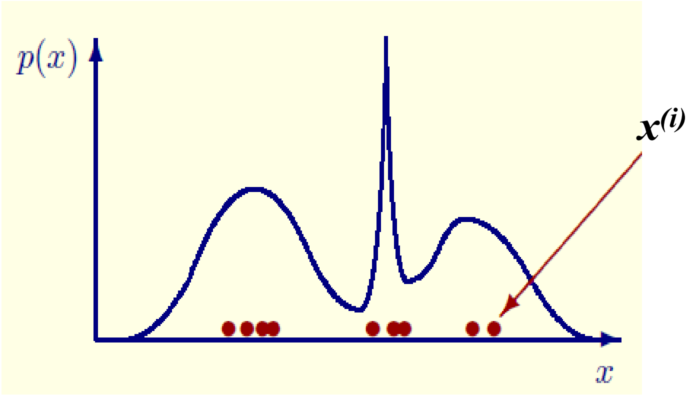
Mechbal N.

STATE ESTIMATION -- SYSTEM IDENTIFICATION

- **Bayesian estimators:**
  - Kalman filter (KF): linear model and Gaussian noise
  - Extended Kalman filter (EKF): non-linear model (low nonlinearities, approximation of order 1) + Gaussian noise
  - Unscented Kalman filter (UKF): non-linear model (linearization of order 2) + Gaussian noise

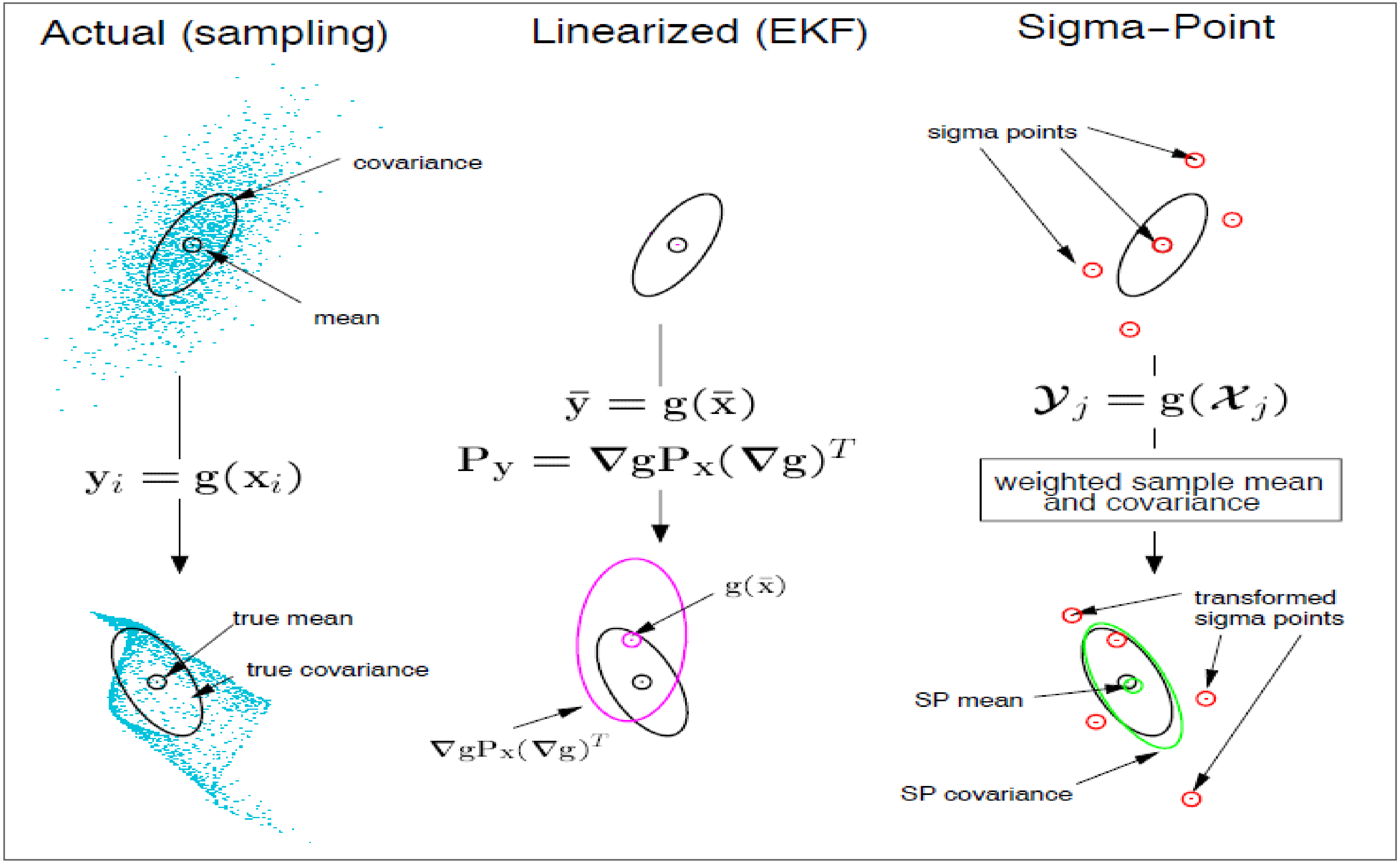
  These estimators make an approximation of the posterior density by a Gaussian distribution. To overcome this limitation, particle filters and their variants have been developed

  - Sequential Monte Carlo Methods (Particle filter): non-linear model , non-Gaussian noises, ….



Bayesian estimators : UKF

Van der Merwe 2004



Mechbal N.

Particle filter

- In recent years, statistical simulation methods have emerged as approaches to make significant advances in areas as diverse as speech processing, bioinformatics, prosecution and the location, the vision , finance, diagnostic systems,...

- The particle filtering methods, simulation methods are sequential Monte-Carlo, in which particles:

  - *Move independently*

  - *interact through selection algorithms (**called resampling**) that channel the particles in the regions of interest.*

- They exploit the information accessible on the processes and their evolution laws.

- Require the generation of several weighted samples or particles, simulating the real trajectory of the system.

Problem statement

- Probabilistic inference is the problem of estimating the hidden variables (states or parameters) of a system in an optimal and consistent fashion as a set of noisy or incomplete observations of the system becomes available online.

- The optimal solution to this problem is given by the recursive Bayesian estimation algorithm which recursively updates the posterior density of the system state as new observations arrive.

- This posterior density constitutes the complete solution to the probabilistic inference problem and allows us to calculate any "optimal" estimate of the state.

Recursive Bayesian estimation

$$\underbrace{p(x_k/Y_k)}_{\text{posterior}} = \frac{\overbrace{p(y_k/x_k)}^{\text{likelihood}}\,\overbrace{p(x_k/Y_{k-1})}^{\text{prior}}}{\underbrace{p(y_k/Y_{k-1})}_{\text{evidence}}}$$

- **Likelihood** expresses the belief in the ***occurrence of an event*** it is defined in terms of ***observation model***
- **Prior** is given by the propagation of past state into future before new observation is made

$$p(x_k/y_{1:k-1}) = \int p(x_k/x_{k-1})\,p(x_{k-1}/y_{1:k-1})\,dx_{k-1}$$

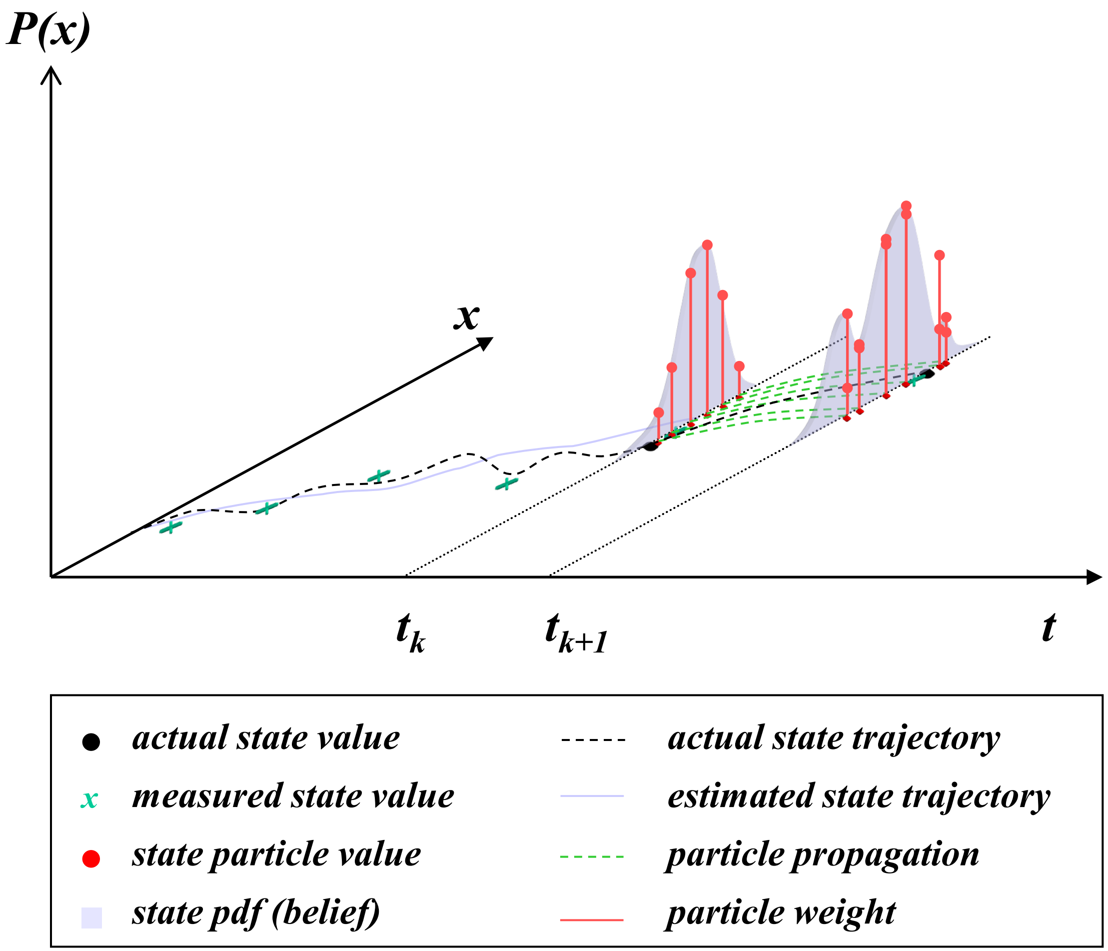
$p(x_k/x_{k-1})$: it is the transition density given by the process model

- **Evidence** :

$$p(y_k/y_{1:k-1}) = \int p(y_k/x_k)\,p(x_k/y_{1:k-1})\,dx_k$$

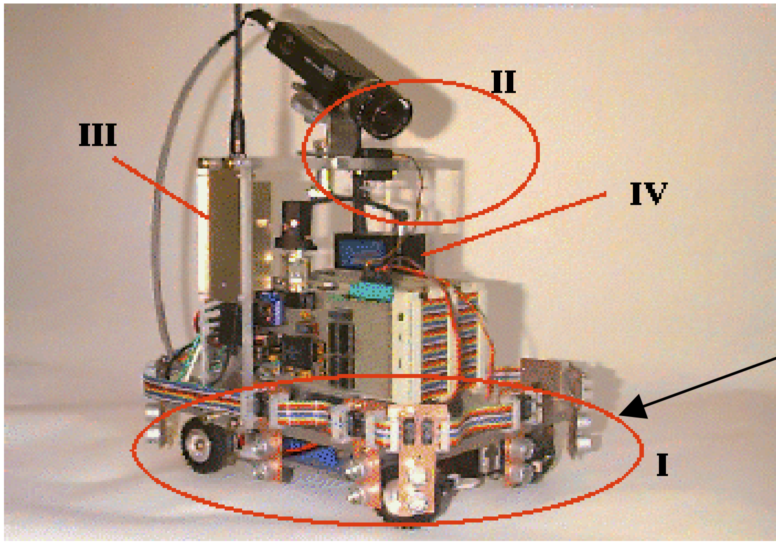
- **In its simplest version, the FP method consists in:**
  - Between two instants of observation, the particles move independently according to the **dynamics** of the hidden state.
  - As soon as a new **observation is available**, a resampling occurs, or the particles are selected for their relevance to the new observation (quantified by the likelihood function).
  - Under the effect of **resampling**, which is the essential step of the method: the particles are concentrated automatically the regions of interest of the state space.
  - The **method is very easy** to implement, since it is sufficient to simulate trajectories independently of the state cache, the interaction taking place only when resampling.

Particle Filtering

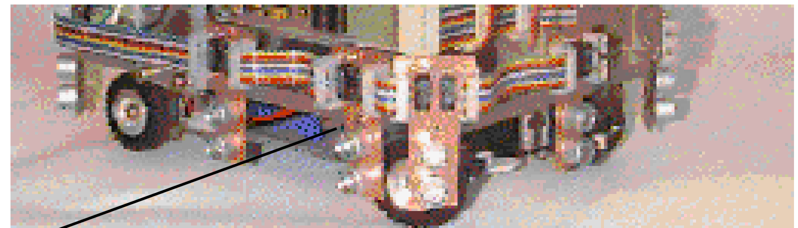


- *represent state as a pdf*
- *sample the state pdf as a set of particles and associated weights*
- *propagate particle values according to model*
- *update weights based on measurement*

✷ Example : Ultrasonic telemeter

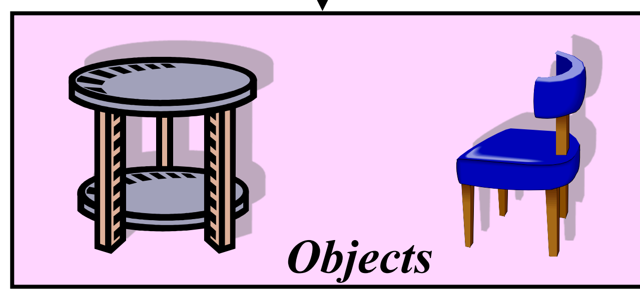


Ultrasonic belt



Distance measurement

d



Goal:

Determine the distance "d" between the robot and the object.

Example : Ultrasonic telemeter

- Principle : distance measurement based on flying time of the wave
- In theory: the distance between the robot and the obstacle:

$$t = \frac{2d}{C}$$

***t* : time between the emission and reception**

***C* : Celerity of the wave**

- In practice: Presence of random uncertainties and noises, ***b*** :

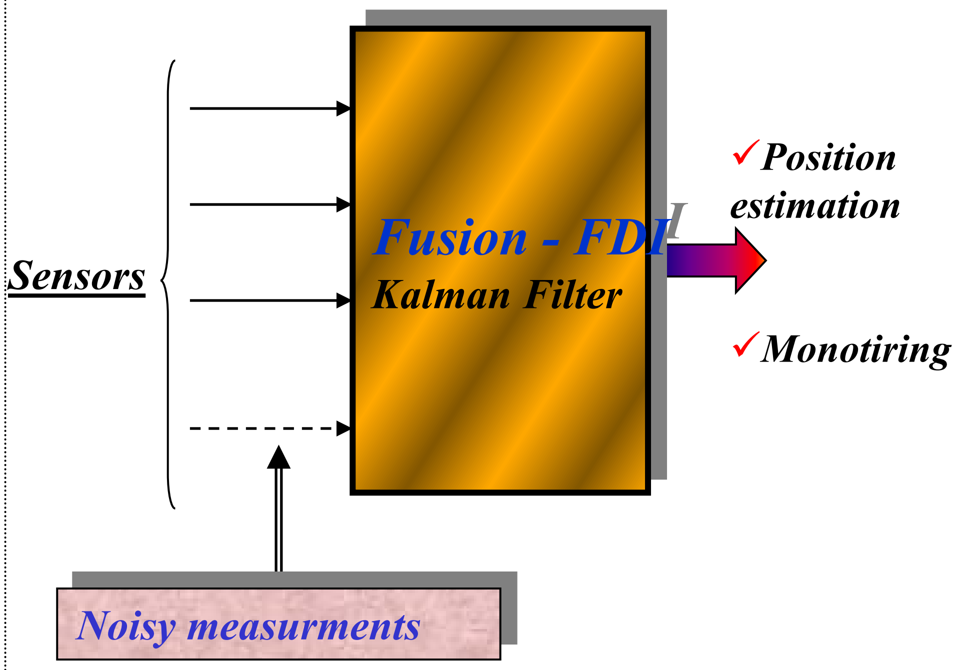
$$t = \frac{2d}{C} + b$$

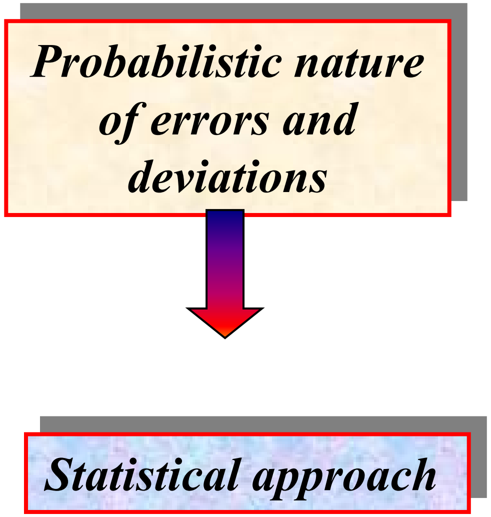
From the measurement of t, we can only seek an approximation of the distance d

Estimation $\hat{d}$ of d

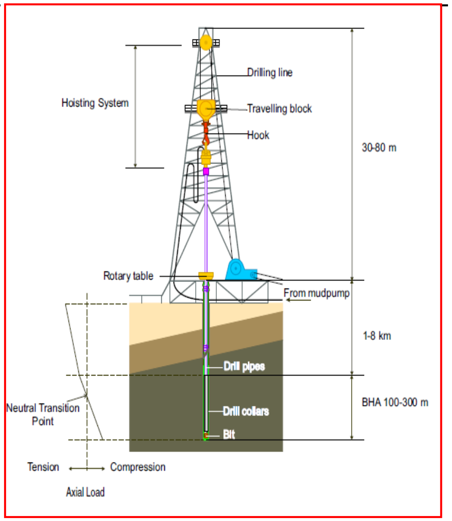
Mechbal N.

- All sensors present drift and possible failures, We must then add to them modules that performing measurements updating and monitoring the registration and monitoring





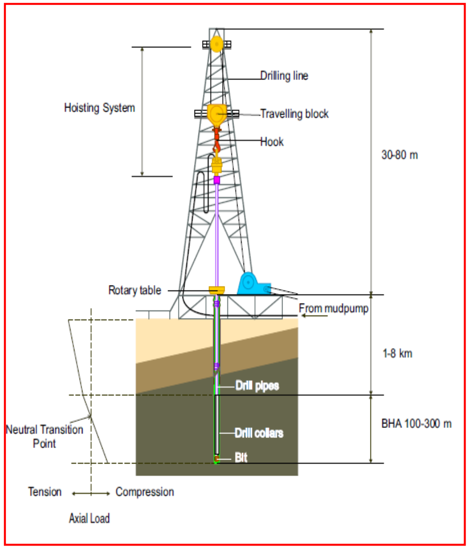
Some introductive xamples

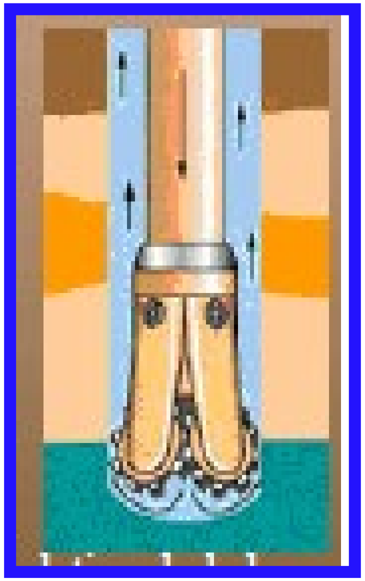


Introductive Examples

Drilling process:

- Drilling processes are very complex owing to the unpredictable environments that they are confronted. Adverse conditions of rocks, problems of cutting transportation, and the impossibility of observing the process during the drilling operation make the deep hole drilling one of the most difficult tasks. Cuttings transportation during drilling operations is vital to enable a correct sequence of operations.
- As a consequence, the hydraulic process integrated in the drilling system plays an essential role; it allows fluid transportation and consequently the evacuation of cuttings which come from the drilling operation. This task is achieved from downhole to a filtering system located at the surface in closed loop form.
- In addition, the drill bit cleaning as well as its cooling is carried out through this fluid circulation. During cuttings transportation the **plugging in the hydraulic system might occur**. When this situation happens the drilling process can not be achieved as expected, consequently a lost of time becomes unavoidable owing to the required maintenance procedure.



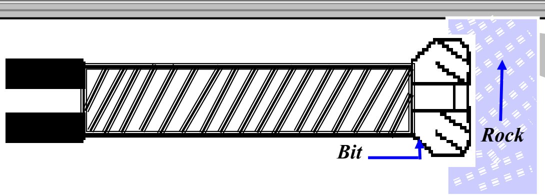


Mechbal N.

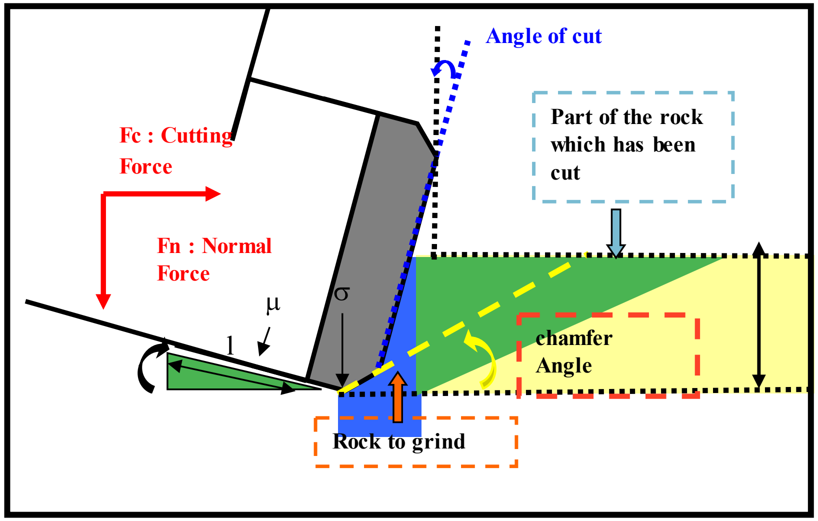
✷ Context of the study: balling, plugging

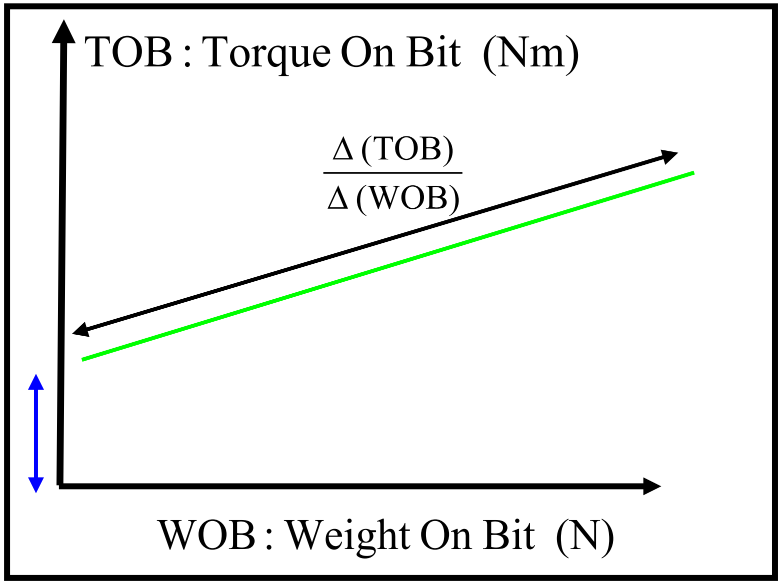
- Cuttings stick to the bit surface and cannot be washed away by the mud circulation. At the end the bit is completely covered by the stuck cuttings and the bit is no longer in contact with the formation

✷ Bit rock interaction model



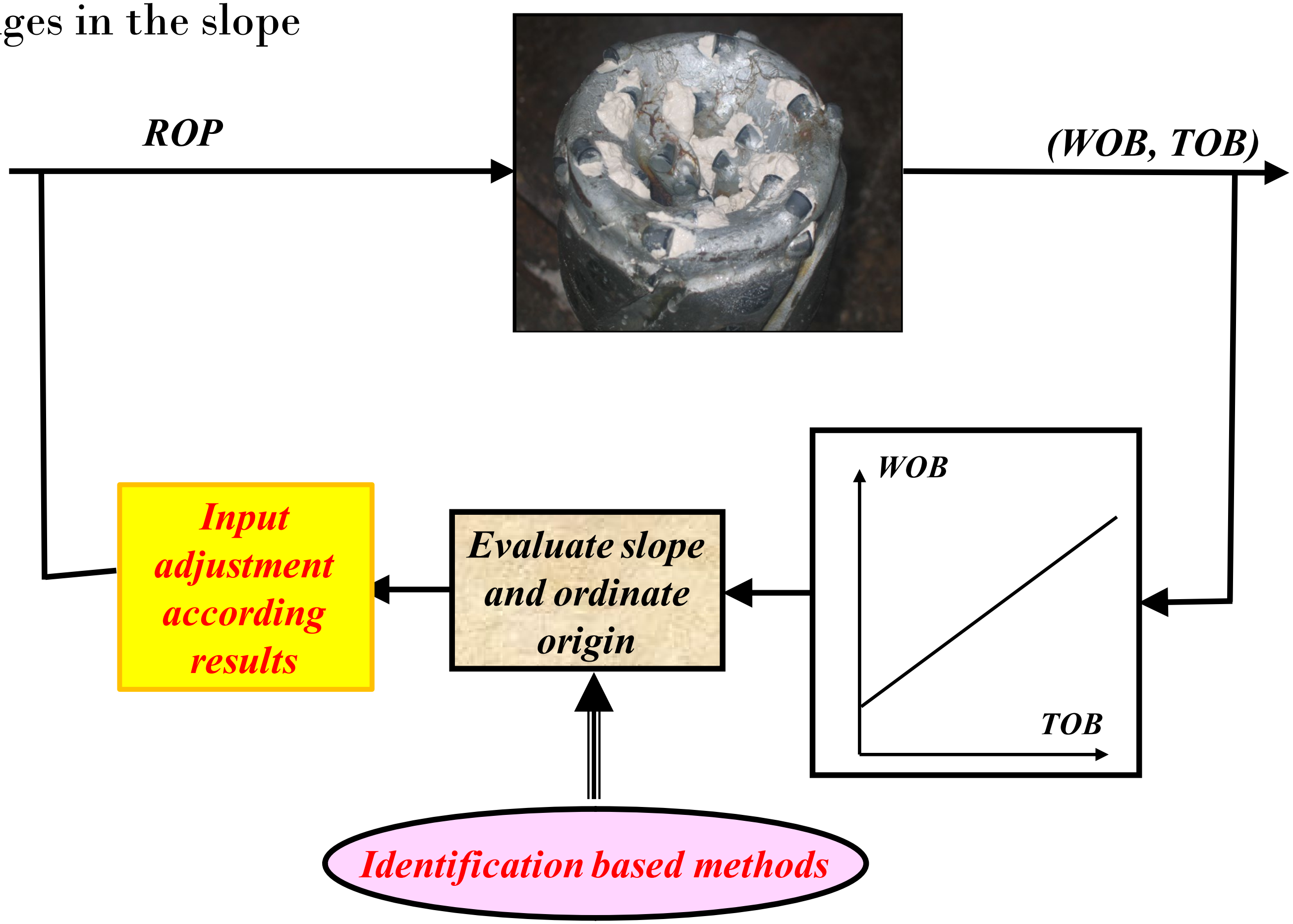
- Detournay's model (2002): the slope is only dependent on the bit shape factor and the friction coefficient. Hence, the rock properties are not involved.



