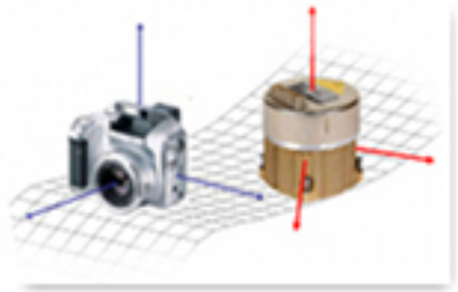
EPFL



Sensor Orientation – Random Processes

Jan SKALOUD

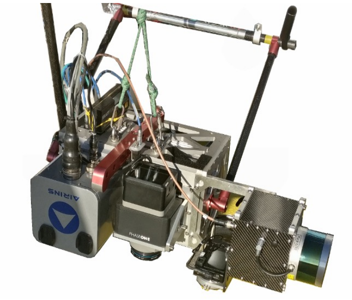
École polytechnique fédérale de Lausanne

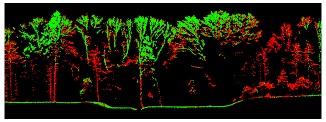
EPFL, spring semester

Sensor orientation – main topics

J. Skaloud, ESO

This translates into three rough big areas





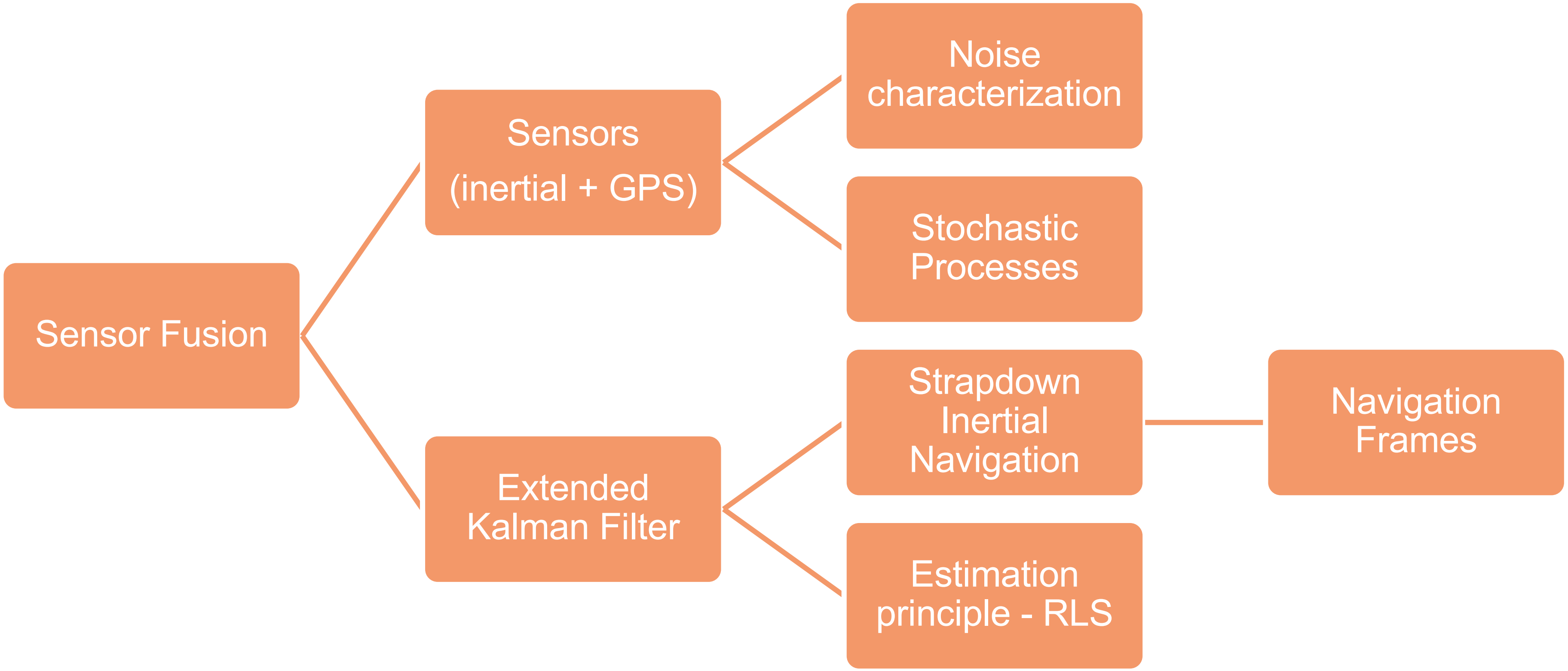
1. Fundamentals
   - How to characterize sensor noise
   - How to transform from the sensed signals to navigation frame?
2. Position, velocity, attitude (navigation)
   - How to formulate navigation equation in different frames?
   - How to resolve them numerically?
3. Sensor fusion
   - How to formulate models for sensor fusion?
   - How to implement it in optimization and use it for mapping?