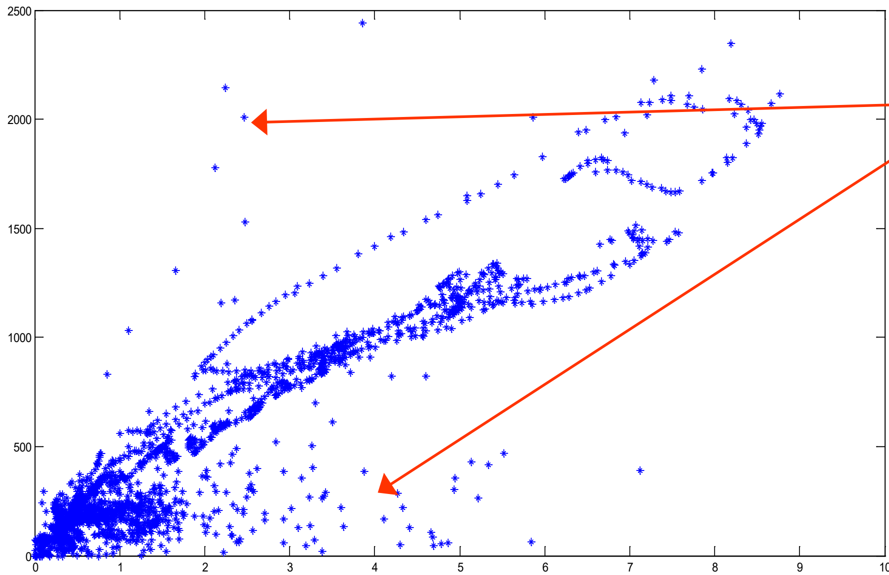


Strategy :

- When a change occurs on the bit, for example the bit wear or break the slope changes, the fault detection approach developed is based on the identification of changes in the slope



Test campaign and data processing



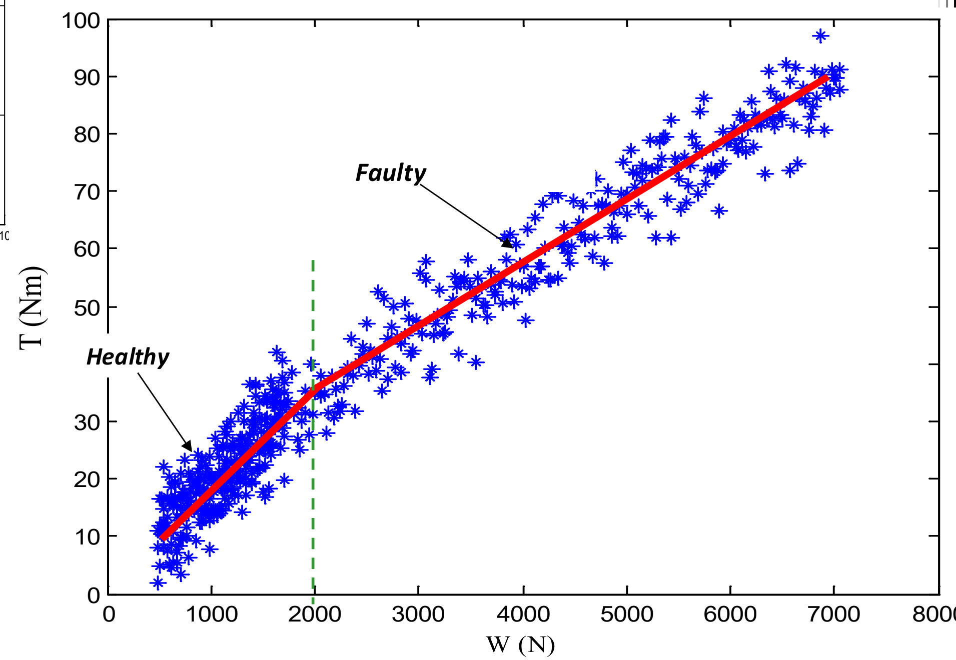
Raw measurements

Remarks:

-Outliers measurements

- No time indication in this specific space

A first signal processing and slope changes

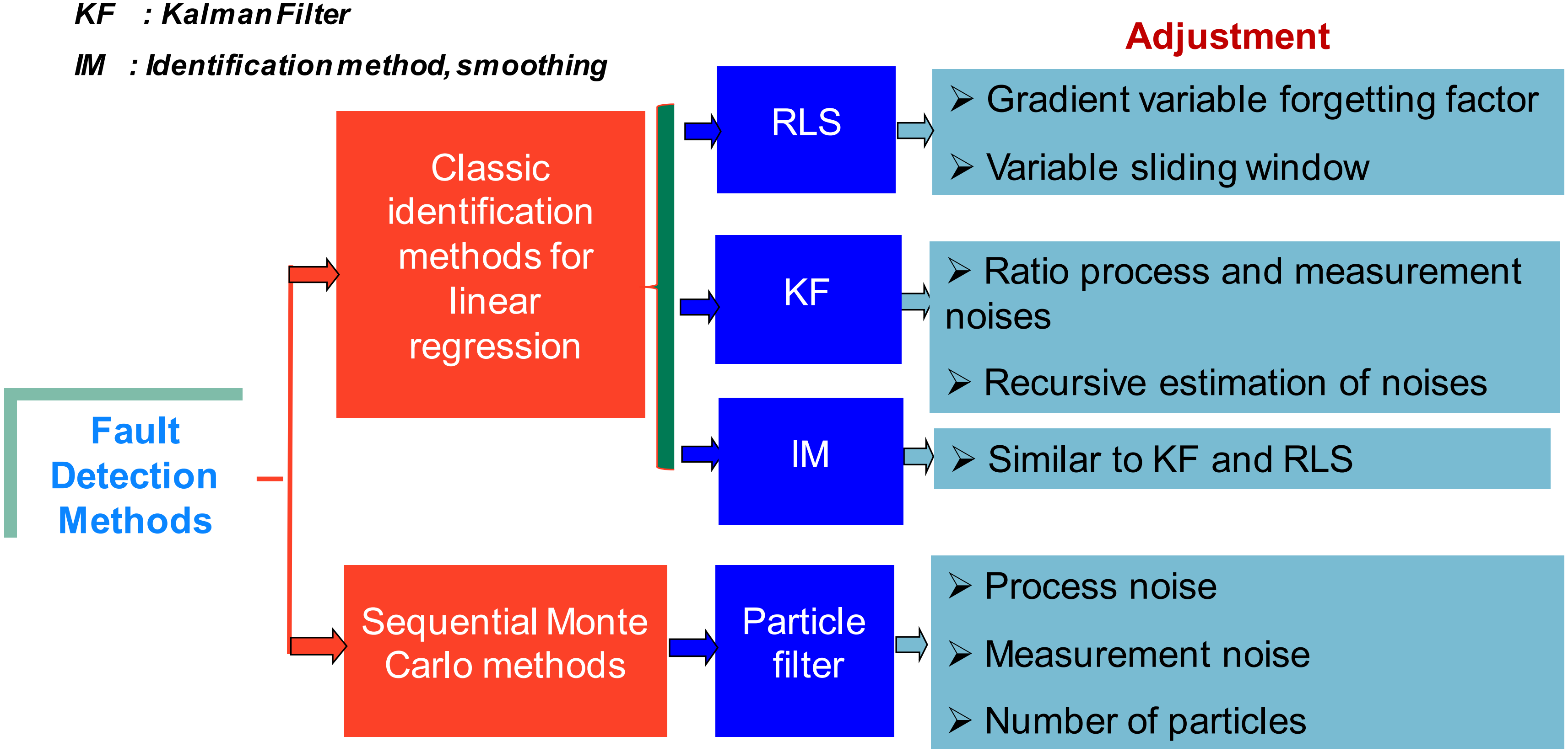


Methods for slope identification

RLS : Recursive Least Square method

KF : Kalman Filter

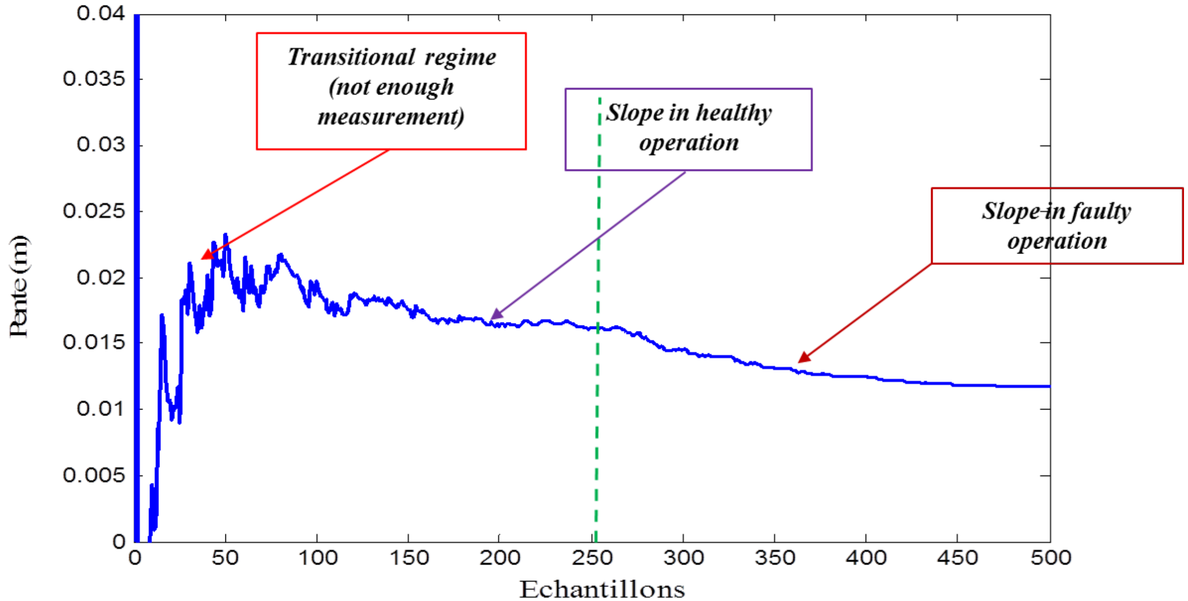
IM : Identification method, smoothing



Mechbal N.

Results with Rao-blackwellised Particle Filter (RBPF)

- Temporal evolution of the Slope given by the estimator

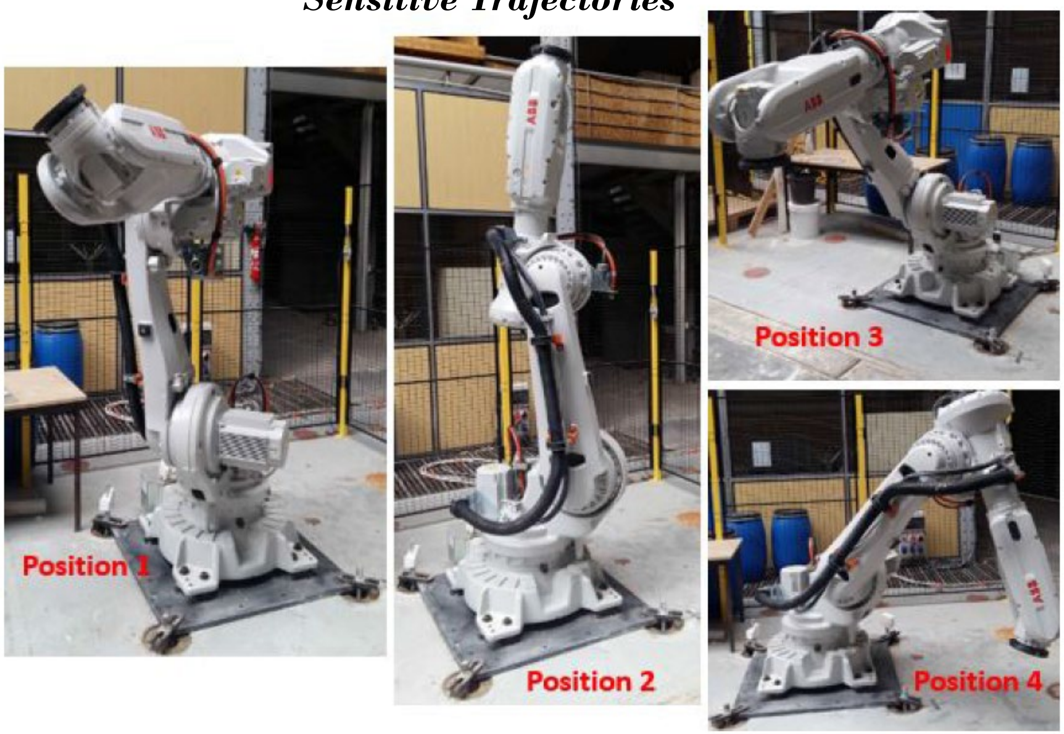


Mechbal N.

Identification of the dynamic model of a robot

- Rigid or flexible joints ?
- Open or closed loop identification approach ? Off-line or On-line ?
- The joint positions are not measured on industrial robots.
- Manufacturers do not want to give the control laws of their controllers
- External external measurement → Expensive !

Sensitive Trajectories



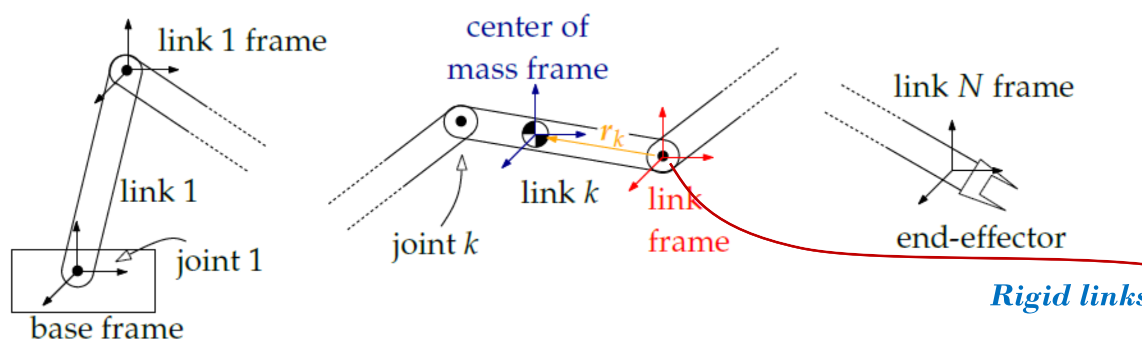
Dynamic Models Robot Identification

✷ Rigid body and rigid links dynamic model

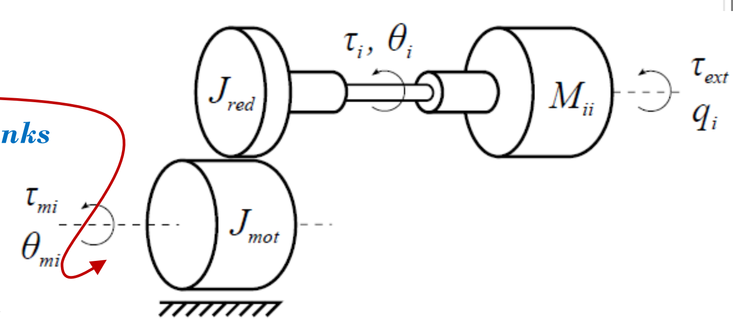
- The dynamic model of a rigid robot with N moving relates the full motion of its joints—positions, velocities and accelerations ($\boldsymbol{q}, \dot{\boldsymbol{q}}$ *and* $\ddot{\boldsymbol{q}}$) — with the ($N \times 1$) vector of forces, $\boldsymbol{\tau}$ being applied to those joints:

$$\boldsymbol{M_g}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})\dot{\boldsymbol{q}} = \boldsymbol{\tau}$$

- $\boldsymbol{M_g}(\boldsymbol{q})$ is the generalized inertia matrix of the robot
- $\boldsymbol{C}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is the vector modelling the disturbances or perturbations, including friction forces, gravity effects and other non-linearities depending on the studied robot.



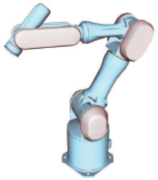
Under the hypothesis of ideal transmissions and rigid joints, the joint and motor positions are confounded $\theta = q$.



*- **E**quations can be expressed in a linear form with respect to the physical parameters. This property is fundamental for the identification and adaptive control.*

- The inertia matrix $M_{rigid}(q)$ is symmetrical, defined positive.

✷ Rigid body and flexible links dynamic model

Robot	Spécifications*	Technologie d'actionnement et instrumentation
 KUKA-DLR LWR	$m_r = 16$ kg $m_c = 7$ kg $l^p = 0.936$ m	• réducteurs Harmonic drive® • jauges de contrainte intégrées, capteurs de position moteurs et articulaires
 Barrett WAM	$m_r^{\text{bras}} = 5.8$ kg $m_r^{\text{total}} = 27$ kg $m_c = 3$ kg $l^p = 1$ m	• transmissions à câbles sans engrenages • capteurs de position moteurs (et articulaires)
 Mitsubishi PA10-7CE	$m_r = 38$ kg $m_c = 10$ kg $l^a = 0.93$ m	• servo-moteurs AC, réducteurs Harmonic drive® • capteurs de position moteurs
 Bras ASSIST (CEA-LIST)	$m_r = 9.3$ kg $m_c = 3$ kg $l^a = 0.8$ m	• transmissions à câbles, moto-réducteurs • capteurs de position moteurs

(*) : m_r - masse du robot, m_c - charge utile
l - dimension caractéristique (distance maximale entre articulations l^a ou

(Makarov, 2013)

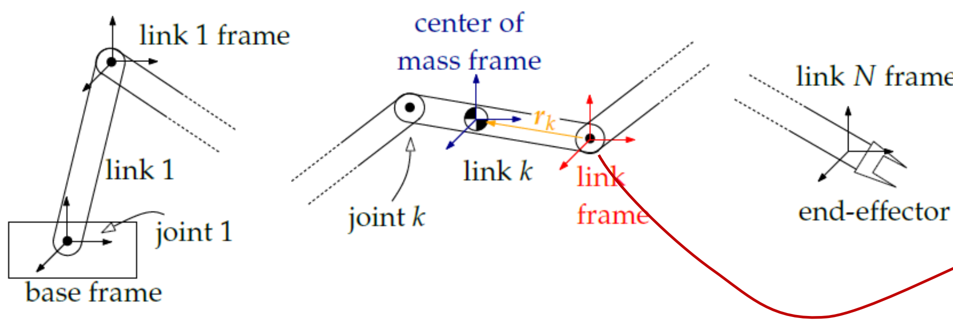
Rigid body and flexible links dynamic model

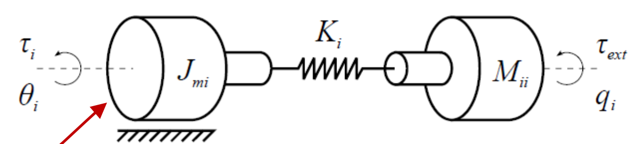
- By considering some simplifying assumptions (e.g. inertial couplings between the motors and the robot segments are neglected), we obtain the following reduced dynamic model

$$\boldsymbol{M_g}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{C}(\boldsymbol{q},\dot{\boldsymbol{q}})\dot{\boldsymbol{q}} + \boldsymbol{K}(\boldsymbol{q}-\boldsymbol{\theta}) = \boldsymbol{\tau_{ext}}$$
$$\boldsymbol{J_m}\ddot{\boldsymbol{\theta}} - \boldsymbol{K}(\boldsymbol{q}-\boldsymbol{\theta}) = \boldsymbol{\tau_{mot}}$$

- $\boldsymbol{K}$ is the diagonal matrix of joint stiffnesses

For very high joint stiffnesses ($K \to \infty$), the robot is considered to be entirely rigid: $\theta = q$.





Flexible links

- *In the case of a fully rigid robot, a number of <u>control inputs equal to the number of degrees of freedom of the robot is available</u> and knowledge of the output variables of the drive shafts is sufficient to measure the state of the system,*
- *This is no longer the case for a robot with flexible joints, which has additional degrees of freedom due to <u>elastic deformations</u>. The state of the system consists of both the motor variables and the articular variables, and more complete instrumentation is required to access the complete state.*

Mechbal N.

Inverse dynamic model (IDM) – IDMI (Identification)

- The goal is to identify the standar dynamic parameters, $\boldsymbol{\theta}$, of a robot

$$\boldsymbol{\tau} = \boldsymbol{IDM}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \ddot{\boldsymbol{q}})\boldsymbol{\theta}$$

- Well known and used approach: the **IDMI-LS approach**

$$\boldsymbol{Y} = \boldsymbol{\Phi}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \ddot{\boldsymbol{q}})\boldsymbol{\theta} + \boldsymbol{e}_{LS}$$

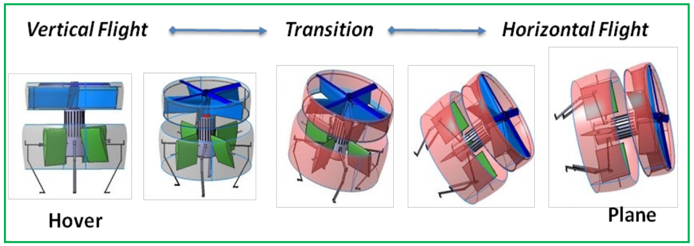
  - $\boldsymbol{Y}$: measurement vector, $\boldsymbol{\Phi}$ is the observation matrix and $\boldsymbol{e}_{LS}$ the random vector of errors of zero mean and variance σ^2,
  - Based on Leas Square optimization. The LS solution is

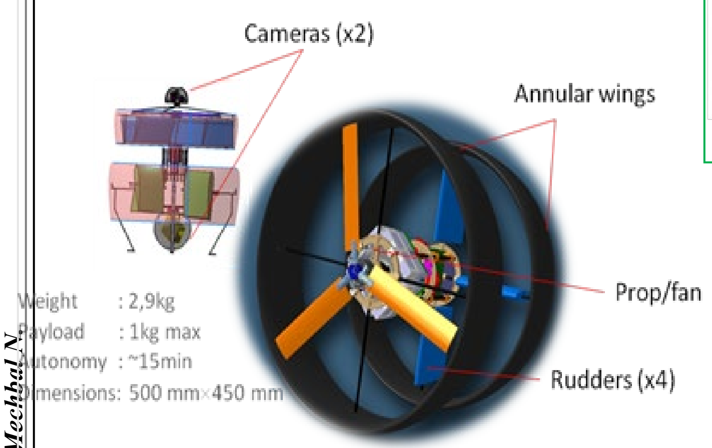
$$\hat{\boldsymbol{\theta}}_{ls} = \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^T\boldsymbol{Y} = \boldsymbol{\Phi}^{\dagger}\boldsymbol{Y} \qquad \boldsymbol{\Phi}^{\dagger} = Pseudo\ Inverse\ matrix$$

- In the case of flexible joints:
  - It is then necessary to identify accurately the stiffnesses to control and simulate precise and reliable motion
  - complementary measurement could be necessary

Nonlinear Estimation for UAV sensors fusion and control

- we have proposed a daring, secure and original autonomous MAV named **AMP**. This UAV has two annular wings that enable both the hovering, thanks to the propeller, and the near horizontal flight at high speed, through the lift of the wings.
- The transition flight is carried by pitching over, to a near horizontal attitude, using 4 rudders. The other benefits of these annular wings are the decrease of the propeller noise signature, its protection against damage and increase autonomy by using buoyancy in forward flying mode.
- Such vehicles have inherent advantages due to hover and plane flights capabilities. These capabilities greatly enlarge the autonomy of the UAV, allowing long time continued missions and also perch and stare maneuvers.
- In plane flight, and thanks to the lift generated by the two annular wings, this UAV presents an estimated gliding of 6-8 and about 80% of energy saving compared to hover mode.





- **Problem:** Loosely coupled GPS/INS integration and State estimation for guidance, navigation with a nonlinear controller
- **Sensor fusion:**
  - It is equipped with sensors and computational devices:
    - Two onboard cameras that can perform aerial photographing and mapping.
    - XSens MTi-G MEMS inertial measurement unit (IMU) for estimating the attitude of the vehicle
    - GPS
    - Altimeter sensor (camera)
    - Gumstix Verdex: Linux-based embedded PC.

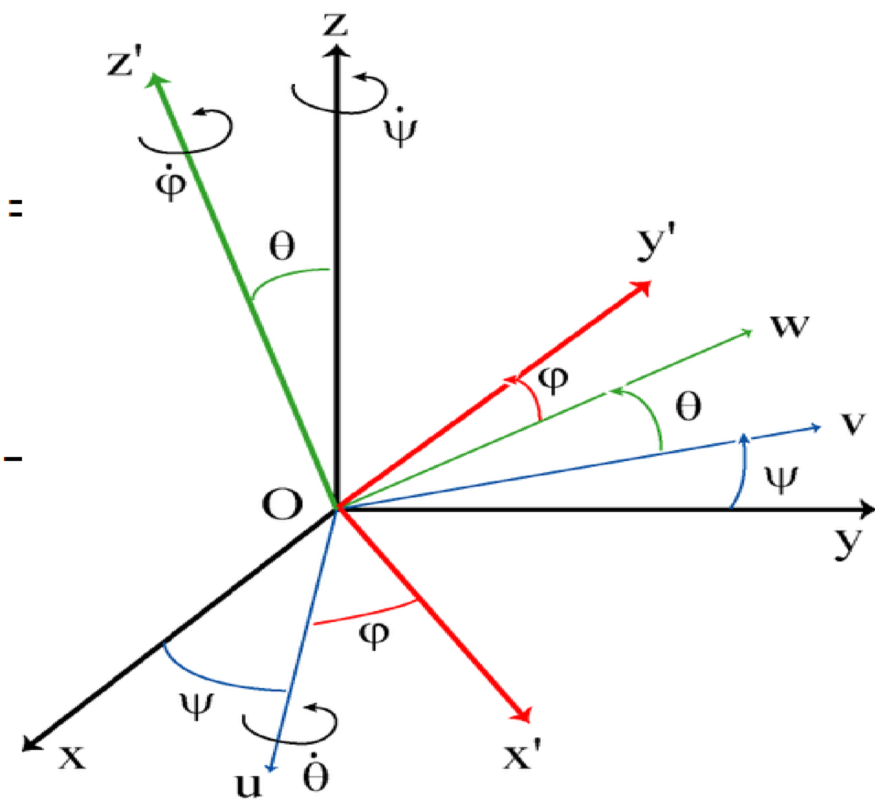
Xsens sensor	**: 25Hz**
GPS	**: 1Hz**
altimeter sensor	**: 1Hz**

Quaternion representation

- Attitude

$\widehat{\Lambda}_{Bb}$ =

ion: rotation from
to B



$\Re_{Bb} = \left(2\Lambda_{Bb0}^{2} -$

Rotation tensor

base

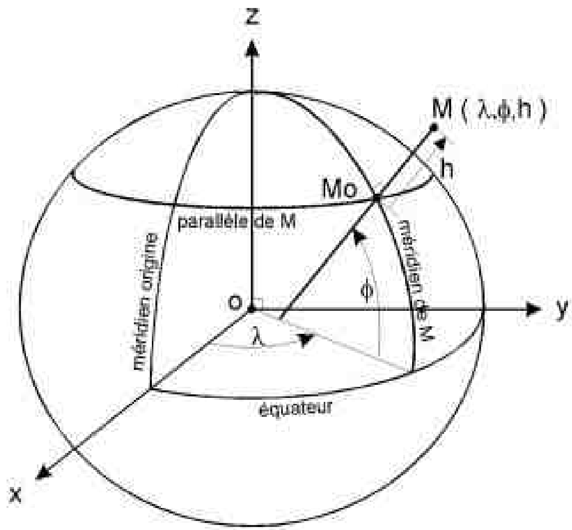
n

Mechbal N.

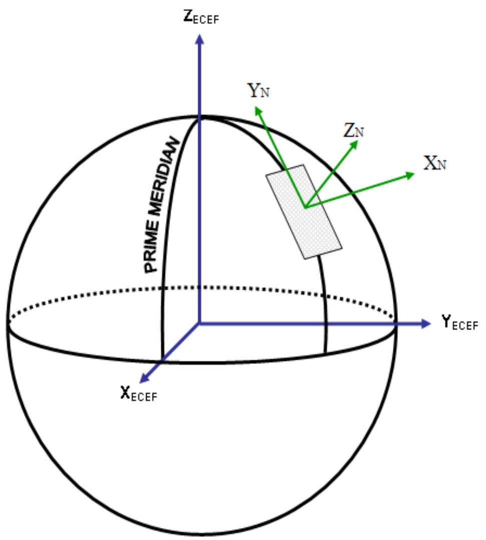
GPS:

- Sensor latency supposed unknown
- Frame transformation

WGS84



ECEF (earth-centered earth fixed)



$$\begin{cases} X_{ECEF} = \left(\frac{a}{\chi} + h\right) cos\phi . cos\lambda \\ Y_{ECEF} = \left(\frac{a}{\chi} + h\right) cos\phi . sin\lambda \\ Z_{ECEF} = \left(\frac{a(1-e^2)}{\chi} + h\right) \sin\ \phi \end{cases}$$

(ϕ, λ, h); latitude, longitude, altitude

Kinematic sensors model

- Positions are estimated by this model and fused with the GPS measurements that are lagged (**sensor latency**)

NON LINEAR

Acceleration Bias

Noise

Acceleration from Xsens

Quaternion

$$\begin{cases} \dot{\underline{p}} = \underline{v} \\ \dot{\underline{v}} = \underline{C_{NB}}.\left(\overline{a} - \underline{b}_a - \underline{q}_a\right) + \begin{bmatrix}0\\0\\1\end{bmatrix}.g \\ \dot{\underline{\hat{\Lambda}}}_{\mathrm{BN}} = \frac{1}{2}.\widetilde{\Omega}_{(\overline{\omega} - \underline{b}_{\omega} - \underline{q}_{\omega})}.\underline{\hat{\Lambda}}_{\mathrm{BN}} \\ \dot{\underline{b}}_a = \underline{q}_{ba} \\ \dot{\underline{b}}_{\omega} = \underline{q}_{b\omega} \end{cases}$$

Noise

Rotation rate from Xsens

Rotation bias

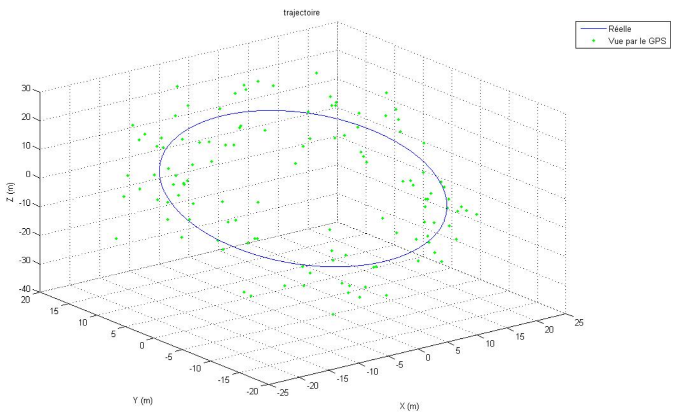
- Discret model

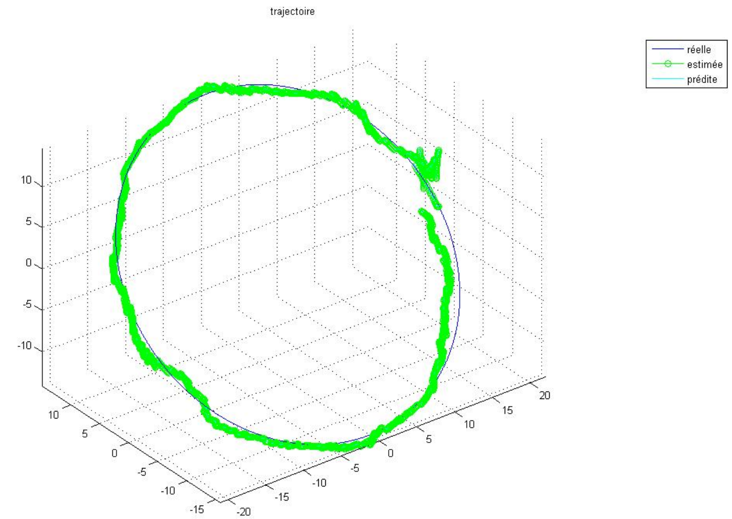
$$\begin{cases} \underline{p}(k+1) = \underline{p}(k) + T_e.\underline{v}(k) + \underline{q}_p \\ \underline{v}(k+1) = \underline{v}(k) + T_e.\left(\underline{C_{NB}}(k).\left(\overline{a}(k) - \underline{b}_a(k) - \underline{q}_a\right) + \begin{bmatrix}0\\0\\1\end{bmatrix}.g\right) + \underline{q}_v \\ \underline{b}_a(k+1) = \underline{b}_a(k) + T_e.\underline{q}_{ba} \\ \underline{b}_{\omega}(k+1) = \underline{b}_{\omega}(k) + T_e.\underline{q}_{b\omega} \\ \underline{\hat{\Lambda}}_{\mathrm{BN}}(\mathrm{k}+1) = \left[(\cos(s) + \eta.\lambda.T_e)I_4 - \frac{1}{2}\Phi_{\Delta}.\frac{\sin(\mathrm{s})}{\mathrm{s}}\right].\underline{\hat{\Lambda}}_{\mathrm{BN}}(\mathrm{k}) + \underline{q}_{\Lambda} \end{cases}$$

Mechbal N.

Simulation results :

- UKF: GPS+Altimeter





- PF : only with GPS

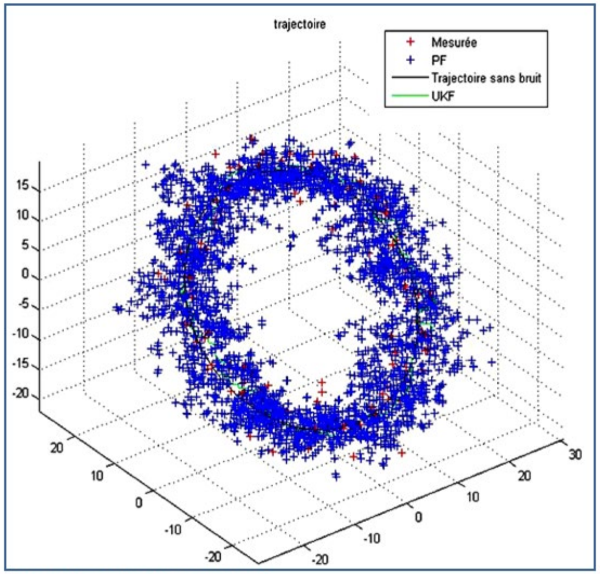


Figure 116: Trajectoire, FP vs GPS

Control model :

- Using the representation of the attitude by quaternion, the equations of motion are described by:

Translation

$$\begin{cases} \dot{x} = v \\ m\dot{v} = -mge_3^0 + F_e + P \end{cases}$$

Rotation

$$\begin{cases} \dot{q} = Q(q)\omega \\ \mathbb{I}\dot{\omega} = -\omega \times (\mathbb{I}\omega) + I_{Rz}\omega_R\left(e_3^1 \times \omega\right) + M_e + u_g \end{cases}$$

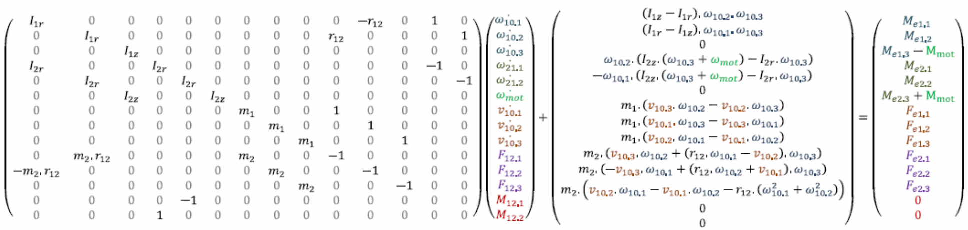
$$P = u_p R e_3^1 = u_p (q \otimes e_3^1 \otimes \bar{q})$$

$$\textit{with} \quad Q(q) = \frac{1}{2}\begin{pmatrix} -\epsilon^T \\ \eta . I_3 + S_\epsilon \end{pmatrix}$$

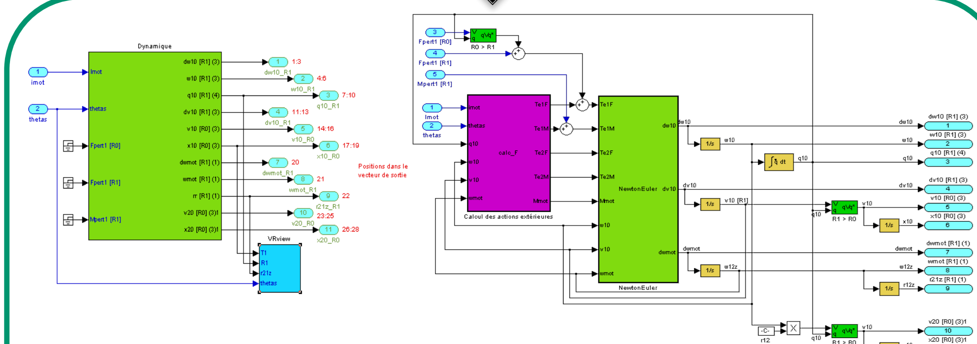
Adaptive backstepping controller

- Backstepping is a nonlinear control design method that provides an alternative to feedback linearization. The present UAV has only one fan and doesn't present any specific orientation in forward flying mode due to its symmetric shape
- In adaptive backstepping, a parameter estimate update law is designed so that closed loop stability is guaranteed.

Steps

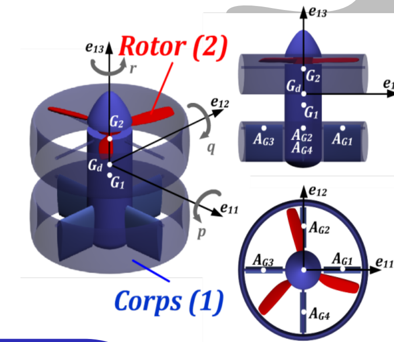


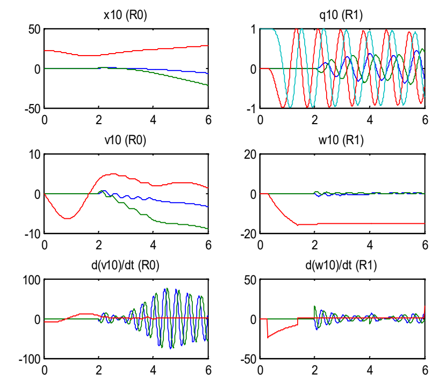
Analytical Madel



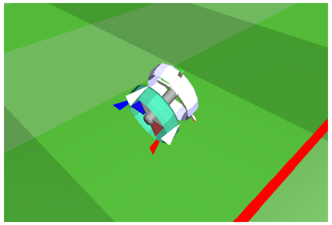
Matlab/Simulink

Parameter estimation from measurement



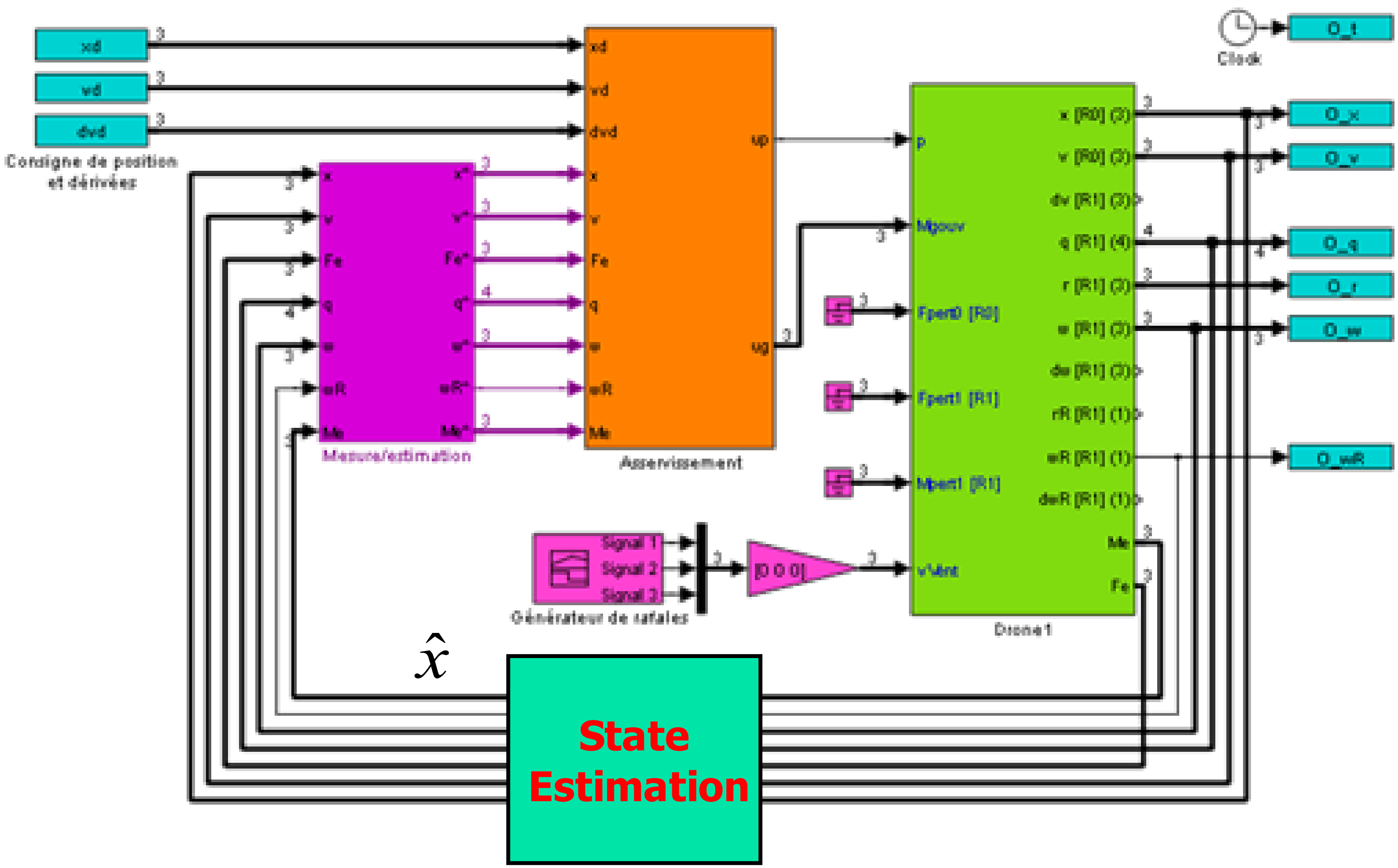


Results

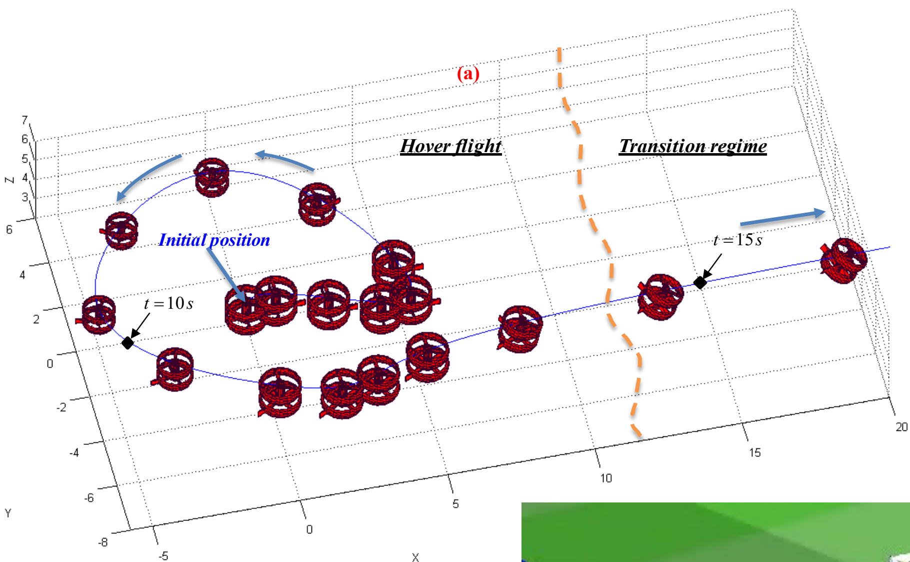


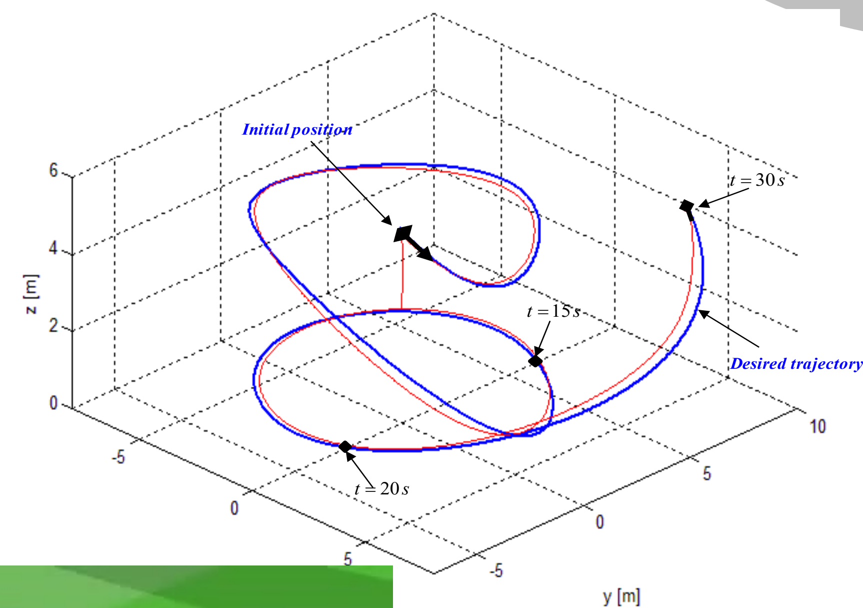
3D Visualization

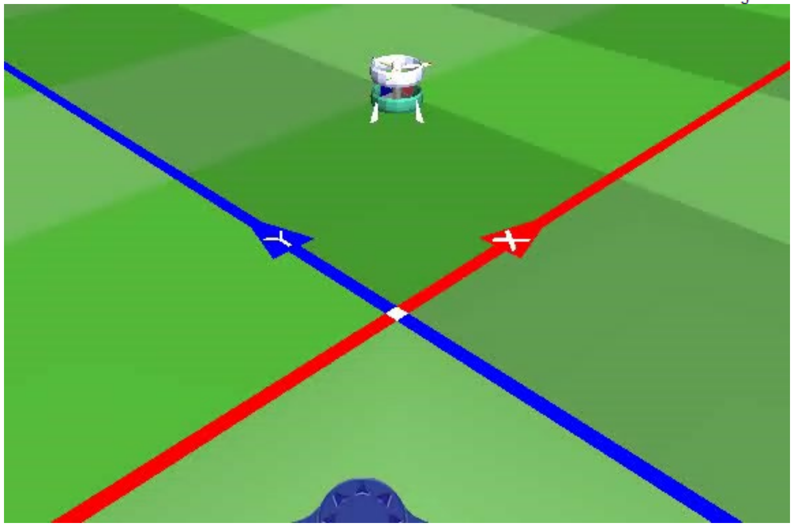
Simulation results



Simulation results



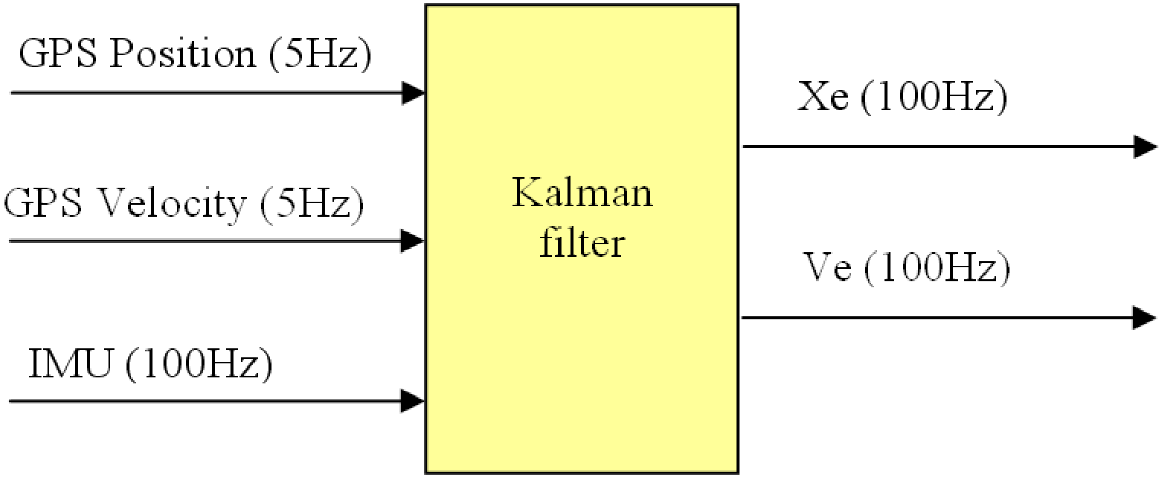




The X4-Flyer : LIST-CEA - GPS Localization & navigation

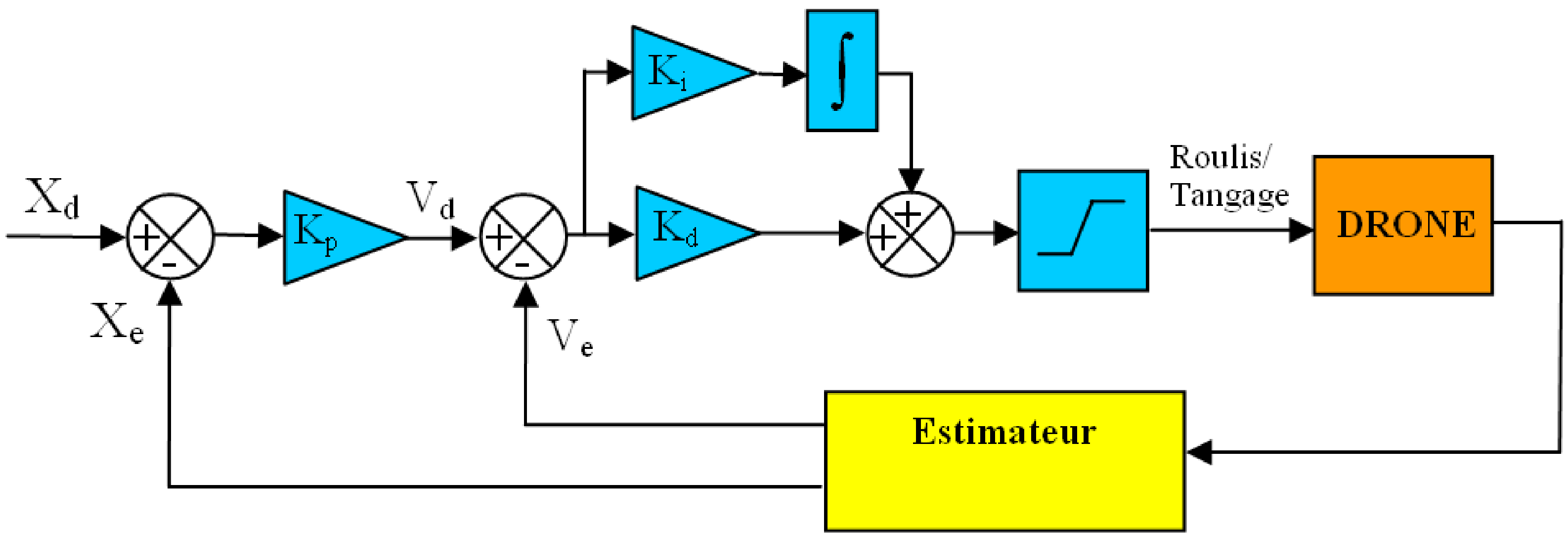
- Waypoints trajectory
- Constraints
  - 5Hz GPS data
  - Delayed GPS data (>1s)
  - Perturbations: wind