You need the frames

You need the navigation quantities and the noise properties

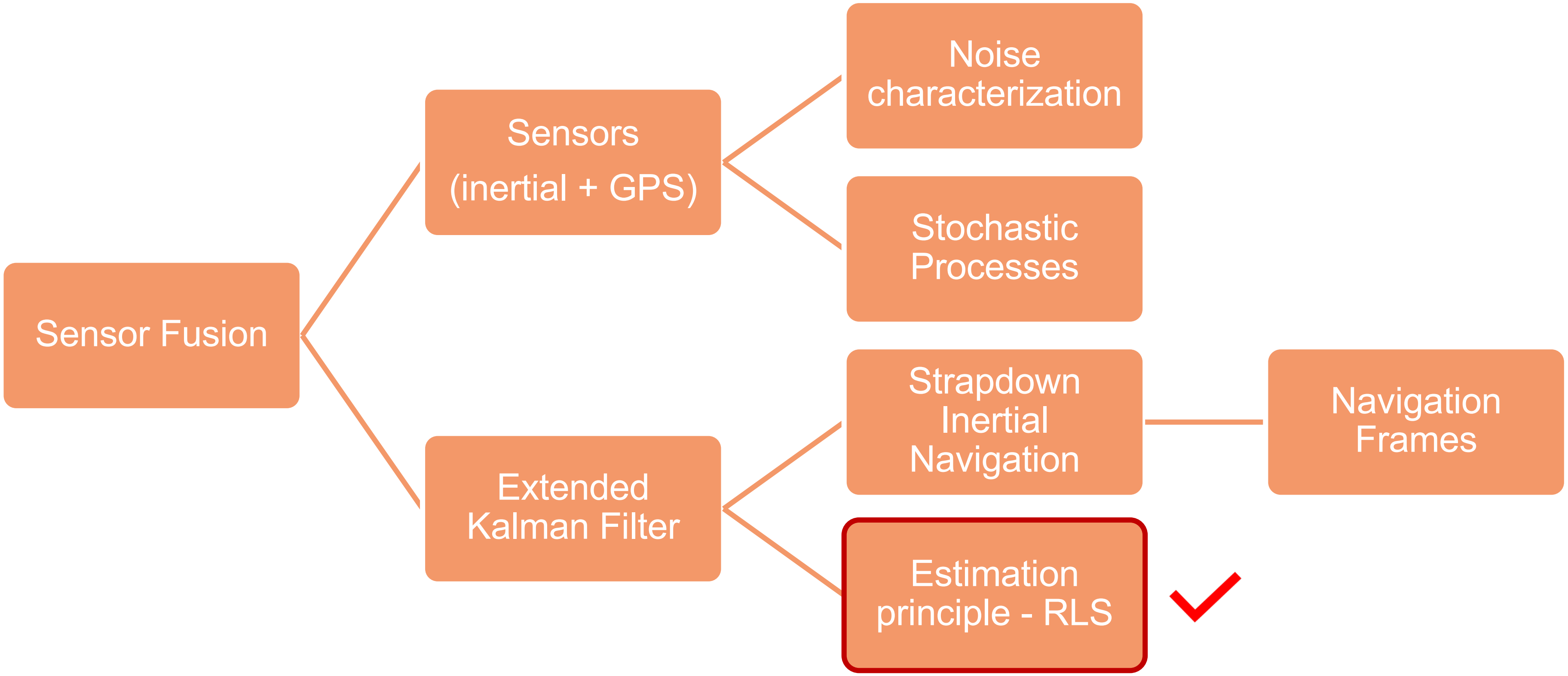
Cockpit view of SO course's topics

How to reach *integrated* sensor orientation?



Cockpit view of SO course's topics

How to reach *integrated* sensor orientation?



Similarities RLS vs. Kalman Filter (2.2.3)

Recursive Least Square (RLS)

(2.28-2.30) gain, state, covariance:

$$\begin{aligned}
\mathbf{K}_j &= \mathbf{P}_{j-1}\,\mathbf{H}_j^{\mathrm{T}}(\mathbf{H}_j\,\mathbf{P}_{j-1}\,\mathbf{H}_j^{\mathrm{T}}+\mathbf{R}_j)^{-1}\\
\hat{\mathbf{x}}_j &= \hat{\mathbf{x}}_{j-1}+\mathbf{K}_j\,(\mathbf{z}_j-\mathbf{H}_j\,\hat{\mathbf{x}}_{j-1})\\
\mathbf{P}_j &= (\mathbf{I}-\mathbf{K}_j\,\mathbf{H}_j)\,\mathbf{P}_{j-1}
\end{aligned}$$