- ***Localisation***
  - ✓ ***Use of Kalman filter***

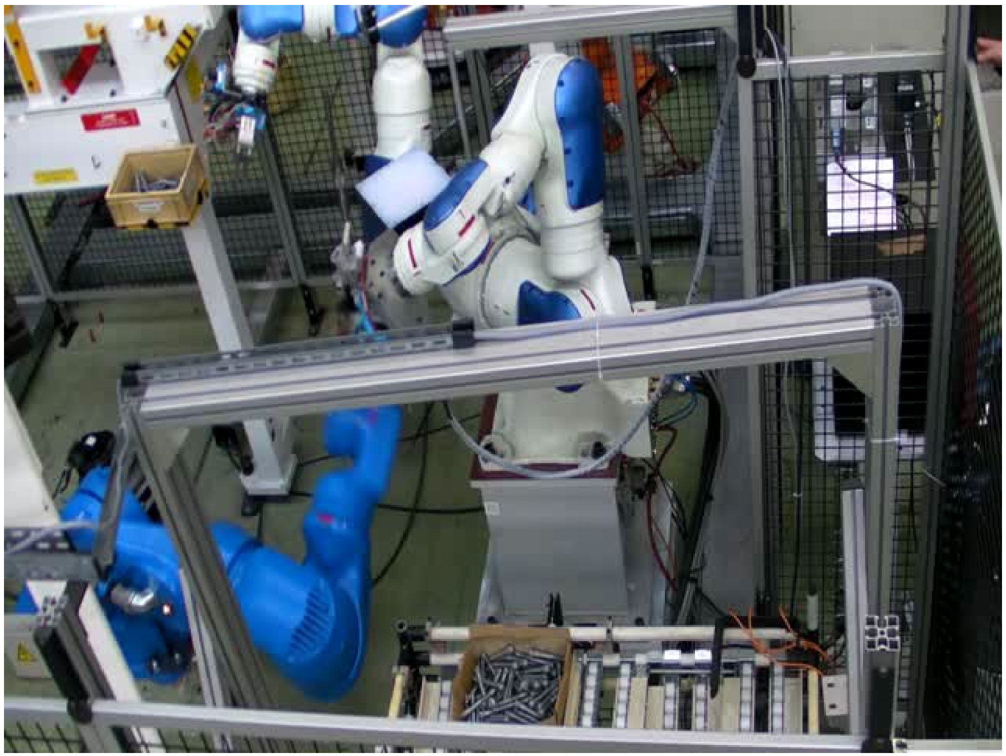


The X4-Flyer : LIST-CEA - GPS Localization & navigation

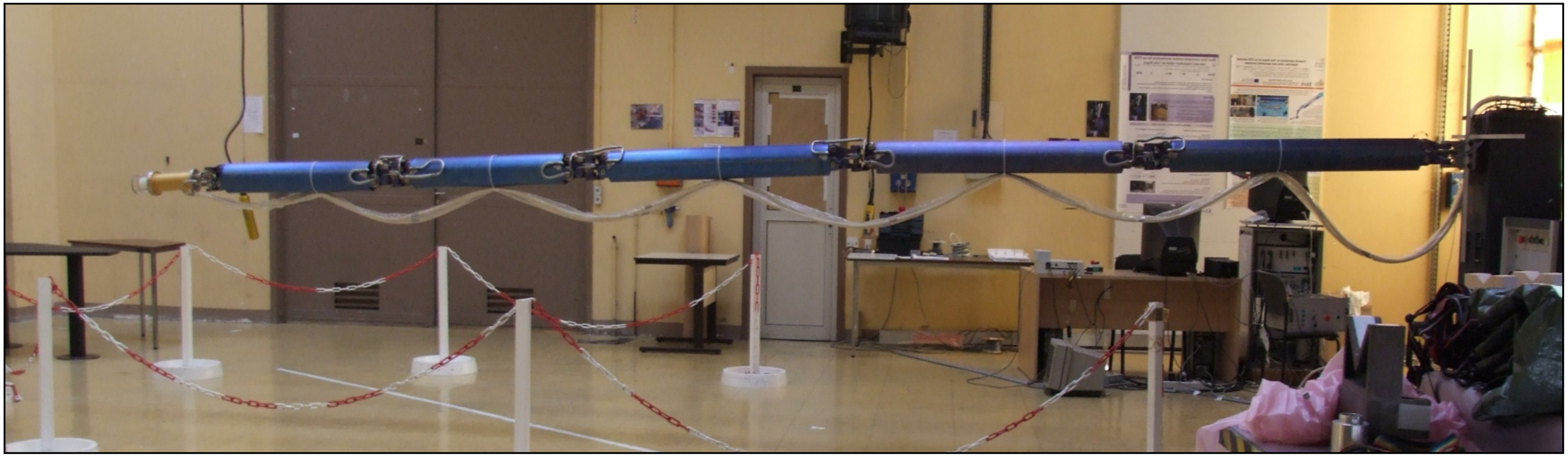
- Control loop with GPS waypoint trajectory generation
- Controller P(PI) : Speed / Position

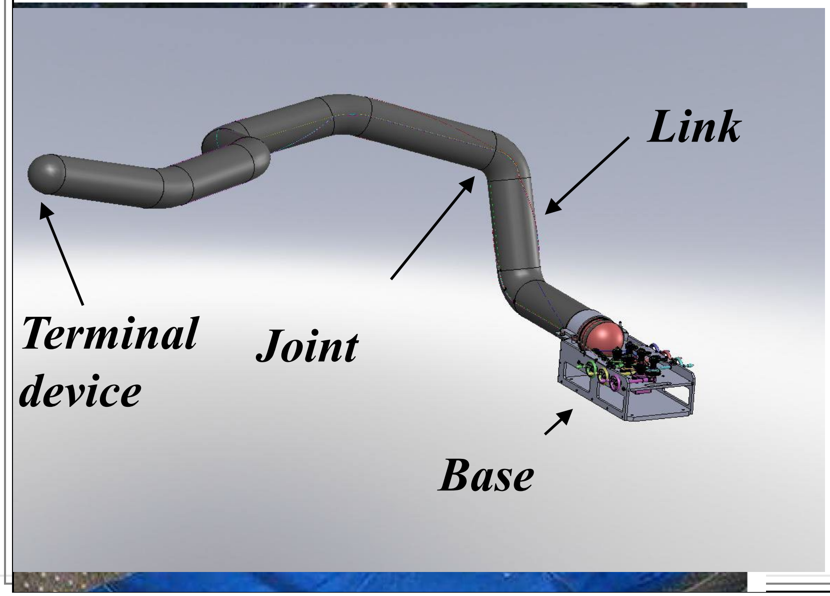


- **Robotic Screwing**
  - Random Bin Picking
  - *Multi-arm robots*



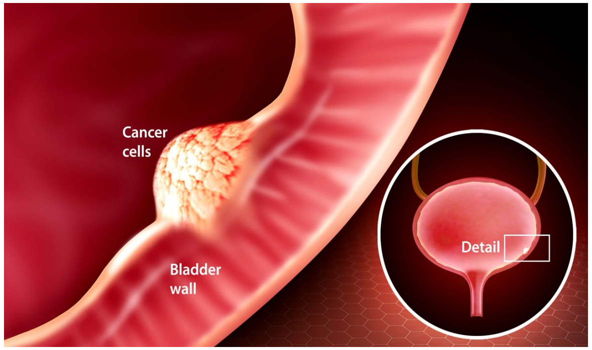
- Inspection in hostile environment (radioactivity, toxicity, …)

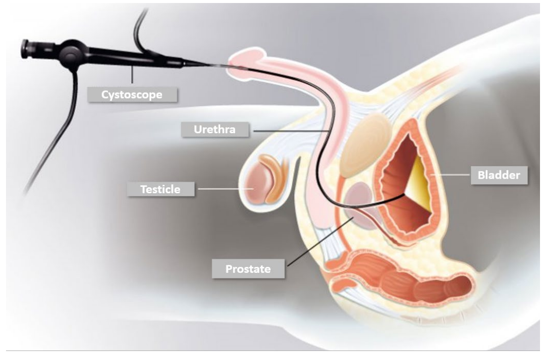


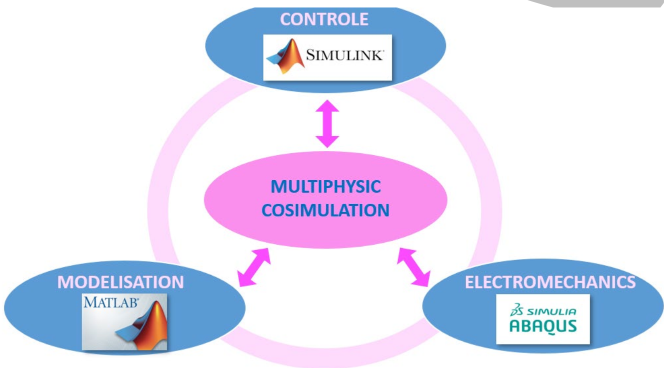


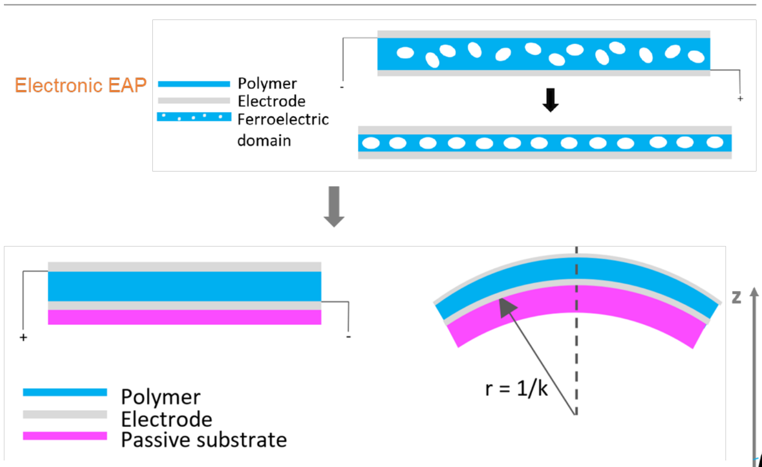
Mechbal N.

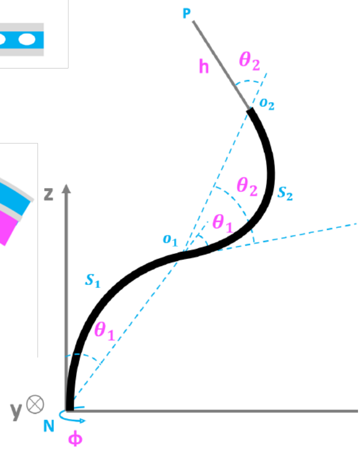
Electroactive Polymer Actuator For Endoscopic Robot

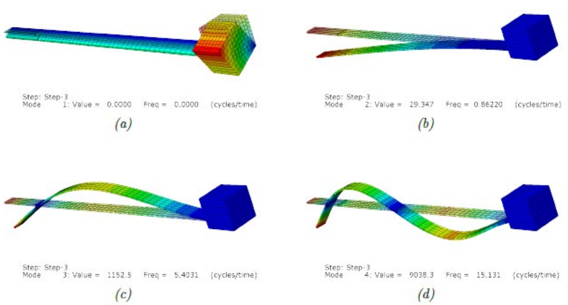


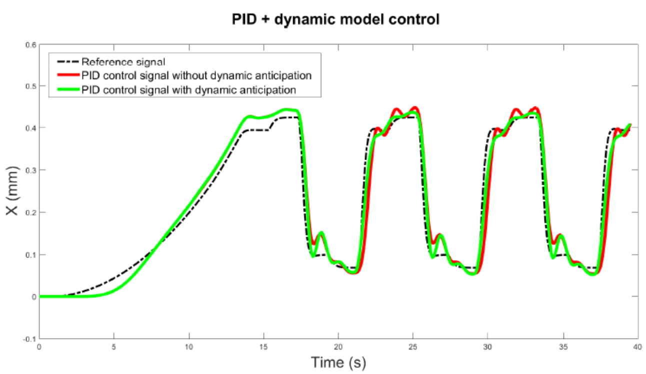


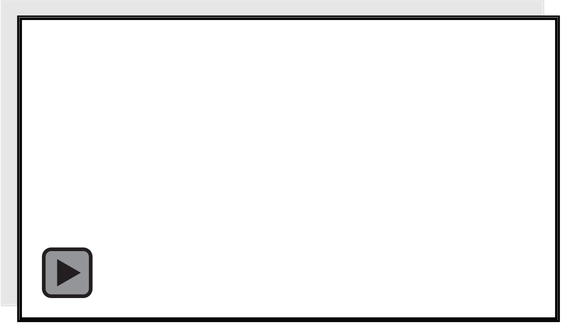






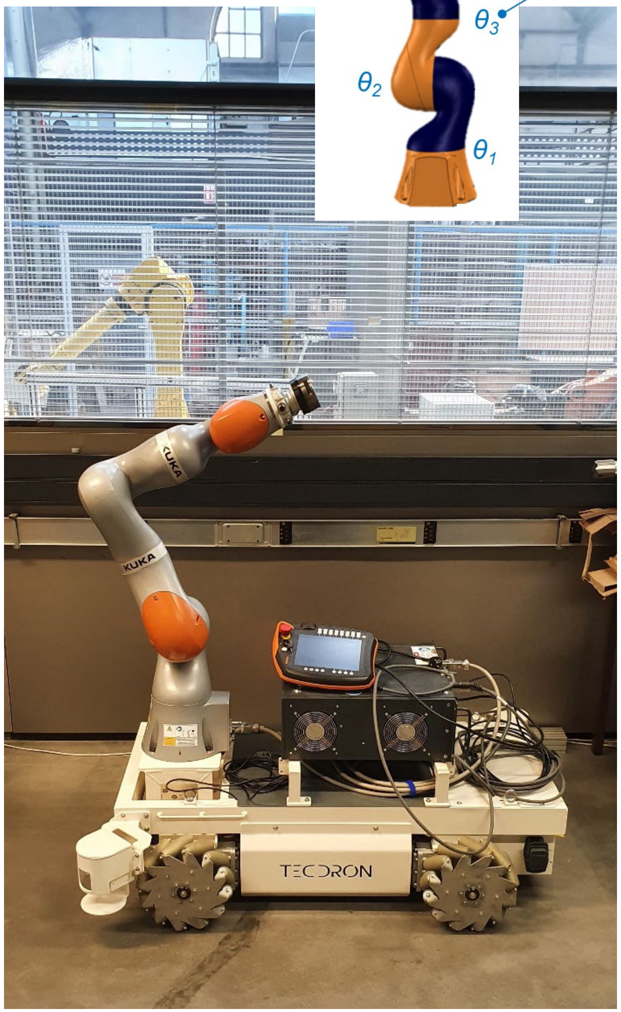


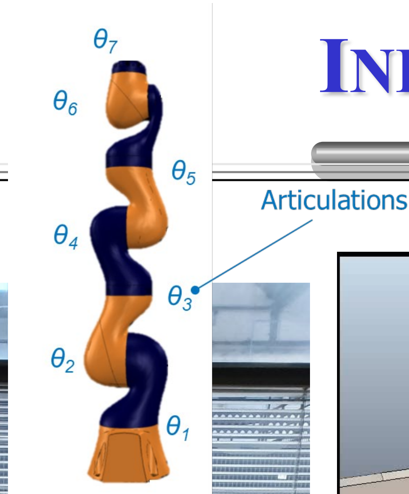


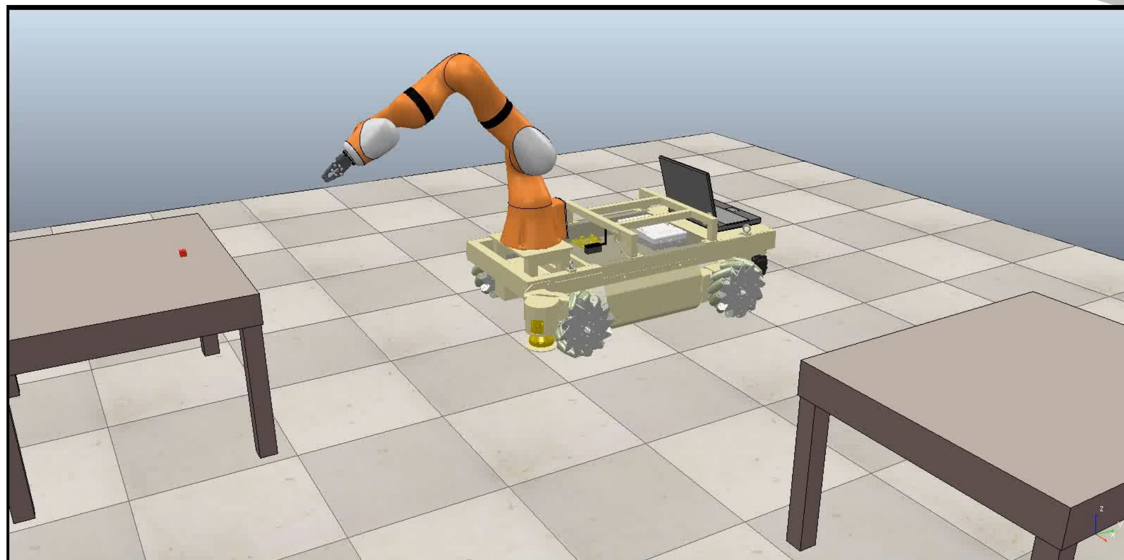


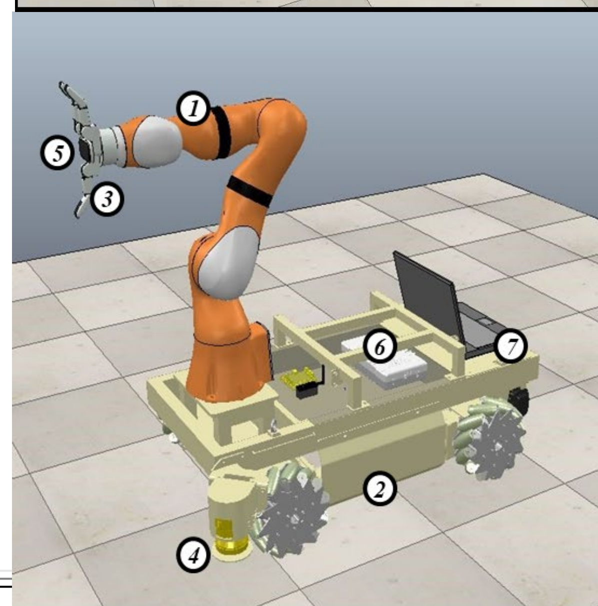
Mechbal N.

INDUSTRIAL MOBIL ROBOT ARM









1. *Bras manipulateur*
2. *Base mobile*
3. *Effecteur (pince)*
4. *Capteur (lidar)*
5. *Capteur (Camera)*
6. *Batterie et électronique de commande*
7. *Automate de commande*

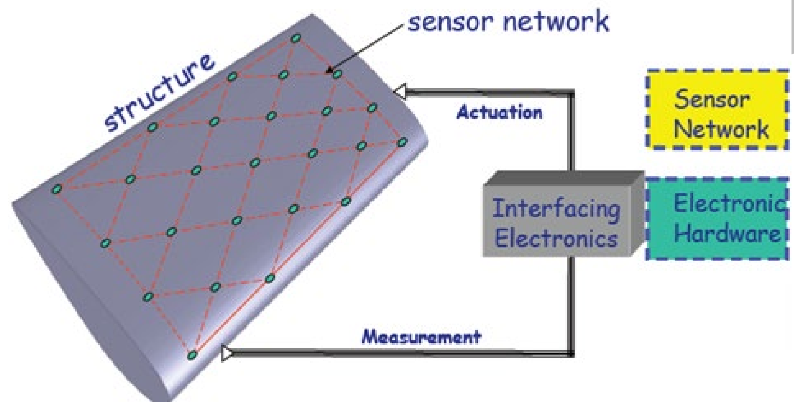
✷ Structural Health Monitoring

- SHM is a technology *to automate the inspection* process to assess and evaluate the health condition of structures in real-time or at specified time intervals.
- SHM is a system technology which integrates sensors/actuator networks with structures, software to interpret sensor signals, and hardware to process and manage the signals.
- *SHM is the key technology to enable the transition from traditional schedule-driven maintenance to Condition-Based Maintenance (*CBM).

SHM - SMART STRUCTURES

Smart structures

*A **smart structure** is an instrumented structure with embedded sensors (nerves) and actuators (muscles) with a central processor (brain) that try to mimic living beings systems in the presence of internal or external forces.*



Active elements / Smartness:

- Shape memory alloys, optical fiber (Fiber Bragg Grating, FBG), MEMS, …, and **Piezoelectric elements – PZT-** (actuation/sensing, on surface or embedded into).

Damage in composite structures:

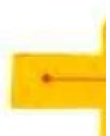
- Composite materials with fiber reinf
- There are three main ty

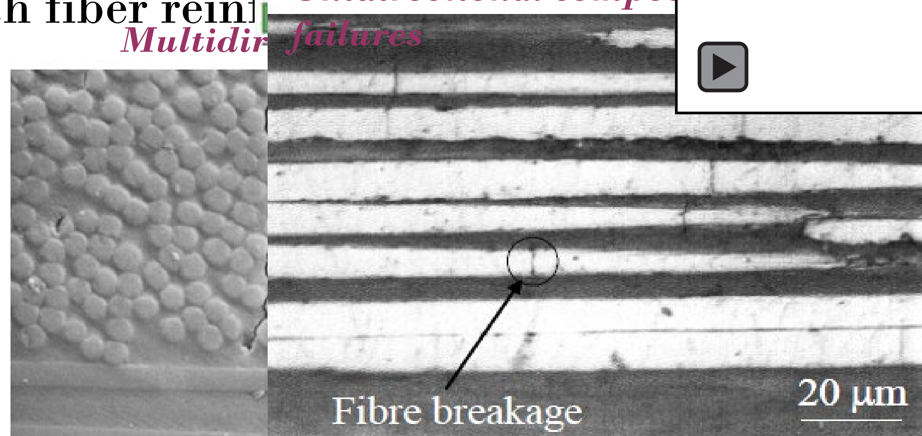
GFRP (glass)

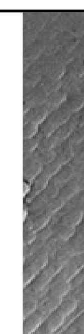
Unidirectional compos

Multidir

failures





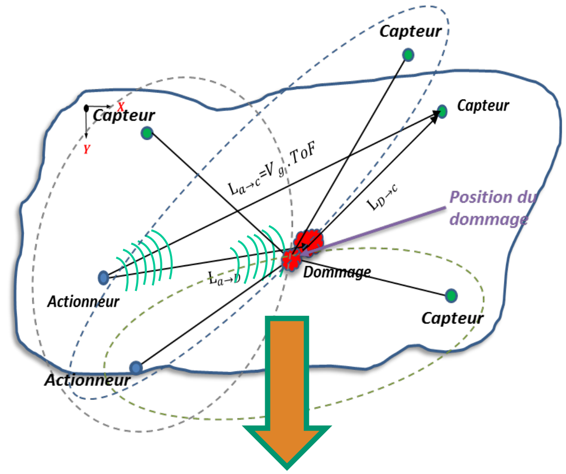


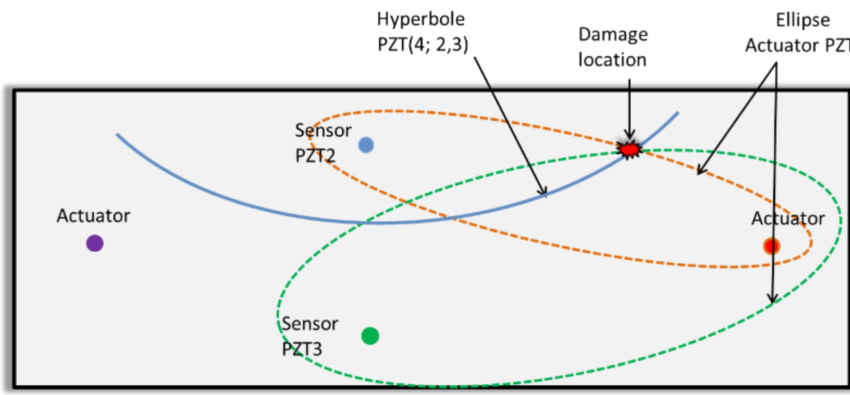
Focus: Bayesian framework for SHM (Fendzi, Phd)

Lamb waves-based damage localization :

- *Time of flight (Tof) based principle*:
  - Ellipse method: time of arrival (ToA) :
  - Hyperbola method: time difference of arrival (TDoAs)

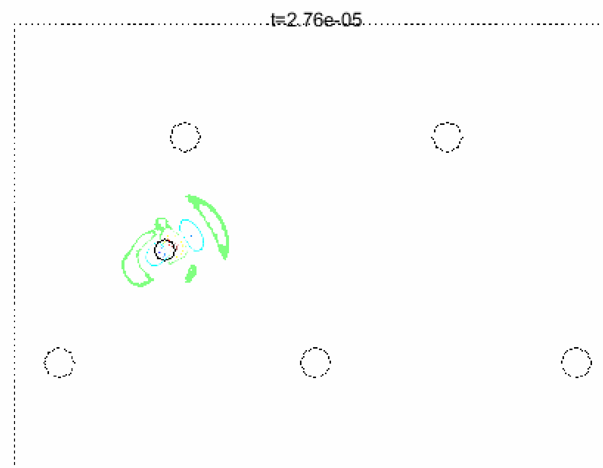
$$d_{a-D} + d_{D-C} = \boldsymbol{Vg}\, ToF\ (scatt.signal)$$





Problems:

- The size and extent of the damage are unknown → its ef **uncertain**
- The **sizes of the** actuator and sensor are usually not algorithms
- The **dispersive** nature of Lamb waves → the waveform distorted
- Composite structures → **angular** dependent group velo
- Robust decision: **Spatial probability of localization**



Mechbal N.

Focus: Bayesian framework for SHM

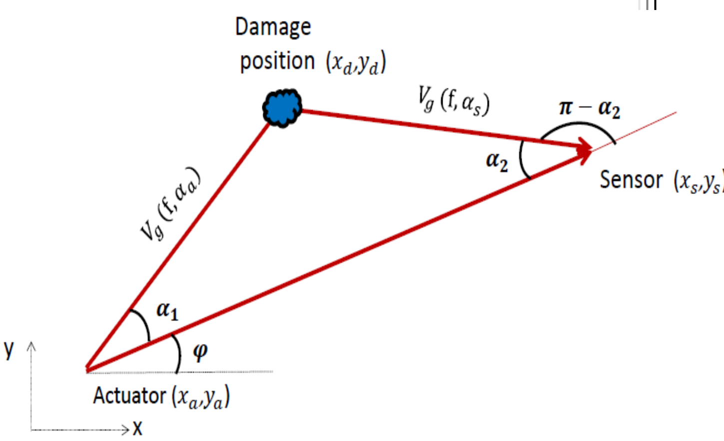
Approach: Bayesian estimation

- Bayesian formulation of the Tof:

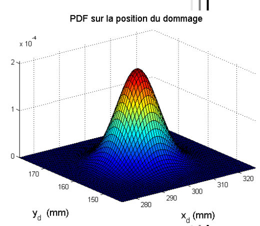
$$\boxed{Tof_m = Tof_c(\boldsymbol{\theta}) + \varepsilon}$$

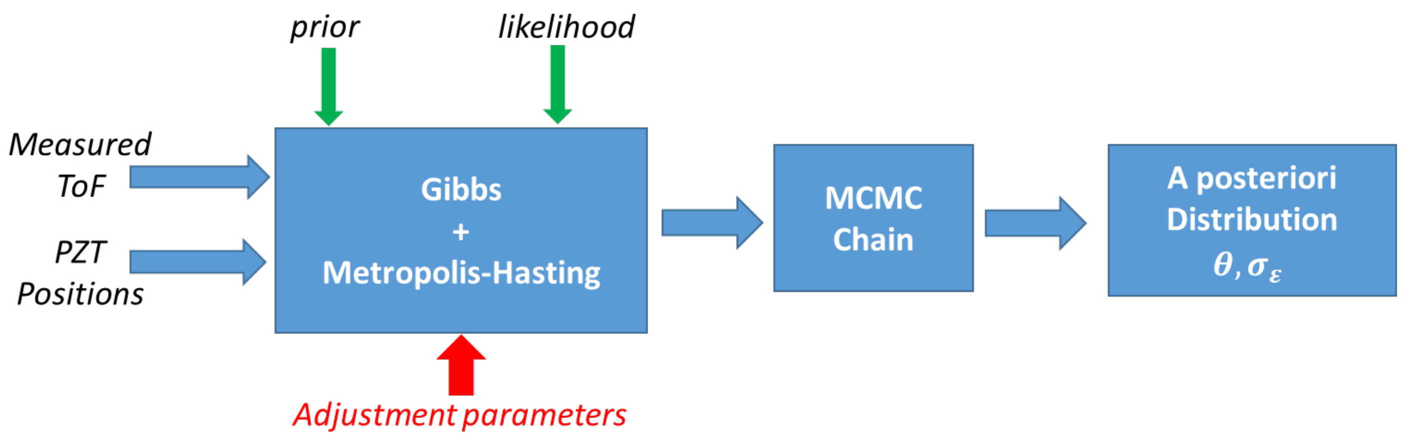
$$\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2) = \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left(\frac{(Tof_m - Tof_c(\boldsymbol{\theta}))^2}{2\sigma_\varepsilon^2} \right)$$

$$\boldsymbol{\theta} = [x_d, y_d, V_g(f, \alpha_a), V_g(f, \alpha_s)]$$



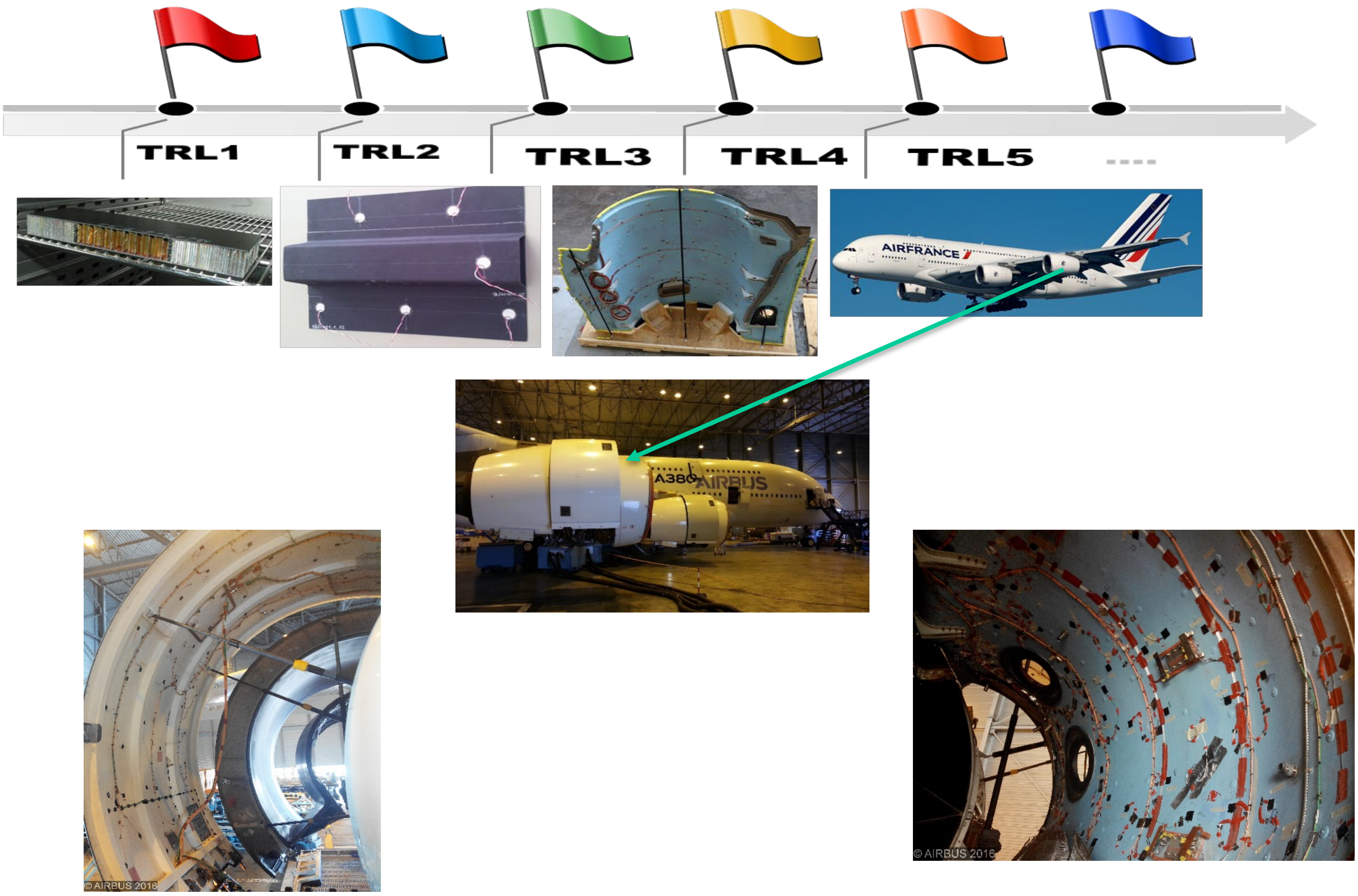
- Estimation: Sampling methods
  - Joint Gibbs sampling
  - Metropolis-Hastings algorithms





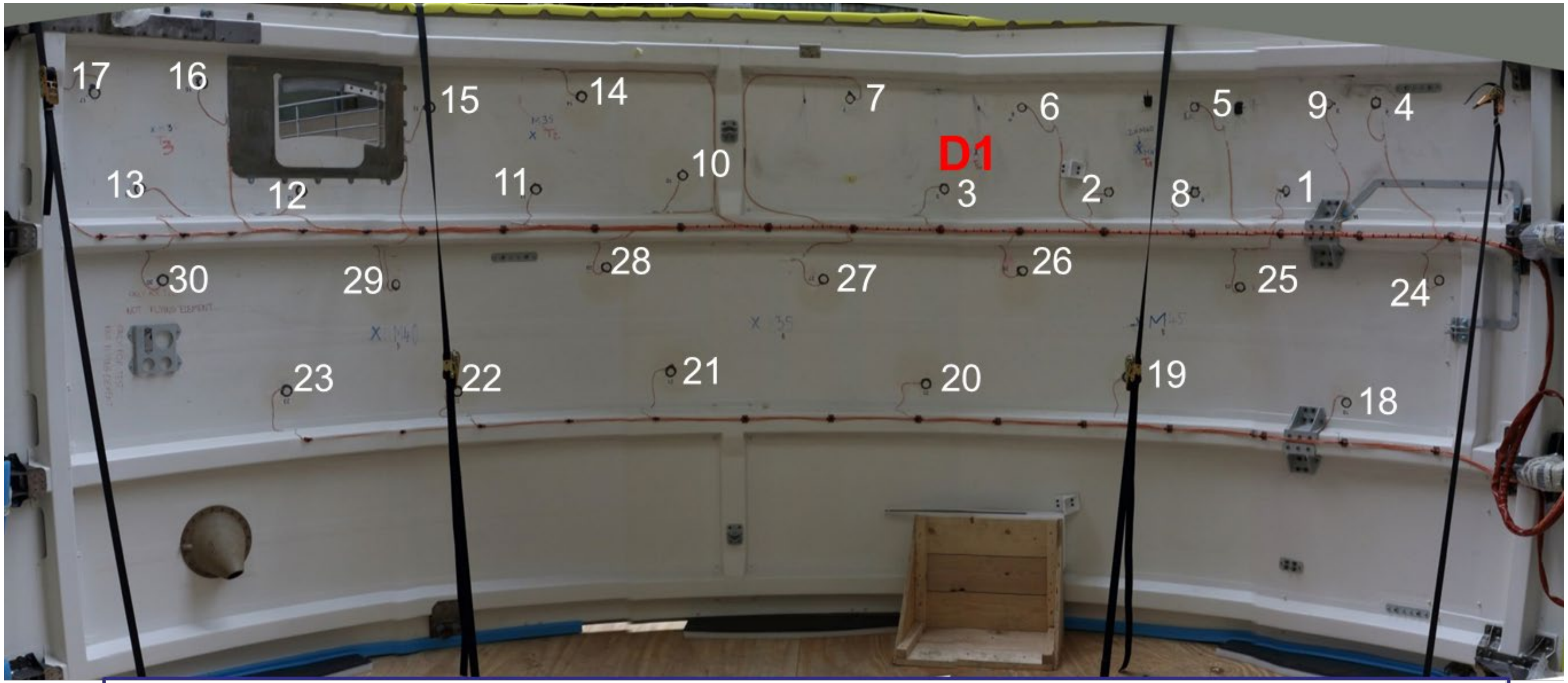
Focus: Bayesian framework for SHM

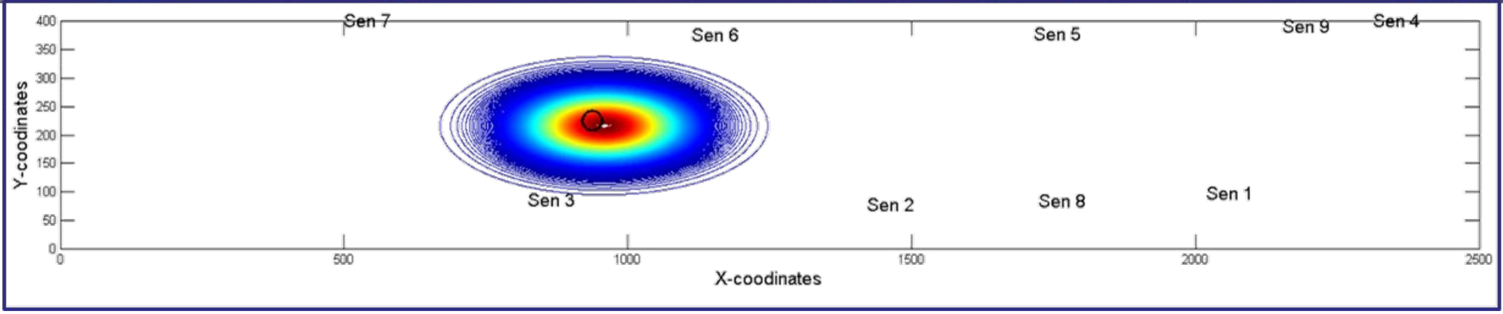
* *Application: SHM for Nacelles*



FOCUS: BAYESIAN FRAMEWORK FOR SHM

- Localization: real nacelle with one damage





COURSE OBJECTIVES

- **The purpose is to give the participant comprehension on some methods to perform advanced system (parameter and state) estimation and experimental identification:**
  - Bayesian estimators for linear and nonlinear system
  - Sequential Monte Carlo estimators for nonlinear, non-Gaussian system
  - Experimental System identification with application to Industrial Robot

- **The course reviews the basic of signal processing and vibration theory related to measured input-output data and explore the particular case of parameters identification of flexible robot manipulators in robotic machining.**

- **Students should be able to compute this techniques using dedicated software as MATLAB.**

PART I: OUTLINE TOPICS

- There are 3 Topics in this part:
  - Chapter 1 Preliminaries: Probability Theory
  - Chapter 2 Linear estimation theory
  - Chapter 3 Non Linear Estimation

Do not put your faith in what statistics say until you have carefully considered what they do not say.

PART I: SYLLABUS

- **Lecture 1:**
  - Review of Probability
- **Lecture 2:**
  - Estimation theory
  - Exercises
- **Lecture 3:**
  - Linear estimator: Kalman Filter / Observer
  - Exercise
- **Lecture 4:**
  - Linear estimator: Kalman Filter / Observer
  - Matlab class
- **Lecture 5:**
  - Nonlinear estimation: EKF and UKF
  - Exercises
  - Matlab class
- **Lecture 6:**
  - Nonlinear estimation: PF and RBPF
  - Exercises
  - Matlab Class

Section I: Review of probability theory

It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge.
... The most important questions of life are indeed, for the most part, really only problems of probability.
... The theory of probabilities is at bottom nothing, but common sense reduced to calculus.

Laplace, Theorie analytique des probabilites, 1820

I. ***Probability: intuitive approach***
II. ***Probability trough sets***
III. ***Random variable and vector***
IV. ***Random process***

PROBABILITY: INTUITIVE APPROACH

- We attribute the phenomena that seem to arrive and take over without any order, causes and hidden variables, whose action has been designated by the word chance, a word which is basically just an expression of our ignorance. *The probability is related in part to this ignorance, in part, to our knowledge.*
- Probability theory consists in reducing all the events that can occur in a given circumstance to *a number of cases equally possible*, that is to say as we are also undecided about their existence and to determine from among these case, the number of those who support the event which we seek probability.
- The ratio of the number to that of all possible cases, is the measure of this probability is therefore only a fraction whose numerator is the number of favorable cases and whose denominator is that all possible cases (Laplace Pierre-Simon, 1814).

Mechbal N.

- **Terminology :**
  - Experiment: Experience leading randomly to several possible outcomes
  - Outcomes: possible results of the experiment.
  - Sample space : set of all possible outcomes in a given experiment, noted Ω.
  - Event: subset of the sample space
- **Intuitive approach to probability:**
  - The most important property of random phenomena is statistical regularity. Indeed, consider an event A associated with a given random experiment. And suppose we repeat N times the experience. Noting n_A or $n(A)$ the number of times A is performed, the *frequency of realization* of A on the N shot is:

$$f_A = \frac{n_A}{N} = \frac{\text{number of favorable cases}}{\text{number of all possible cases}}$$

the ratio of the number of favorable cases to that of all cases possible

  - This frequency:

$$\begin{cases} f_A \in [0,1] \\ f_A \geq 0 \; ; \; f_\Omega = 1 \\ f_{A,\mathbf{B}} = f_A + f_B \; (A \text{ et } B \text{ exclusifs}) \end{cases}$$

Mechbal N.

Définition :

- The intuitive approach consists in associating with each event A, a number that will be called "probability of occurrence" of A. This number laying between 0 and 1 represents the chance that this event be realized. It is defined by

$$\boxed{P(A) = \lim_{N\to\infty} f_A}$$

- For example: in a game of dice, one can calculate the set:

$$\Omega = \{1,2,3,4,5,6\} = \{x_i\}$$

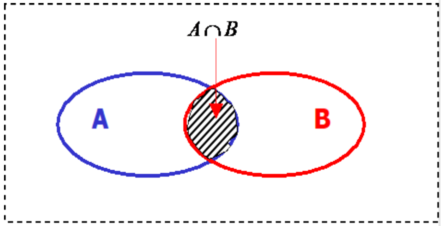
and the following likelihood of occurrence:

$$P(x_i) = \frac{1}{6} \qquad et \qquad P(even\ number) = \frac{3}{6} = \frac{1}{2}$$

Mechbal N.

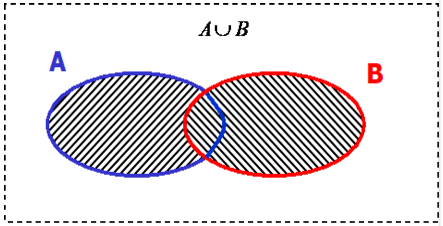
Set operations:

- The Intersection of two sets A and B :



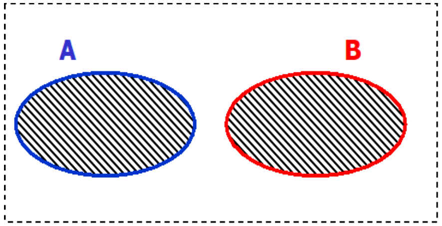
$$A \cap B = A \text{ and } B$$

- The Union of two sets A and B



$$A \cup B = A \text{ or } B$$

- Two sets A an B are mutually exclusive if



$$A \cap B = \varnothing \Rightarrow A \text{ et } B \text{ sont exclusifs}$$

- the complement of a set A :

$$C(A) = A^C = \text{not in } A$$

Definition :

Consider the set Ω of N possible events. To each event $A_i \in \Omega \; i = \ldots N$, we assign a nonnegative number called *probability* this number is noted $P(A_i)$ for "probability of event A" and has to satisfy the following three axioms:

$$(i) \qquad P(\Omega) = 1$$

$$(ii) \qquad P(A_i^C) = 1 - P(A_i)$$

$$(iii) \qquad P(A_1 \; ou \; A_2) = P(A_1 \cup A_2) = P(A_1) + P(A_2) \qquad \left(A_1 \cap A_2 = \varnothing\right)$$

(i) recognizes that the sample space itself is an event "the certain event".
(ii) gives the complement.
(iii) states that the probability of the event equal to the union of any number of mutualy exclusive events.

Properties:

$$P(\varnothing) = O$$

$$P(A_2 \cap A_1^C) = P(A_1) - P(A_1 \cap A_2)$$

if A_1 is include in A_2 then we have: $A_1 \subset A_2 \Rightarrow P(A_1) \le P(A_2)$

Mechbal N.

Joint and conditional probability

- In some events are not mutually exclusive because of common elements. These elements correspond to the *simultaneous or joint* occurrence.

- **Joint probability**: noted $P(A_1 \, and \, A_2)$ or $P(A_1, A_2)$ and given by:

$$P(A_1 \, and \, A_2) = P(A_1) + P(A_2) - P(A_1 \, or \, A_2) \leq P(A_1) + P(A_2)$$

$$or$$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2 \cap A_1^C) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

- **Conditional probability**: noted $P(A_1 \, if \, A_2)$ or $P(A_1/A_2)$ it is the probability of $P(A_1)$ knowing that the event A_2 has occurred, *i.e.*,

$$P(A_1 \, if \, A_2) = \frac{P(A_1 \, and \, A_2)}{P(A_2)} \qquad or \qquad P(A_1 / A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}$$

Statistically independence:

- This important concept of *conditional probability*, introduces another key concept, that is statistical independence.
- Thus if events are independent of one another, the probability of their combined existence is the product of their respective probabilities.

$$P(A_1 \ and \ A_2) = P(A_1 \cap A_2) = P(A_1)P(A_2)$$

- So using the conditional probability, this means that the probability of occurrence of one event is not affected by the occurrence of the other event:

$$\Rightarrow \quad \boxed{P(A_1 / A_2) = P(A_1)}$$

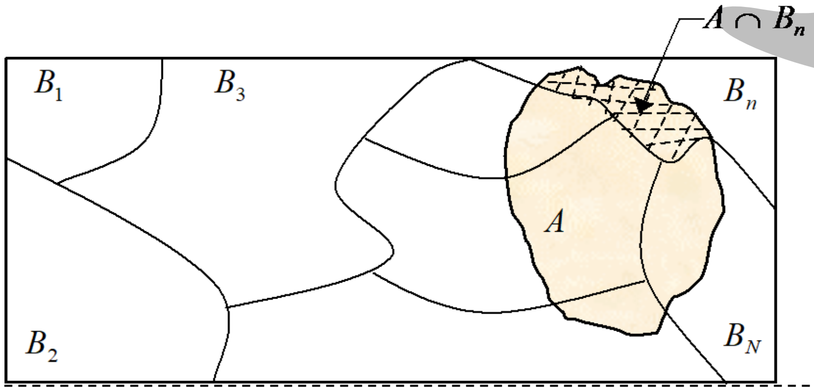
- Independence is fundamental to our later work because it a hypothesis that greatly simplified probability problems.

- **Remark**: two mutually exclusive events cannot be statistically independent:

$$P(A_1 \cap A_2) = 0 \neq P(A_1)P(A_2)$$

Mechbal N.

Total probability:



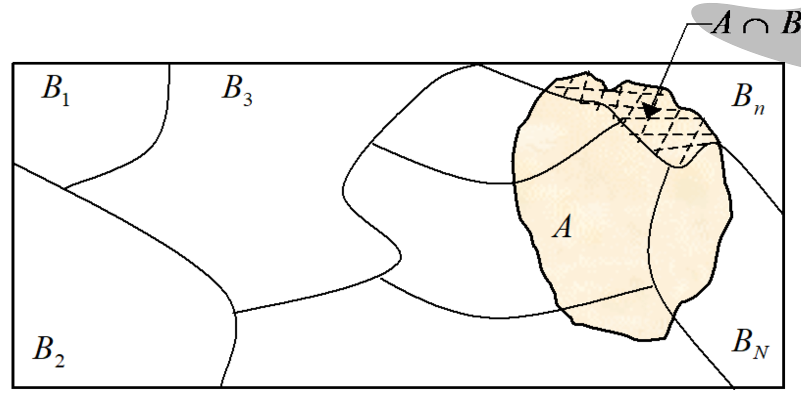
- ***Total probability***: The probability $P(A)$ of any A defined on a sample space S set can be expressed in terms of conditional probabilities. Suppose we are given N mutually exclusive events B_n, $n = 1, 2, \ldots, N$, whose union equals S. These events satisfy.

  - $B_m \cap B_n = \varnothing \quad pour\ m \neq n$ et $\bigcup_{n=1}^{N} B_n = S$

- The total probability of event A is defined by (car $A \cap S = A$) :

$$P(A) = \sum_{n=1}^{N} P(A/B_n)P(B_n)$$

Bayes' Theorem



- ***Bayes' Theorems***: let B_n ***and*** A two events defined above. From the definition of the conditional probability we can write:

$$P(B_n / A) = \frac{P(B_n \cap A)}{P(A)} \qquad if \quad P(A) \neq 0$$

The Bayes rule is then obtained by the following expression:

$$P(B_n / A) = \frac{P(A/B_n)P(B_n)}{P(A)}$$

or using the total probability:

$$P(B_n / A) = \frac{P(A/B_n)P(B_n)}{P(A/B_1)P(B_1) + \ldots + P(A/B_N)P(B_N)}$$

Summary

$$P(A,B) = P(B,A) = P(A\,|\,B)P(B) = P(B\,|\,A)P(A)$$

$$P(A\,|\,B) = \frac{P(B\,|\,A)P(A)}{P(B)} = \frac{P(B\,|\,A)P(A)}{P(B,A) + P(B,\bar{A})}$$

$$P(A\,|\,B,C) = \frac{P(B,C\,|\,A)P(A)}{P(B,C)} = \frac{P(B\,|\,C,A)P(C\,|\,A)P(A)}{P(B\,|\,C)P(C)} = \frac{P(B\,|\,A,C)P(A\,|\,C)}{P(B\,|\,C)}$$

$$P(A\,|\,B) = \frac{P(A,B)}{P(B)} = \frac{\sum_{n=1}^{N} P(A,B,C_n)}{P(B)}$$

$$= \sum_{n=1}^{N} \frac{P(A,C_n,B)P(C_n,B)}{P(C_n,B)P(B)} = \sum_{n=1}^{N} P(A\,|\,C_n,B)P(C_n\,|\,B)$$

Random Variable:

- The goal is to associate a quantifiable value to the various possible outcomes of a random experiment. A random variable is a quantity that depends on the outcome of the experiment. It is a real function of the element of a sample space.
- We represent a random variable, X, by a capital letter and its value, x, by lowercase letter.

Distribution Function:

- Noted $F(x)$ and it is a number that depends on x. We call this function the cumulative probability distribution function of X and it is defined by:

$$F(x) = \Pr ob(X \leq x) = \lim_{N \to \infty} \frac{n(x)}{N}$$

It is the probability of the event $X \leq x$. *It has some specific proprieties:*

(i) $F(+\infty) = 1$

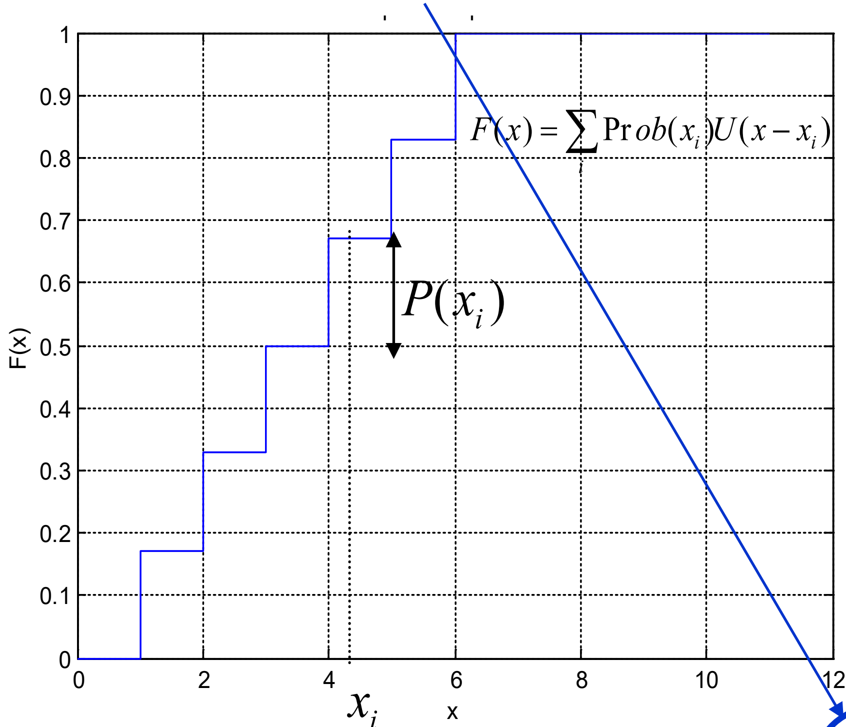
(ii) $F(-\infty) = 0$

(iii) $F(x)$ *is a nondecreasing function of* x.

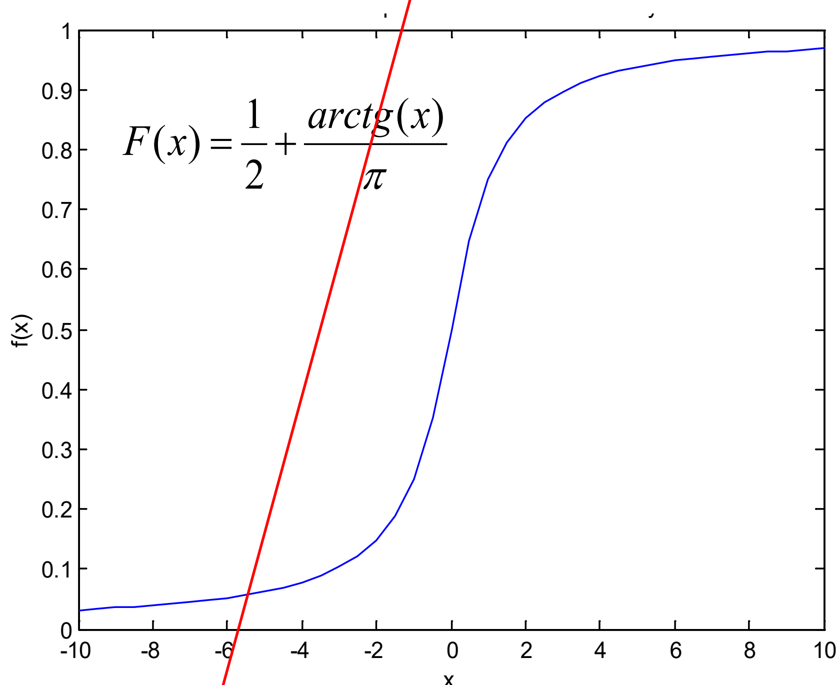
Distribution Function

- Depending on the nature of the random variable, the distribution can be either continuous or discrete

***Discrete* : craps game**



***Continuous:* Cauchy function**



- General case

$$F(x) = a_1 F_1(x) + a_2 F_2(x)$$

Mechbal N.

Density function

- Noted $f(x)$, it is the derivative of the distribution function

$$\boxed{f(x) = \frac{dF(x)}{dx}} \Rightarrow \boxed{F(x) = \int_{-\infty}^{x} f(u)du}$$

- It is a probability density in the sense that the probability that X must belong to the interval [a, b] is given by:

$$\boxed{\Pr ob\big(X \in (a,b)\big) = \int_{a}^{b} f(x)dx = F(b) - F(a)} \Leftrightarrow \boxed{\int_{-\infty}^{\infty} f(x)dx = 1 \quad area = 1}$$

- It is a non-negative function and not necessarily less than 1.

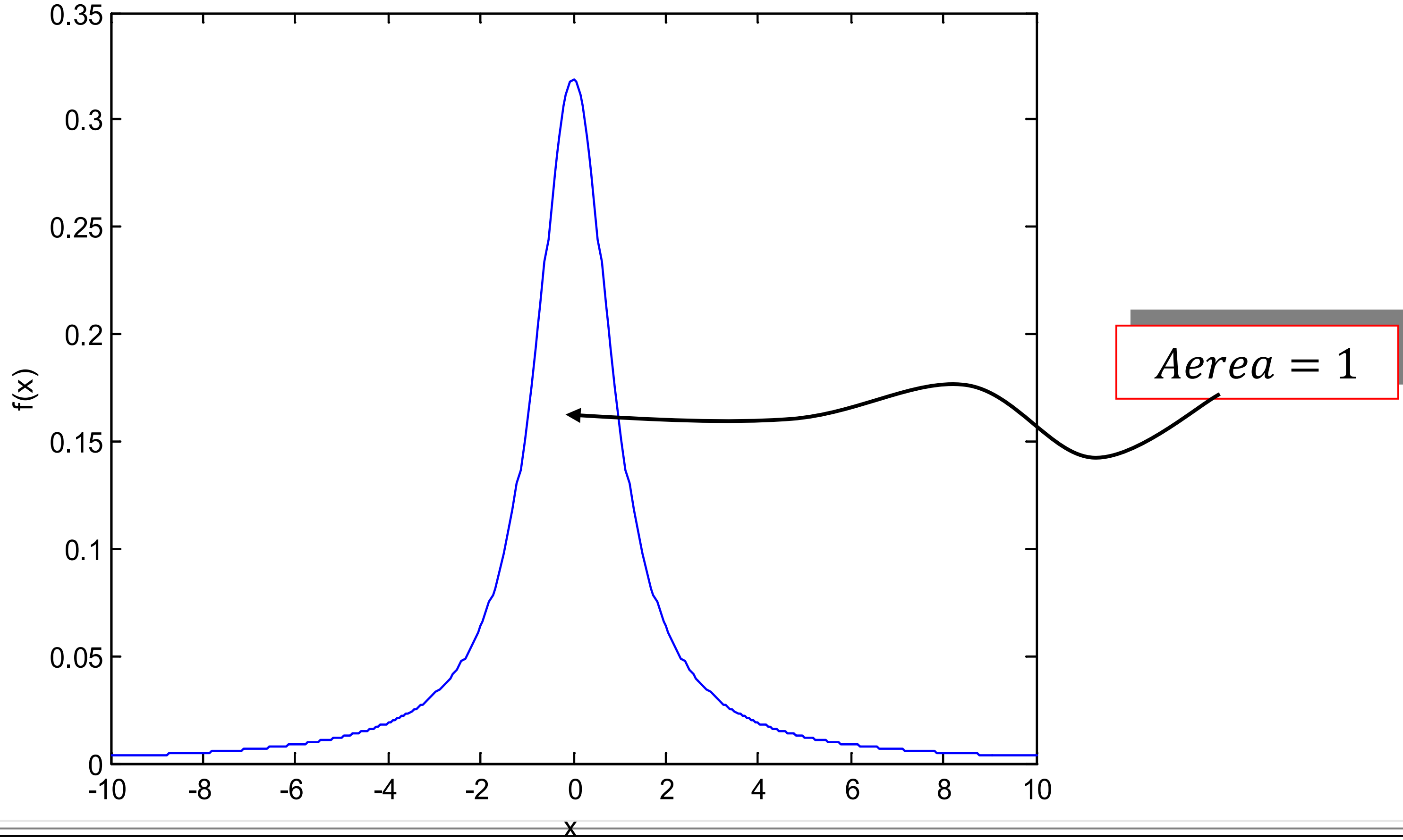
$$0 \leq f(x) \quad for \;\; all \;\; x$$

- It can be either continuous or discrete

Density function

- Continuous example: Cauchy function

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$$

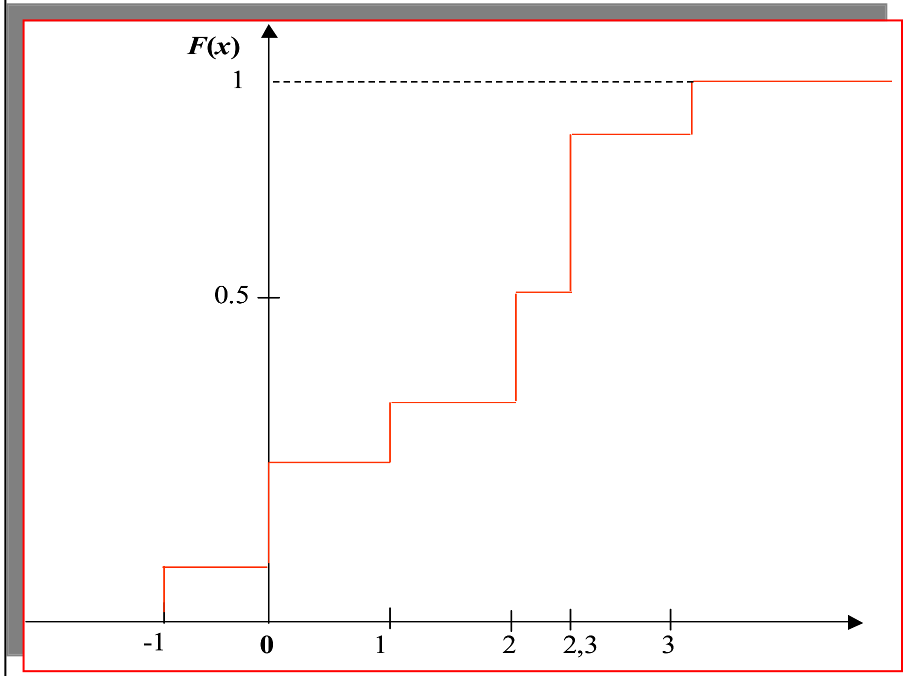


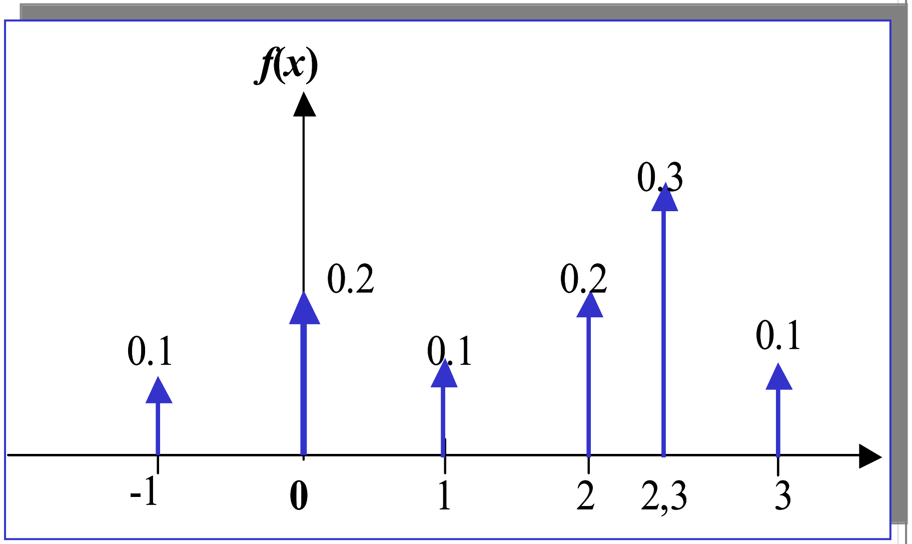
Example

- <u>Discrete random variable</u> : *consider the random variable X having its values in the set* $\{-1,0,1,2,2.3,3\}$ and the corresponding probabilities are:

$$\{0.1,0.2,0.1,0.2,0.3,0.1\}$$

- The distribution and density functions are given by :





Step $F(x) = \sum_i \boldsymbol{Prob}(x_i)U(x - x_i)$

Dirac $f(x) = \sum_i \boldsymbol{Prob}(x_i)\,\delta(x - x_i)$

Expectation of a variable (or Espérance mathématique)

- It is the process of averaging when a random variable is involved. The expected value of X is noted $E(X)$.

N measurements are performed, and we have n_i times the value x_i, we calculate then the mean by:

$$\overline{X} = \frac{1}{N} \sum_i n_i x_i$$

***frequency of realization** = Probability*

- We can show that the calculus of the mean of a random variable can be generalized to incorporate *probability information* through the use of its *density function*. The expected value is then given by:

$$E(X) = \overline{X} = \int_\Omega X dF(x) = \int_{-\infty}^{+\infty} x f(x) dx \qquad \textit{continuous}$$

$$E(X) = \sum_i x_i f(x_i) \qquad \textit{discrete}$$

Expectation of a function of variable

- If we have now a function of a random variable the expected value is given by:

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) f(x) dx \qquad continuous$$

$$E(X) = \sum_i g(x_i) f(x_i) \qquad discrete$$

Conditional event and expectation

- To define a conditional event A in term of random variable X let:

$$A = \{X \le a\} \qquad -\infty < a < +\infty$$

- Then using conditional law, conditional distribution is given by

$$F(x / X \le a) = \Pr ob(X \le x / X \le a) = \frac{\Pr ob(X \le x \cap X \le a)}{\Pr ob(X \le a)}$$

- Two cases must be considered:

$$\begin{cases} x \ge a \;\Rightarrow\; \{X \le x \cap X \le a\} = \{X \le a\} \;\Rightarrow\; F(x / X \le a) = \dfrac{\Pr ob(X \le a)}{\Pr ob(X \le a)} = 1 \\ x < a \;\Rightarrow\; \{X \le x \cap X \le a\} = \{X \le x\} \;\Rightarrow\; F(x / X \le a) = \dfrac{\Pr ob(X \le x)}{\Pr ob(X \le a)} = \dfrac{F(x)}{F(a)} \end{cases}$$

Conditional event and expectation

- Using a similar statement and derivative we have the conditional density, *i.e.*,

$$f(x/X \le a) = \begin{cases} \dfrac{f(x)}{F(b)} = \dfrac{f(x)}{\int_{-\infty}^{a} f(x)dx} & x < a \\ 0 & x \ge a \end{cases}$$

- From these results, the conditional expectation is defined by:

$$E(X/A) = \int_{-\infty}^{+\infty} xf(x/A)dx \iff E(X/X \le a) = \frac{\int_{-\infty}^{b} xf(x)dx}{\int_{-\infty}^{a} f(x)dx}$$

which is the mean value of X when *it* is constrained to the set $\{X \le a\}$

Moments

- The nth moment:

$$m_n = E\{X^n\} = \int_{-\infty}^{+\infty} x^n f(x)dx$$

- If $n = 1$, we have the mean :

$$m_1 = E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

> ➢ *The average is a measure of the dynamics of a signal*
> ➢ *The variance characterizes the dispersion around the mean*
> ➢ *The third order moment characterizes the symmetry of the pdf*
> ➢ *The moment of order four measures the Gaussianity*

- Variance : It characterizes the ***dispersion*** (the range of variation) around the mean. It is noted $var(X)$ or σ^2 and defined by:

$$\boldsymbol{var}(\boldsymbol{X}) = E\{(X - E(X))^2\} = \int_{-\infty}^{+\infty} (X - E(X))^2 f(x)dx = m_2 - m_1^2$$

- The square roots of variance is called standard deviation of X:

$$\boldsymbol{\sigma}(\boldsymbol{X}) = \sqrt{Var(X)}$$

- Skewness: is the ***standardized centered third moment***. It is a measure of the ***asymmetry*** of de density function about its mean

$$S = \boldsymbol{m}_3^c/\sigma^3 \quad with \quad \boldsymbol{m}_3^c = E\{(X - E(X))^3\}$$

if a pdf is symmetric the skewness is null.

Mechbal N.

Moments

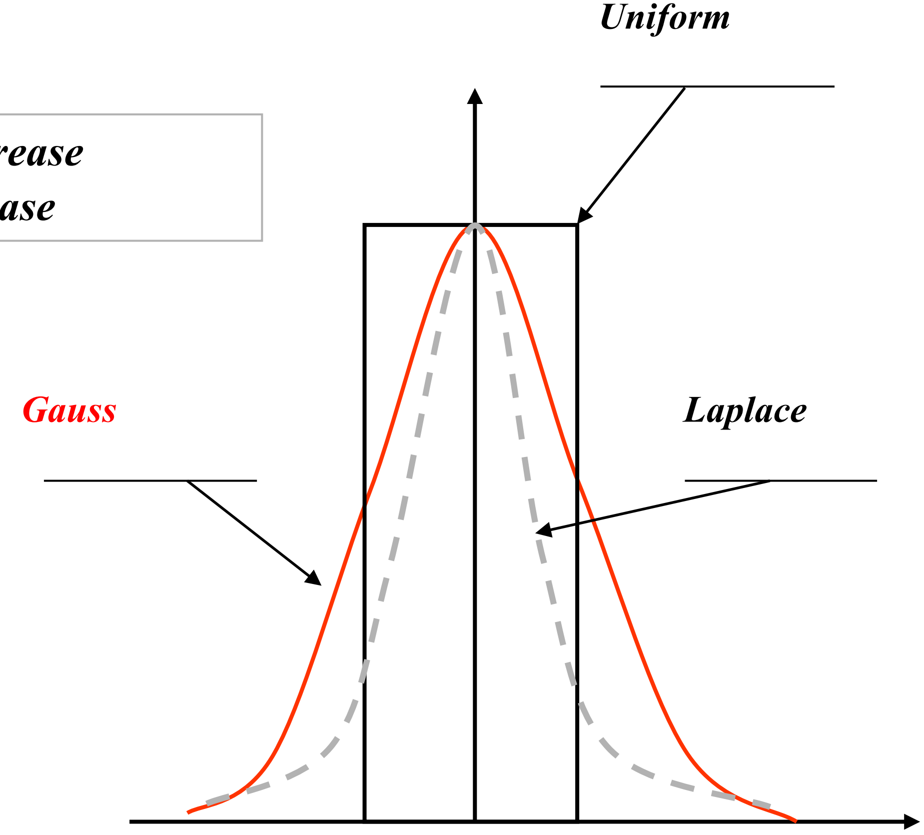
KURTOSIS: standardized fourth moment

- *is a measure of the "peakedness" (aplatissement) of the probability distribution. It is a descriptor of the shape of a probability distribution,*

$$K_x = \frac{E(X^4)}{\sigma^4}$$

$K_x > 3$ ***pdf quick decrease***
$K_x < 3$ ***pdf slow decrease***

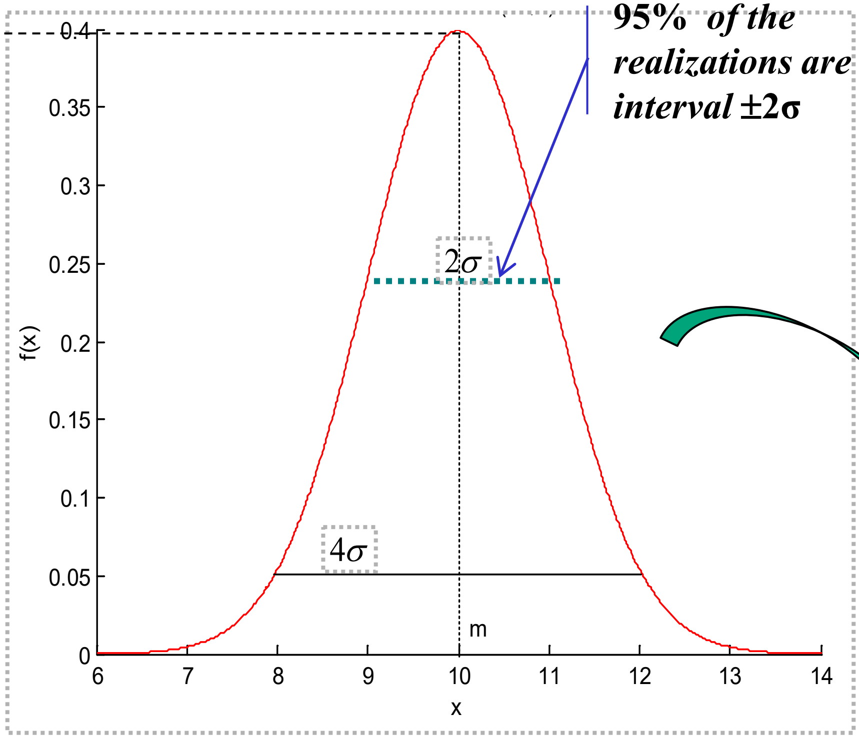
Loi	Kurtosis
Gauss	3
Uniform	1.8
Laplace	6



Gaussian random variable

- Gaussian density function: $$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$
- In this statistical distribution, most of the achievements are spread about the point $x=m$ (mean) and are related to σ (variance)
- Most important density encountered in nearly all engineering areas

$\frac{1}{\sqrt{2\pi\sigma^2}}$



95% of the realizations are in the interval ±2σ

$$\begin{cases} Pob(|X-m|) > \sigma = 32\% \\ Pob(|X-m|) > 2\sigma = 5\% \\ Pob(|X-m|) > 3\sigma = 0.3\% \\ Pob(|X-m|) > 4\sigma = 0.0063\% \end{cases}$$

Gaussian distribution function

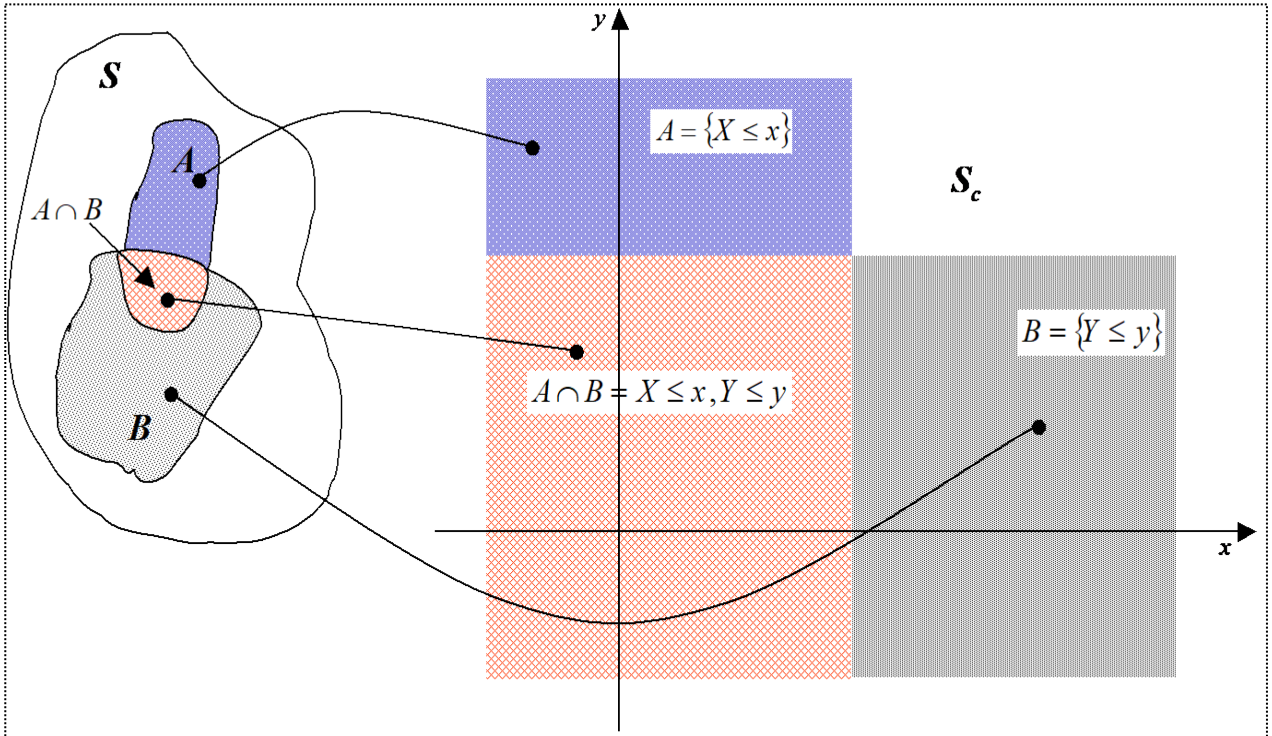
$$\Pr o\, b(-k\sigma < x < k\sigma)$$
$$= F(k\sigma) - F(-k\sigma) = \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{k} e^{-\frac{x^2}{2}} dx$$

Mechbal N.

- **Random Vector:** for simplicity and notation, we restrict ourselves to the case of two dimensions, the formulas can be easily generalized.
  - *Consider two random variables defined on a sample space S: :*

$$A = \{X \le x\} \quad et \quad B = \{Y \le y\}$$

  - The plane of all random point (x, y) is defined as the Joint sample space S_c (crosshatched)



Mechbal N.

Joint distribution and density:

Joint distribution:

$$\boxed{\begin{aligned} F(x,y) &= \text{Prob}(X \le x, Y \le y) \\ &= \text{Prob}(X \le x_1 \cap Y \le y) \end{aligned}}$$

$$\begin{cases} F(x,y) \ge 0 & F(+\infty,+\infty) = 1 \\ F(x,-\infty) = 0 & F(-\infty,y) = 0 \\ 0 \le F(x,y) \le 1 & \end{cases}$$

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(\alpha_1,\alpha_2)\,d\alpha_1 d\alpha_2 \qquad \textit{Continuous}$$

$$F(x,y) = \sum_{ij} p_{ij} U(x-x_i)U(y-y_j) \quad \forall\, i,j \ \ \textit{tels que} \begin{cases} x_i \le x \\ y_j \le y \end{cases} \qquad \textit{Discrete}$$

Joint density: it is the second derivative of the joint distribution

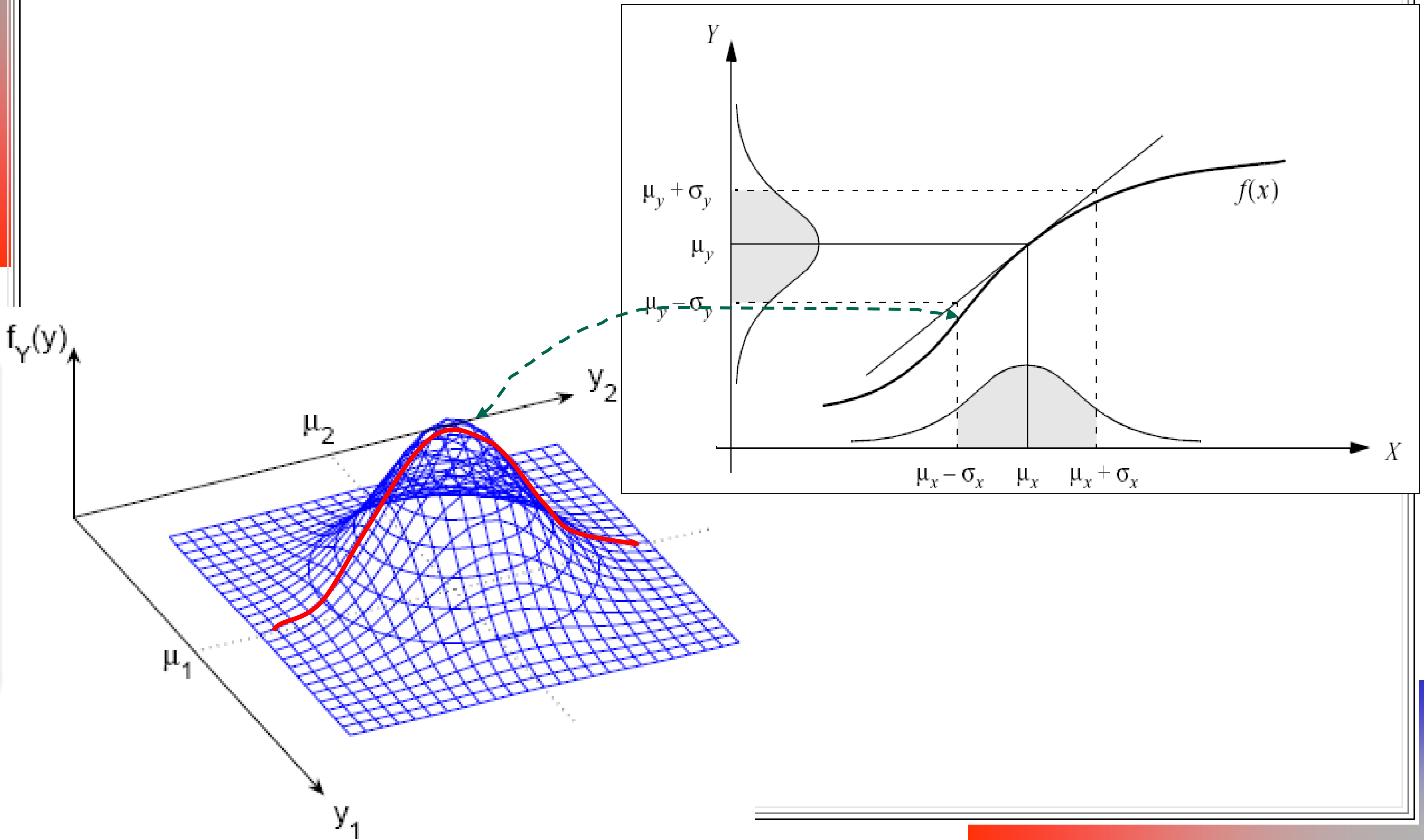
$$\boxed{f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}} \qquad f(x,y) = \sum_{ij} p_{ij}\delta(x-x_i, y-y_j)$$

$$\begin{cases} f(x,y) \ge 0 & \displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dxdy = 1 \\ \text{Prob}\{x_1 < X \le x_2, y_1 < Y \le y_2\} = \displaystyle\int_{x_1}^{x_2}\int_{y_1}^{y_2} f(x,y)\,dxdy & \end{cases}$$

Mechbal N.

RANDOM VARIABLE AND VECTOR

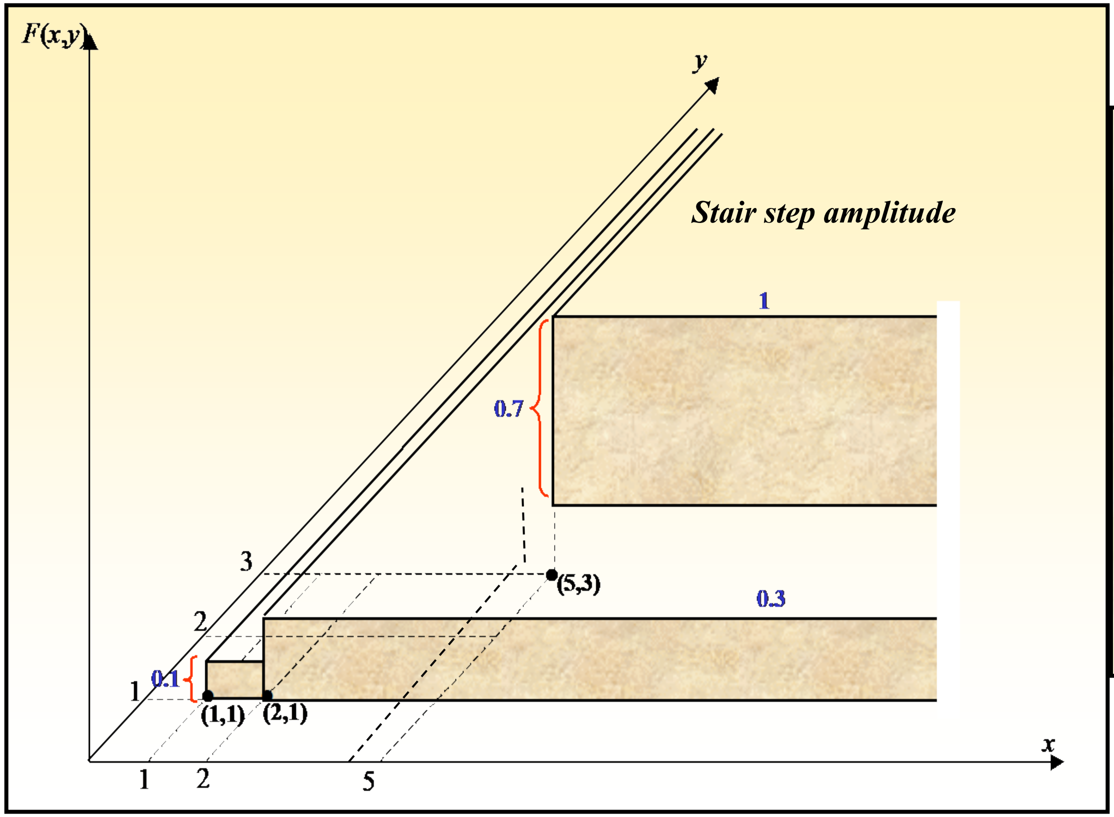
- **Joint distribution and density:**

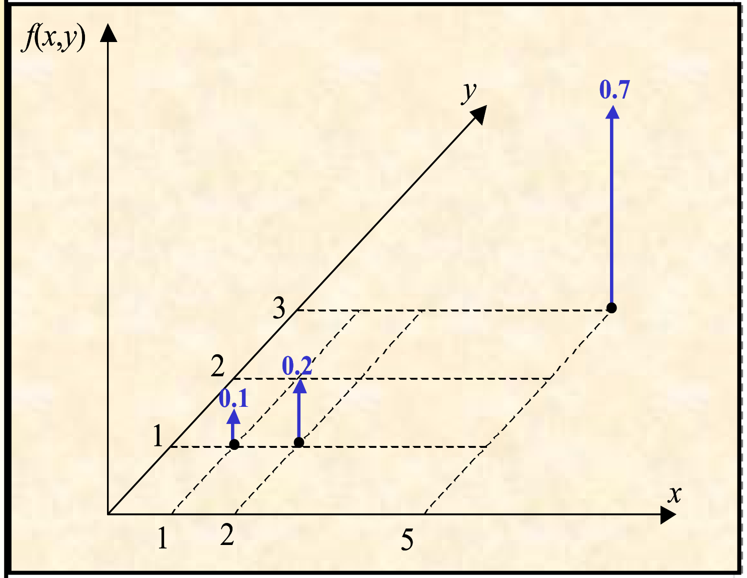


Joint distribution and density:

- Example : let a joint sample space S_c composed of 3 possible elements: (1,1), (2,1) and (5,3), having the following probability :

$$P(1,1) = 0.1,\ P(2,1) = 0.2 \text{ et } P(5,3) = 0.7$$





Mechbal N.

✷ Marginal distribution and density functions

- We are interested in the projection of the *two laws on one of the two axes*. The Marginal functions are obtained by setting the value of the other variable to infinity:

$$\begin{cases} F(x) = F(x,+\infty) = \mathrm{Prob}(X \le x \ \ \forall Y) \\ F(y) = F(\infty, y) = \mathrm{Prob}(Y \le y \ \ \forall X) \end{cases} \Rightarrow \begin{cases} f(x) = \dfrac{dF(x)}{dx} \\ f(y) = \dfrac{dF(y)}{dy} \end{cases}$$

- We have then the following properties:

Continuous	Discrete
$F(x) = \int_{-\infty}^{+\infty} d\alpha_2 \int_{-\infty}^{x_1} f(\alpha_1,\alpha_2)d\alpha_1$	$F(x) = \sum_{ij} p_{ij} \quad \forall\, j \; et \, \forall\, i \; tel \; que \, x_i < x$
$F(y) = \int_{-\infty}^{+\infty} d\alpha_1 \int_{-\infty}^{x_2} f(\alpha_1,\alpha_2)d\alpha_2$	$F(y) = \sum_{ij} p_{ij} \quad \forall\, i \; et \; \forall\, j \; tel \; que \, y_j < y$
$f(x) = \int_{-\infty}^{+\infty} f(x,y)dy$	$f(x) = \sum_{ij} p_i \delta(x - x_i) \;\; où \;\; p_i = \sum_j p_{ij}$
$f(y) = \int_{-\infty}^{+\infty} f(x,y)dx$	$f(y) = \sum_{ij} p_j \delta(y - y_j) \;\; où \;\; p_j = \sum_i p_{ij}$

Mechbal N.

Conditional distribution and density

- We are interested here in defining distribution and density function of one random variable X conditioned by values of the second random variable Y:

$$\text{Conditional probability} = \frac{\text{Joint probability}}{\text{Marginal probability}}$$

$$f(x/Y \le y) = f(x/y) = \frac{f(x,y)}{f(y)} \qquad \textit{and} \qquad f(y/X \le x) = f(y/x) = \frac{f(x,y)}{f(x)}$$

nce is defined by:

$$\text{Prob}\{X \le x, Y \le y\} = \text{Prob}\{X \le x\}\text{Prob}\{Y \le y\}$$

- It follows:

$$F(x,y) = F(x)F(y) \quad \textit{and} \quad f(x,x_2) = f(x)f(y)$$
$$F(x/Y \le y) = F(x) \quad \textit{and} \quad f(x/Y \le y) = f(x)$$
$$F(y/X \le x) = F(y) \quad \textit{and} \quad f(y/X \le x) = f(y)$$

Moments for random vector:

- Expectation of a function $g(.)$ of random variables of pdf $f(.)$:

$$E\big(g(X,Y)\big)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$$

- Joint moments:

$$m_{nk}=E(X^nY^k)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^n y^k f(x,y)dxdy$$

- Covariance: it is the joint central second order moment and is defined by:

$$C_{XY}=E\big((X-\overline{X})(Y-\overline{Y})\big)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\big(x-\overline{X}\big)\big(y-\overline{Y}\big)f(x,y)dxdy$$

- Correlation: if the means of the variables are nulls, then we have the correlation function:

$$R_{XY}=E\big(XY\big)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy \quad \Rightarrow \quad C_{XY}=R_{XY}-E\big(X\big)E\big(Y\big)$$

Properties of the covariance

- Intuitively, the covariance is a measure of the simultaneous variation of two random variables: the covariance becomes more positive for each pair of values that differ from their mean in the same direction, and negative for those who differ from their mean in the opposite direction .
- The two variables are uncorrelated or independent if:

$$R_{XY} = E(XY) = E(X)E(Y) \quad \Rightarrow \quad C_{XY} = 0$$

- If the two variables are orthogonal then

$$R_{XY} = 0 \quad \Rightarrow \quad C_{XY} = -E(X)E(Y)$$

- *If the two variable are independent, then they are uncorrelated, but the converse is not necessarily true.*
- Correlation coefficient: it is the normalized second-order moment

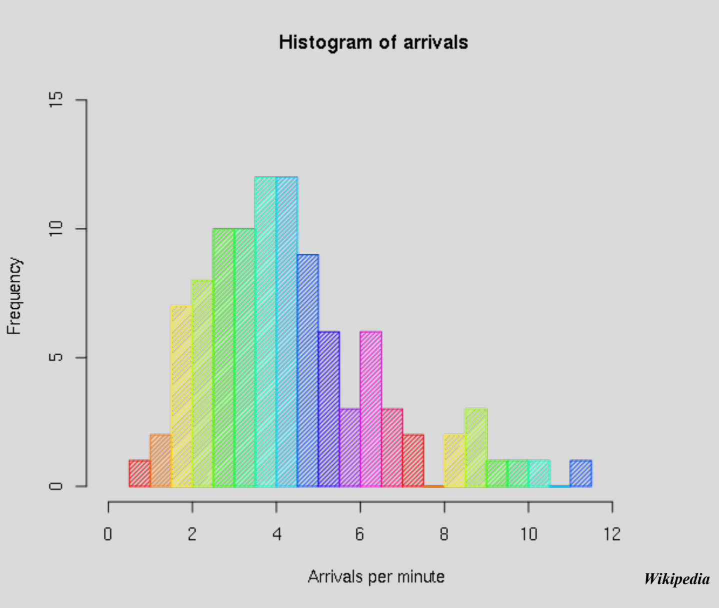
$$\rho = \frac{C_{XY}}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{C_{XY}}{\sigma_X \sigma_Y} \Rightarrow -1 \leq \rho \leq 1$$

- The correlation coefficient *measures the degree of relationship* between two random variables, if there is no relationship, the variables are called independent, and the $\rho = 0$. However, if $\rho = 1$, the variables are then related to each other by a *linear relationship*.

Mechbal N.

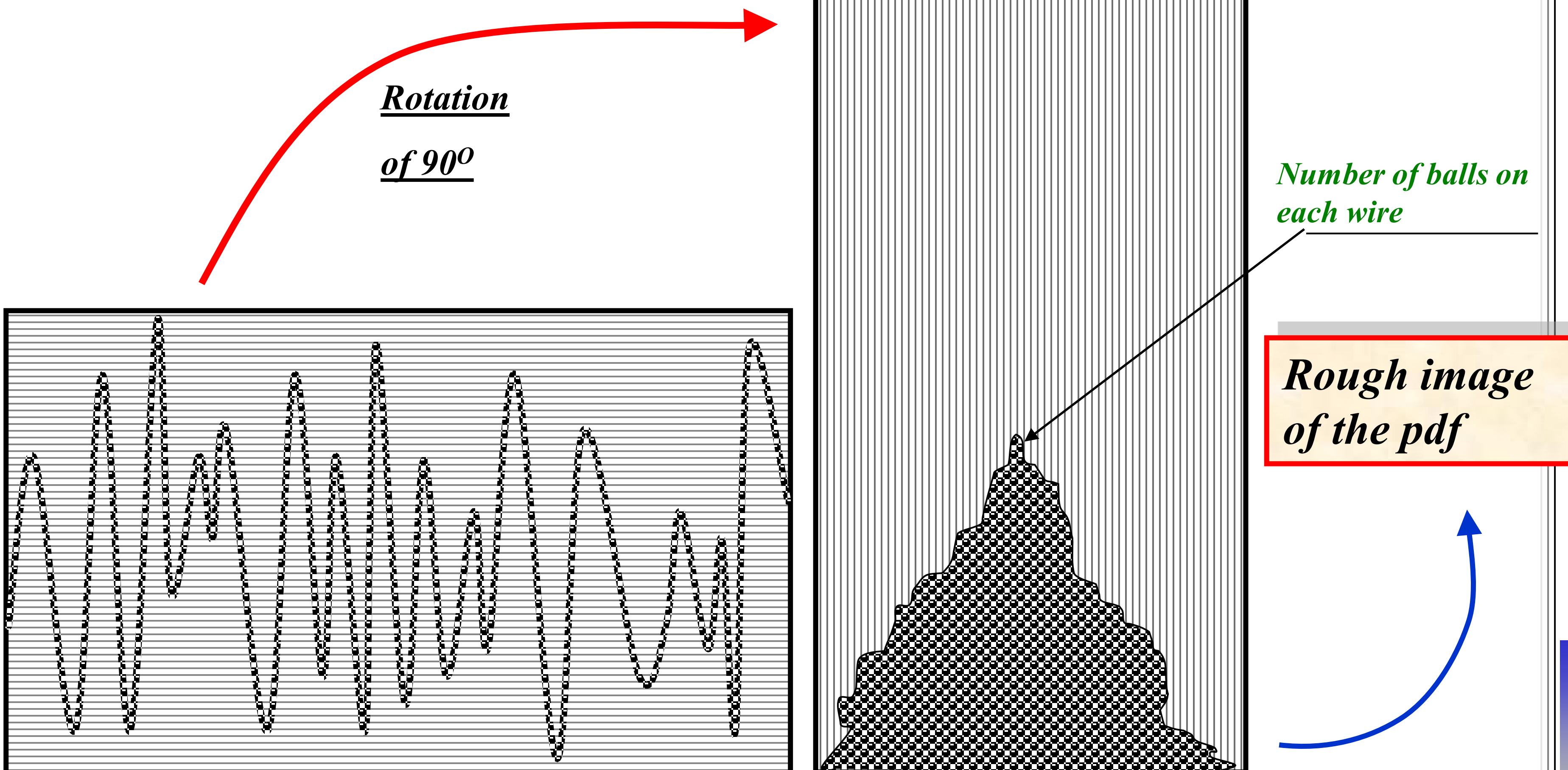
Estimation of probability density function

- The calculation of the probability density is almost always impossible. We then use histogram approach to estimate it. It is a basic tool of statisticians.
- A histogram consists of tabular frequencies shown as adjacent rectangular, erected over discrete intervals (bins), with an area equal to the frequency of the observations in the interval. The height of a rectangle is also equal to the frequency density of the interval, i.e., the frequency divided by the width of the interval. The total area of a histogram used for probability density is always normalized to 1.
- Another way of estimating pdf is kernel density estimation. A range of kernel functions are commonly used: uniform, triangular, biweight, triweight, Epanechnikov, normal, and others.



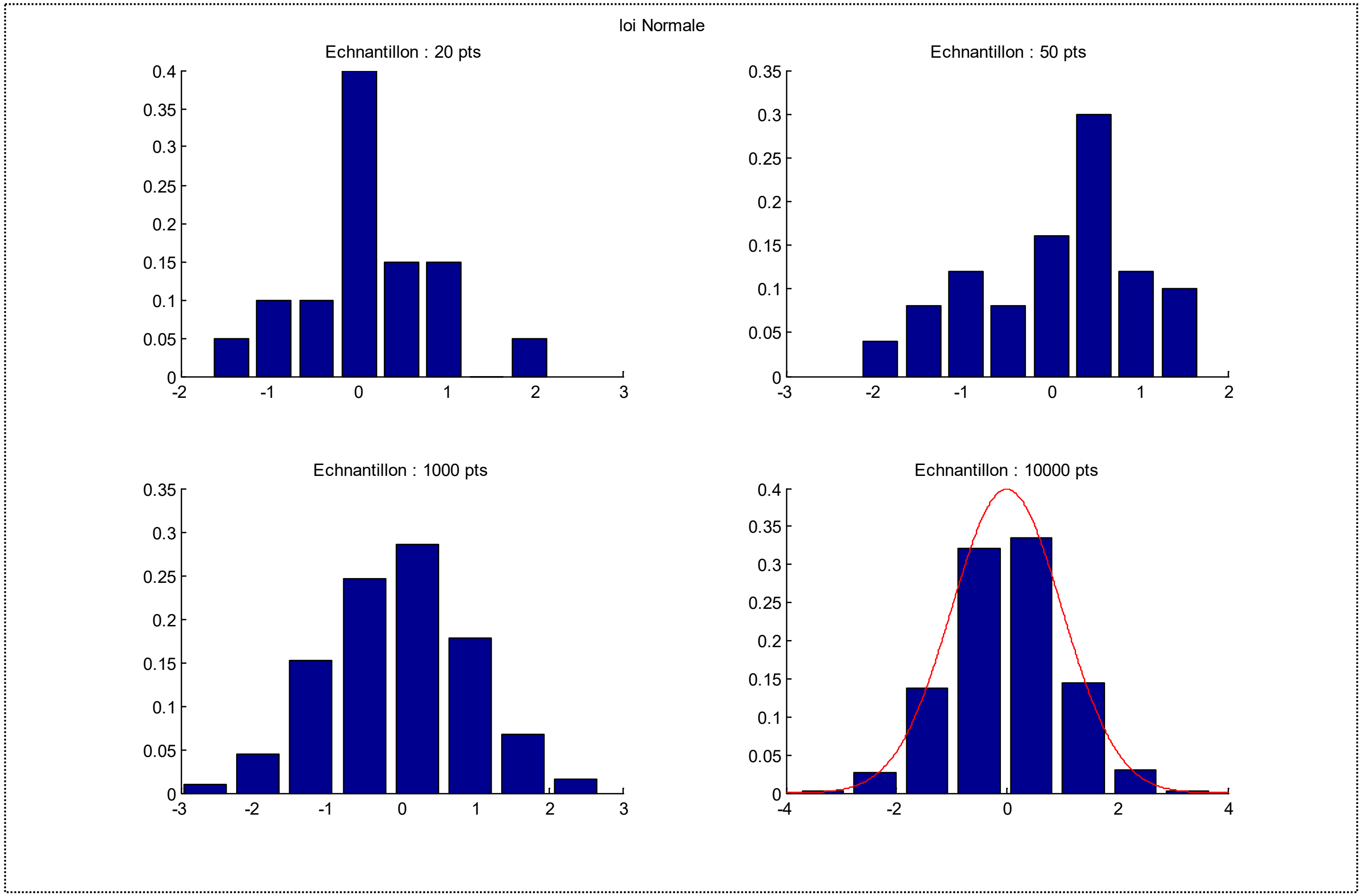
Estimation of probability density function

- Consider a function plotted on an abacus whose rods are close enough to give an impression of continuity



Estimation of probability density function

- Simulation with Matlab: influence of the number of samples



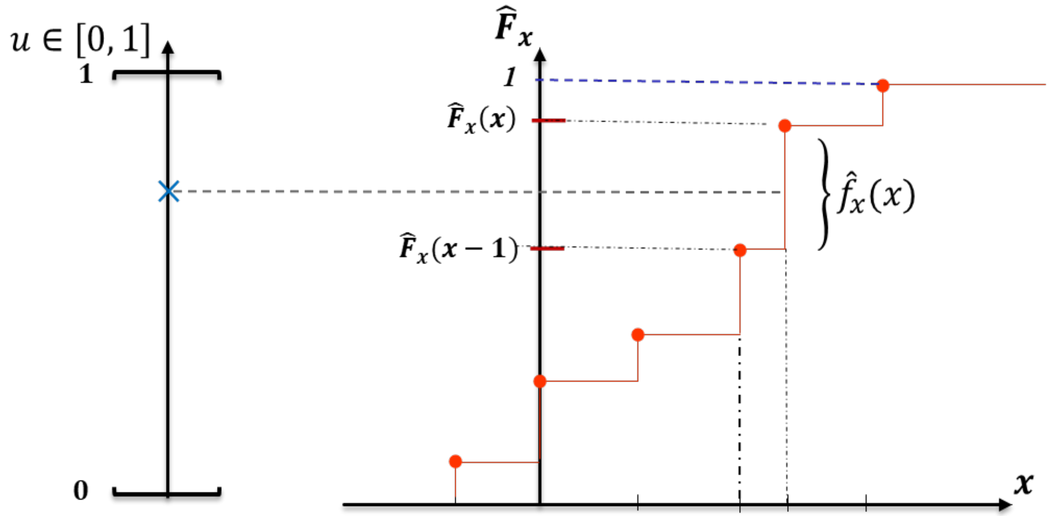
Generation of a random variable (r.v.) with an arbitrary pdf

- Most mathematical libraries have functions to generate a uniform distribution $f_u(u)$ on the interval $[0,1]$ like Matlab's ***rand.m*** function:

$$f_u(u) = \begin{cases} 1 if\ u \in [0,1] \\ 0\ otherwise \end{cases}$$

- If we now wish to generate a discrete r.v. with an arbitrary probability density (**pdf**), we can do so by inferring from this uniform pdf (see figure below):
- **Procedure**: Let there be a desired pdf $\hat{f}_x\ (x)$ associated with the random variable x :
  - ✓ We compute the cumulative distribution function (**cdf**) :

  $\hat{F}_x\ (x) = \sum_{\bar{x}=-\infty}^{x} \hat{f}_x(\bar{x})$ (in $continuous\ \ \hat{F}_x(x) = \int_{-\infty}^{x} f_v(\bar{x})d\bar{x}$)
  - ✓ We generate the r.v. u of the uniform distribution $f_u(u)$
  - ✓ We compute the value x such that $\hat{F}_x(x-1) < u \leq \hat{F}_x(x)$
  - ✓ It is then shown that the randon variable x is a random variable with pdf $f_x(x) = \hat{f}_x\ \ (x)$.



Central limit theorem

- Sum of two random variables: $W = X + Y$
  - The resulted expectation is the sum of the expectation of each variable

$$E(w) = E(x) + E(y)$$

This property generalizes to higher moments if the random variables X and Y are independent and centered.

  - Pdf of the sum:

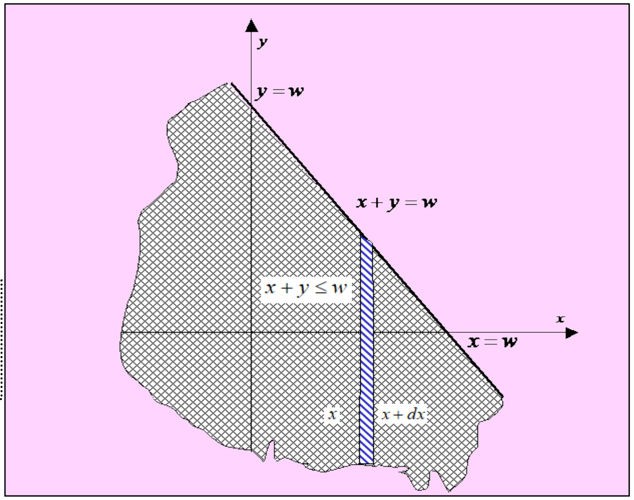
$$F(w) = \mathrm{Prob}\{W \le w\} = \mathrm{Prob}\{x + y \le w\}$$

$$F(w) = \int_{-\infty}^{+\infty} \mathrm{Prob}\{x \le X \le x + dx, Y \le w - x\}$$

$$F(w) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{w-x} f(x,y)dy \right] dy$$

derivation

$$f(w) = \int_{-\infty}^{+\infty} f(x, y = w - x)dx$$



  - If the variables are independent, then the pdf is the convolution product of the pdf of each variable

$$f(w) = \int_{-\infty}^{+\infty} f(x) f(w - x)dx$$

$$f(w) = f(x) * f(y)$$

Central limit theorem

- Statement: "*The probability distribution of the sum of a large number of random variables statically independent approaches a Gaussian distribution*".

let $X_1,\ldots,X_n$*,be n Independent Identically Distributed (i.i.d) random variables with finite first and second order moment (m et* σ^2*). Then as* $n \to \infty$*, the following random variable:*

$$Y_n = \frac{\sqrt{n}}{\sigma}\left(\frac{X_1 + \ldots + X_n}{n} - m\right)$$

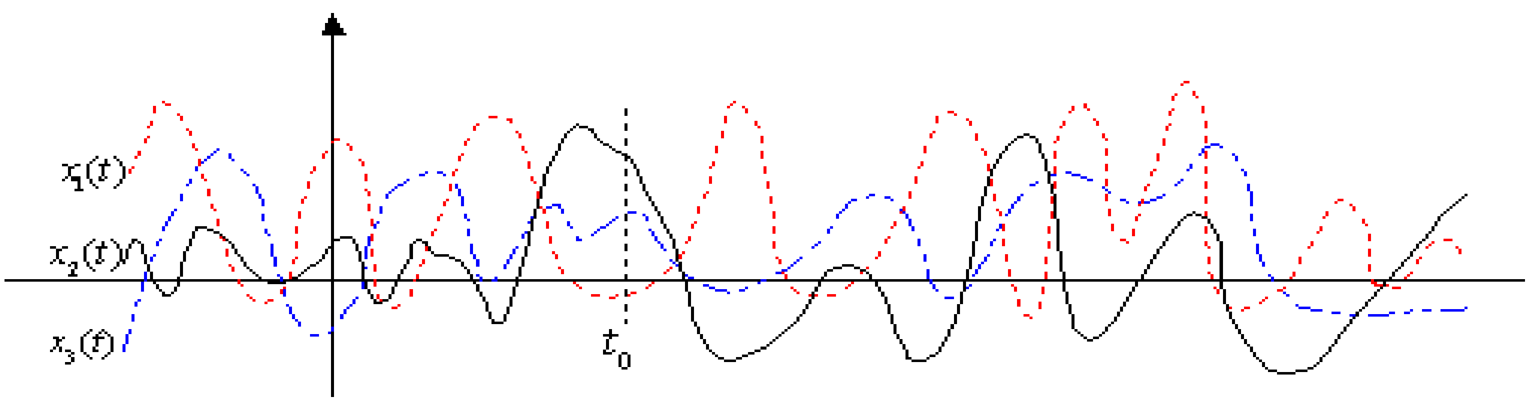
tends, to a Gaussian normal variable $N(0,1)$.

- Indeed, this theorem explains why the Gaussian character is most *prevalent in physics*. Often the observed random variable is the sum of a large number of weak random variables of any identical distribution. In practice, many random phenomena come from the *superposition* of many phenomena which we do not know the exact distribution. It is then convenient to assimilate the resulting distribution to a Gaussian. For example, *measurement errors*, background noise receivers, etc. ..

Concept

- In engineering science, we deal with random time waveform variables. A random variable is a function of possible outcomes, ω, of an experiment that will be now also a function of time. *The idea is to enlarge the random variable concept to include* ***time****.*
- Random process $X(t,\omega)$ is then the family of functions when t, and ω are variables.

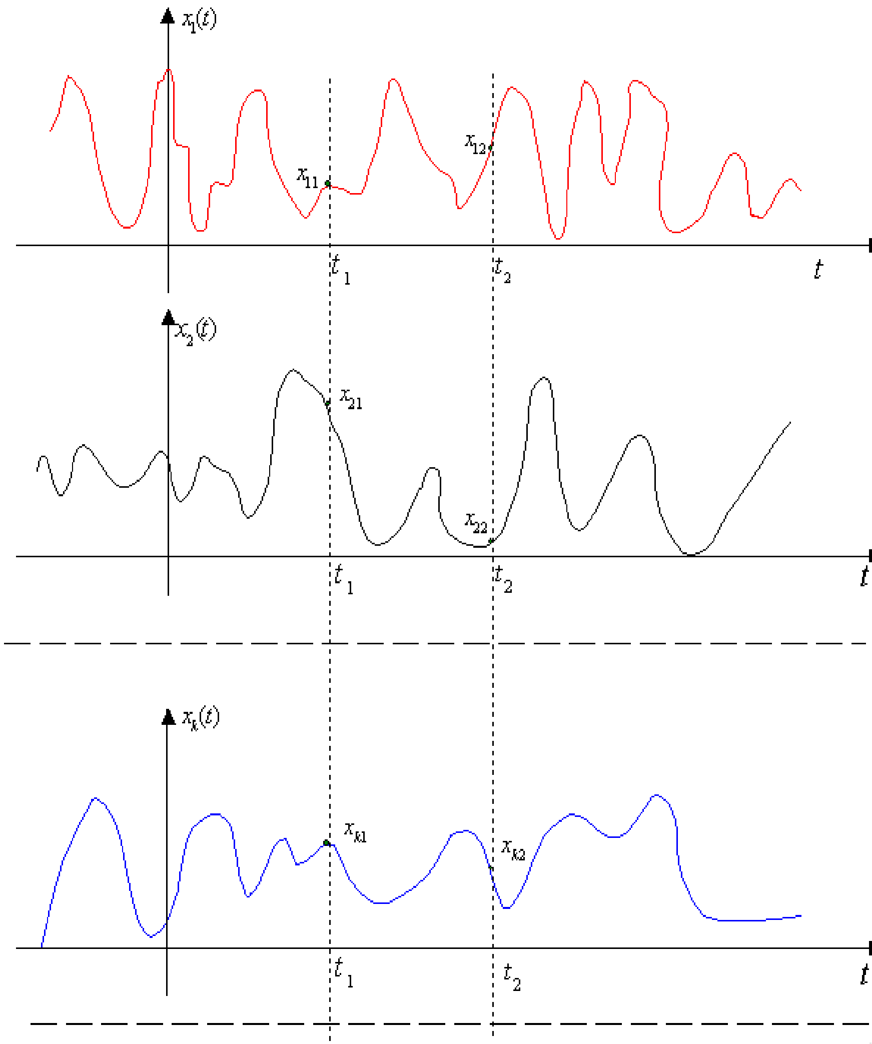
$$\begin{array}{ccc} X(\omega) & \rightarrow & X(t,\omega) \\ \downarrow & & \downarrow \\ x & \rightarrow & x(t) \end{array}$$



✷ **Concept: A random process can be seen**

- As a single time function or realization of the process when t is a variable and ω is fixed at a specific value (outcome)a variable and

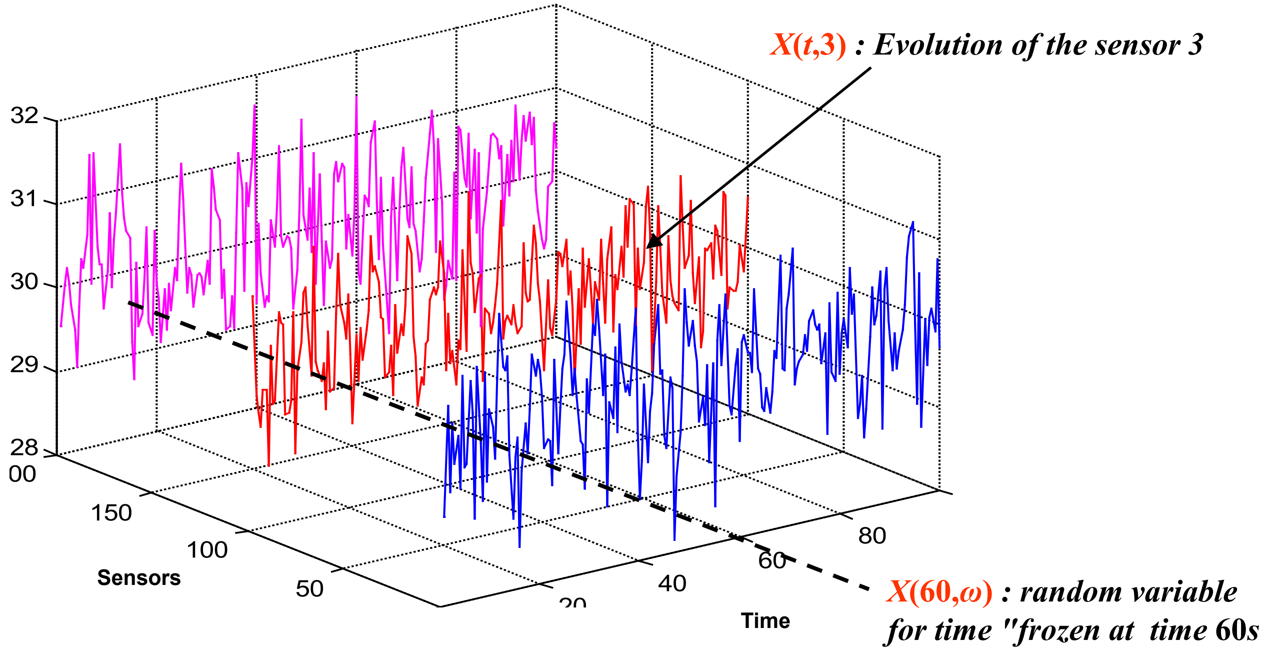
- As a random variable when time t is fixed (frozen) and ω is a variable.



Mechbal N.

Example:

- Consider the case of measuring the altitude of a space launcher. This measure is obtained by merging information from different sensors and each sensor is also redundant. This redundancy provides a natural alternative in case of failure and more accurate measurements:
  - All events or realizations consists of redundant sensors.
  - In steady state, all these sensors must indicate the same value but in practice they are affected by measurement noises.



Mechbal N.

Distribution and density functions

- Distribution:

$$F_X(x_1,t_1) = \text{Prob}\{X(t_1) \le x_1\}$$

- Density:

$$f_X(x_1,t_1) = dF_X(x_1,t_1)/dx_1$$

- For N random variables:

$$X(t_i) \quad i = 1,\ldots,N$$

$$F_X(x_1,t_1;\ldots;x_N,t_N) = \text{Prob}\{X(t_1) \le x_1,\ldots,X(t_N) \le x_N\}$$

$$f_X(x_1,t_1;\ldots;x_N,t_N) = \partial^N F_X(x_1,t_1;\ldots;x_N,t_N)/(\partial x_1 \cdots \partial x_N)$$

Statistical independence

$$f_{X,Y}(x_1,t_1;\ldots;x_N,t_N;y_1,t_1';\ldots;y_N,t_N') = f_X(x_1,t_1;\ldots;x_N,t_N).f_Y(y_1,t_1';\ldots;y_N,t_N')$$

Moments

- Mean value of the random variable $X(t)$ at time t_1

$$\overline{X}(t_1) = \lim_{N\to\infty} \frac{x(t_1)+x_2(t_1)+\ldots+x_N(t_1)}{N}$$

$$E\left[X(t_1)\right] = \sum_{i=0}^{\infty} f_X(x_i,t_1)x_i(t_1)$$

$$\overline{X}(t_1) = E\left[X(t_1)\right] = \int_{-\infty}^{+\infty} x(t_1) f_X(x,t_1)\,dx(t_1)$$

- Second order moment

$$\sigma_X^2(X(t_1)) = E\left[\left(X(t_1)-\overline{X}(t_1)\right)^2\right] = E\left[X^2(t_1)\right] - \overline{X}^2(t_1)$$

- nth-order moment

$$\boxed{\overline{\left[X(t_1)\right]^n} = E\left[X(t_1)^n\right] = \int_{-\infty}^{+\infty} x^n(t_1) f_X(x_1,t_1)\,dx(t_1)}$$

- *Note that the moments of a random function depends on the parameter* ***t***

Mechbal N.

Stationary process

- A random process is said to be stationary to order N if its Nth-order density function is invariant to a time origin shift; that is, if

$$X(t_i) \quad i=1,\ldots,N \Rightarrow f_X(x_1,\ldots,x_N;t_1,\ldots,t_N) = f_X(x_1,\ldots,x_N;t_1+\Delta,\ldots,t_N+\Delta)$$

- Second-order stationary process: the first and the second moments are not function of time.

$$f_X(x_1,x_2;t_1,t_2) = f_X(x_1,x_2;t_1+\Delta,t_2+\Delta) \Rightarrow \begin{cases} E[X(t_1)] = E[X(t_2)] \quad and \quad E[X^2(t_1)] = E[X^2(t_2)] \\ E[X^n(t_1)] \neq E[X^n(t_2)] \quad fo \;\; n>2 \end{cases}$$

- Wide-sense stationary process: it is a process that satisfied the following two conditions:

$$\boxed{\begin{cases} E[X(t)] = \overline{X} \\ E[X(t)X(t+\tau)] = R_{XX}(\tau) \end{cases}}$$

- With the stationarity, the parameters measured after a time t_0, can be considered valid in any moment of the evolving of the process.

⇨ *Thus, the possibility of "**prediction**" on certain characteristics of the process. Although highly variable itself, its static characteristics have a constant value, so* ***predictable*** *!!.*

Mechbal N.

Exercise

- *Consider a random second order stationary function* $X(t)$ *and the following functions:*

$$Y(t) = X(t)\cos(\omega t + \varphi)$$
$$Z(t) = X(t)\cos(\omega' t + \varphi)$$

with ω' *and* ω *are constants and* φ *is a uniform distributed random variable on* $[0,2\pi]$. $X(t)$ *and* φ *are independents variables.*

- *Show that*
  - $\boldsymbol{Y(t)}$ **is a second order stationary function**
  - $\boldsymbol{Y(t) + Z(t)}$ **is not second order stationary function**

Correlation functions

- We will, as in the case of random variables, introduce tools for measuring the *similarity* between values of the same or several random functions at given times
- Autocorrelation: it is the correlation of two random variables $X(t_1)$ and $X(t_2)$ defined by the random process $X(t)$ at times t_1 and t_2:

$$\boxed{R_{XX}(t_1,t_2)=E\left[X_1X_2\right]}$$

- We assign

$$\boxed{t_1=t \quad and \quad t_2=t_1+\tau \quad \Rightarrow \quad R_{XX}(t,t+\tau)=E\left[X(t)X(t+\tau)\right]}$$

- In the case of wide-sense stationary process the autocorrelation function rely only on time difference τ:

$$\boxed{\tau=t_1-t_2 \quad \Rightarrow \quad R_{XX}(\tau)=E\left[X(t)X(t+\tau)\right]}$$

- Cross-correlation: for two random process $X(t)$ and $Y(t)$ we have:

$$\boxed{R_{XY}(t,t+\tau)=E\left[X(t)Y(t+\tau)\right] \;\; or \;\; R_{YX}(t,t+\tau)=E\left[X(t+\tau)Y(t)\right]}$$

Mechbal N.

Correlation functions

- Consider two stationary physical function $x(t)$ and $y(t)$ with a known history over a horizon T (recording). *We can then wonder whether with the knowledge from the stories of these two functions, one can infer a relationship between them*.
- Assume that there is a linear relationship between these two functions, *i.e.* there is a coefficient, a, such as the following criterion is minimized:

$$\overline{\left|\varepsilon(t)\right|^2} = \overline{\left|x(t) - ay(t)\right|^2} \qquad \textit{with} \qquad \overline{x(t)} = \frac{1}{T}\int_{t_0}^{t_0+T} x(t)dt$$

- In fact, the minimization of this criterion will lead to choose a factor a that will ensure that $x(t)$ and $ay(t)$ are superimposed on the best, in the entire time interval. After minimization, this coefficient is given by:

$$a = \frac{\overline{x(t)y(t)}}{\overline{y(t)^2}}$$

- By introducing a correlation coefficient ρ, on have

$$\overline{\left|\varepsilon(t)\right|^2} = \overline{\left|x(t)\right|^2}\left(1-\rho^2\right) \qquad \textit{with} \qquad \rho = \frac{\overline{x(t)y(t)}}{\sqrt{\overline{x(t)^2}\;\overline{y(t)^2}}}$$

- If $\rho = 1$ *then the two function are correlated. But if* $\rho=0$, this **doesn't mean that the two functions are *independent*, *but they do not have, on average, energy of interaction.***

Correlation functions

- We can now wonder if it is not preferable to compare the two functions $x(t)$ and $y(t)$ after time shifting, which is to consider $x(t)$ and $y(t\text{-}\tau)$. We then define, in the case where the signals are stationary, the correlation function or cross-correlation by:

$$\boldsymbol{R}_{\boldsymbol{xx}}(\tau) = \frac{1}{T}\int_{t_0}^{t_0+T} x(t)x(t-\tau)dt \ \text{ and } \ \boldsymbol{R}_{\boldsymbol{xy}}(\tau) = \frac{1}{T}\int_{t_0}^{t_0+T} x(t)y(t-\tau)dt$$

- *This autocorrelation function allows to see how the function at a given time is influenced by what happened at a previous time. It is an image of the memory of a process.*

- Indeed, if for example $x(t)$ has a periodicity then each time the shift τ will be equal to an integer times the period of this periodicity, we will have a maximum of similarity and hence a maximum of the autocorrelation function. *This property has been used to detect the periodicity of the signals.*

✷ **Correlation function properties:** ***If X(t) and Y(t) are two*** **second-order stationary process, then:**

- $R_{XX}(\tau) = R^{*}_{XX}(\tau)$ — ***Symmetry***
- $\left|R_{XX}(\tau)\right| \leq R_{XX}(0)$ — ***Bonded byits value at the origin***
- $R_{XX}(0) = E\left[\left|X^2(t)\right|\right]$ — ***Mean-squared value or power in the process.***

$$\boxed{Si\ \ \overline{X} = 0 \ \Rightarrow \ R_{XX}(0) = \mathrm{var}(X(t))}$$

- $\tau \to \infty \ \Rightarrow \ X(t)$ et $X(t-\tau)$ are becoming uncorrelated

$$\Rightarrow \lim_{\tau\to\infty} R_{XX}(\tau) = E\left[X(t)\right]E\left[X(t-\tau)\right] = \overline{X}^2$$

The present is not influenced by the very distant past

- $\Im\left[R_{XX}(\tau)\right] \geq 0$ — ***Fourier transform is always positive***

Mechbal N.

- $R_{XY}(\tau) = R^{*}_{YX}(-\tau)$
- $|R_{XY}(\tau)| \leq \sqrt{R_{XX}(0)R_{YY}(0)}$
- $R_{XY} = 0$ ⟷ ***X(t) and Y(t) are orthogonal***
- $R_{XY}(t, t+\tau) = E[X(t)]E[Y(t+\tau)]$ ⟶ ***X(t) and Y(t) are Independent***

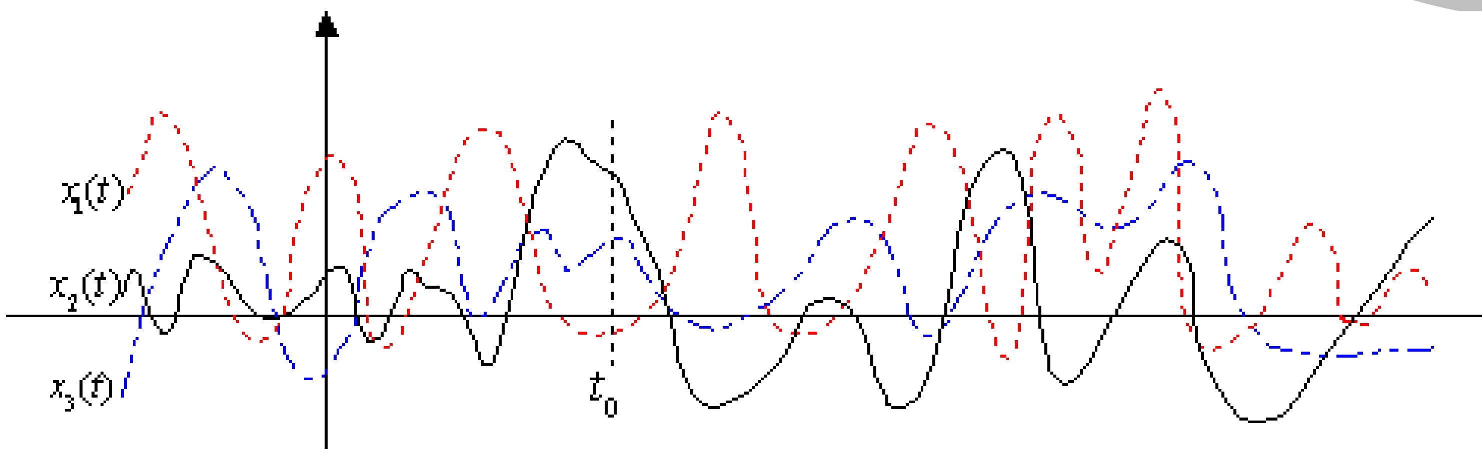
Example :

Consider the autocorrelation function of a stationary process X(t):

$$R_{XX}(\tau) = 25 + \frac{4}{1+6\tau^2}$$

Give the mean and the variance of X(t).

✷ **Concept :**



Problems: How to calculate statistical averages of random processes ?

Solution 1. ***Perform a large number of realization at t_o*** → ***Not always EASY!!***

Solution 2. ***Consider a single realization $x_k(t)$ on a sufficiently long horizon.*** → ***$x_k(t)$ can take all possible values, i.e. all possible values of $x(t_o)$!!***

Ergodicity of a random function

- Time averages :

$$\langle x(t)\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)dt$$

random process for which the time average of one sequence of events is the same as the ensemble average, and vice versa.

Only if

Stationarity to first order of the random function $x(t)$.

$$\langle x(t)\rangle = E\{x(t)\}$$

- and …

$$\langle x^2(t)\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x^2(t)dt$$

$$\Re_{XX}(\tau) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)x(t-\tau)dt$$

Only if

Stationarity to second order of the random function $x(t)$.

Mechbal N.

✷ Spectral analysis

- As random function are not necessarily bonded function, Fourier transform cannot be applied straightforward
- The appropriate spectrum of random process will be related to the description of power in random process as a function of frequency.
- Power spectral density (**psd**): noted S_{XX} and defined by

$$S_{XX}(f) = \lim_{T\to\infty} E\left[\frac{|X(f)|^2}{2T}\right]$$

It is a *density because a power* results from its integration

- Winer-Khinchin relation: we can demonstrate that for ***wide-stationary process*** the psd is given by the Fourier transform of the autocorrelation function

$$S_{XX}(f) = \mathcal{F}[R_{XX}(\tau)] = \int_{-\infty}^{+\infty} R_{XX}(\tau)e^{-j2\pi f\tau}d\tau \qquad R_{XX}(\tau) = \int_{-\infty}^{+\infty} S_{XX}(f)e^{j2\pi f\tau}df$$

- Cross-power spectrum: we can show that using the cross-correlation function we have:

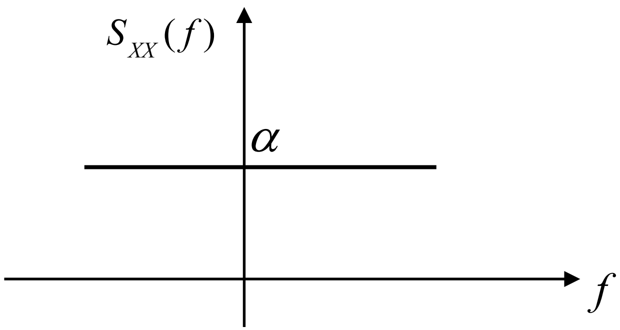
$$S_{XY}(f) = \mathcal{F}[R_{XY}(\tau)] = \int_{-\infty}^{+\infty} R_{XY}(\tau)e^{-j2\pi f\tau}d\tau$$

White Noise

- It is a random process whose spectral density is constant regardless of frequency. The word white has its origin in the analogy with white light, whose power is uniformly distributed over all the optical frequencies.
- It has a psd that is constant over frequency range and a

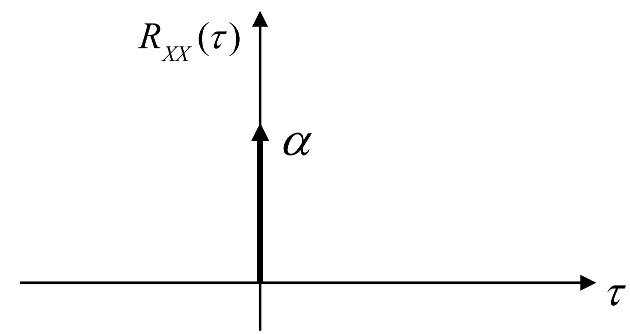
$$S_{XX}(f) = \alpha \qquad \forall f$$

$$R_{XX}(\tau) = \int_{-\infty}^{+\infty} S_{XX}(f) e^{j2\pi f\tau} df = \alpha\delta(\tau)$$



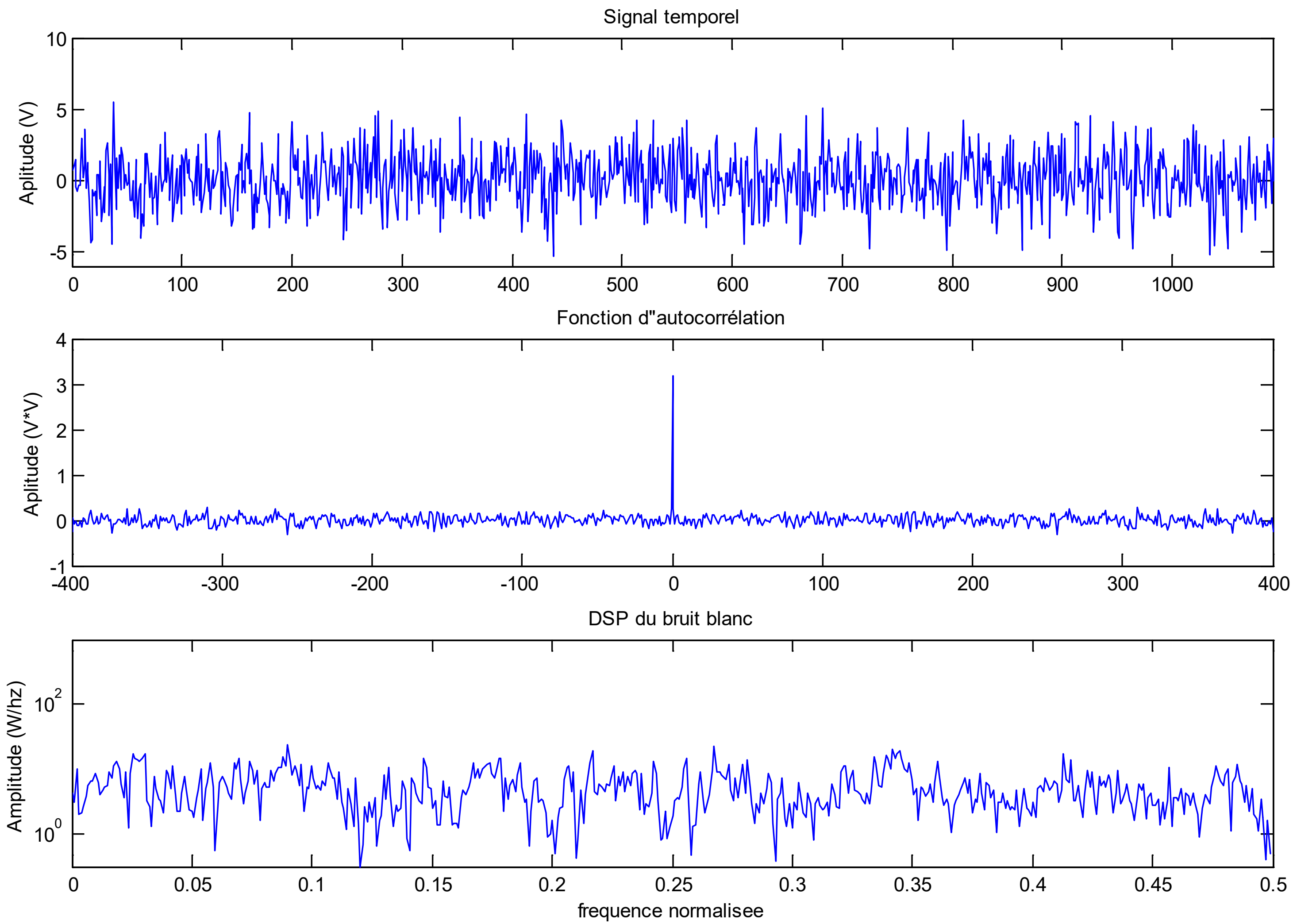
Infinite power:

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} S_{XX}(f) df = \infty$$



- It is a very important signal in identification and modulization

White noise


Signal temporel
Aplitude (V)
10
5
0
-5
0
100
200
300
400
500
600
700
800
900
1000
Fonction d"autocorrélation
Aplitude (V*V)
4
3
2
1
0
-1
-400
-300
-200
-100
0
100
200
300
400
DSP du bruit blanc
Amplitude (W/hz)
10^2
10^0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
frequence normalisee

Effect on a system

- When a signal passes through a linear system, we know that the output signal is given by the *convolution* between the input signal and the impulse response of the system. We can generalize this result to the case of random process, $X(t)$, used in a deterministic system (g(t)). We then have the random process output $Y(t)$ given by:

$$Y(t) = \int_{\tau} X(t-\tau) g(\tau) d\tau$$

- If the process is wide-stationary then we have:

$$E[Y(t)] = \int_{\tau} E[X(t-\tau)] g(\tau) d\tau = \overline{X} \int_{\tau} g(\tau) d\tau = \overline{X} G(0)$$

- We have also the following important relations (***Winer-Hopf***)

$$\boxed{\begin{aligned} S_{yy}(\omega) &= H(j\omega) S_{ee}(\omega) H^{T}(-j\omega) \\ R_{YX}(\tau) &= E\{Y(t)X(t-\tau)\} = \sum_{k=1}^{\infty} g(k) R_{XX}(k-\tau) \end{aligned}}$$

Mechbal N.

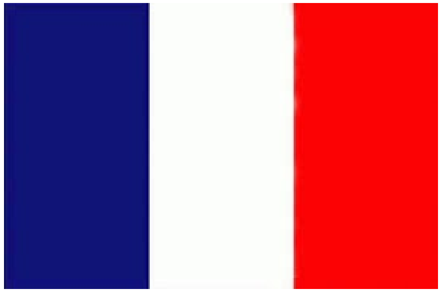
<u>Brief History:</u>

The first attempts to systematically approach the estimation problem, as it is known today, were taken by Gauss and Legendre in studying astronomical problems during the late 18th and the early 19th century. More specifically, they tried to estimate the positions of planets and comets using telescopic measurements. Gauss made use of the method of least-squares for the first time in 1795 at the age of 18. However, it was not until 1809 that he published his results in his book "Theoria Motus Corporum Celestium" (Gauss, 1809). A few years earlier, in 1805 Legendre had independently invented and published the method in his book "Nouvelles méthodes pour la determination des orbites des comètes" This gave rise to a big dispute between Gauss and Legendre, concerning who was the inventor of the least-squares method (Sorenson, 1970). A thorough discussion of the early contributions to estimation theory is provided by Seal (1967) and Sorenson (1970). The next major development in the study of the estimation problem came in the 1940s, with the filtering work of Wiener (1949) and Kolmogorov.

…. and The breakthrough came with the Kalman filter, Kalman(1960) ……..

<u>Estimation Theory</u>

Traduction / Translation



- *Densité de probabilité : ddp*
- *Fonction de répartition*
- *Variable / vecteur aléatoire*
- *Loi composée*
- *Loi marginale*
- *Processus aléatoire*
- *Densité spectrale de puissance*
- *Bruit Blanc*

- *Probabilité density function : pdf*
- *Cumulative probability distribution*
- *Randol Variabl / Vector*
- *Joint distribution*
- *Marginal distribution*
- *Stochastic process*
- *Power spectral density (psd)*
- *White noise*