Discrete Kalman Filter (KF)

(2.32-2.34) gain, state, covariance:

$$\begin{aligned}
\mathbf{K}_k &= \tilde{\mathbf{P}}_k\mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k\tilde{\mathbf{P}}_k\mathbf{H}_k^{\mathrm{T}}+\mathbf{R}_k)^{-1}\\
\hat{\mathbf{x}}_k &= \tilde{\mathbf{x}}_k+\mathbf{K}_k\,(\mathbf{z}_k-\mathbf{H}_k\tilde{\mathbf{x}}_k)\\
\mathbf{P}_k &= (\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\,\tilde{\mathbf{P}}_k
\end{aligned}$$

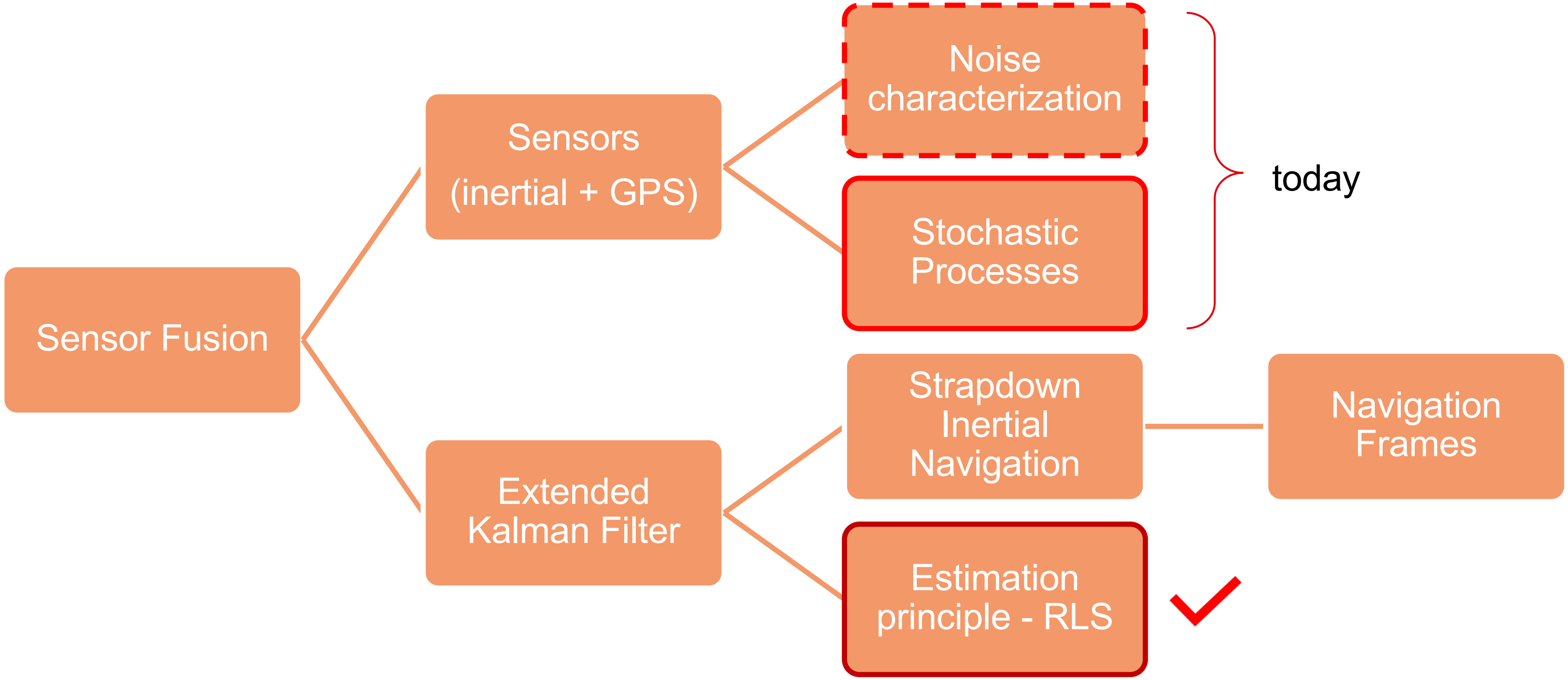
Prediction (for time update)

$$\begin{aligned}
\tilde{\mathbf{x}}_{k+1} &= \mathbf{\Phi}_k\hat{\mathbf{x}}_k\\
\tilde{\mathbf{P}}_{k+1} &= \mathbf{\Phi}_k\mathbf{P}_k\mathbf{\Phi}_k^{\mathrm{T}}+\boxed{\mathbf{Q}_k}
\end{aligned}$$

system / sensor noise !

Cockpit view of SO course's topics

How to reach *integrated* sensor orientation?

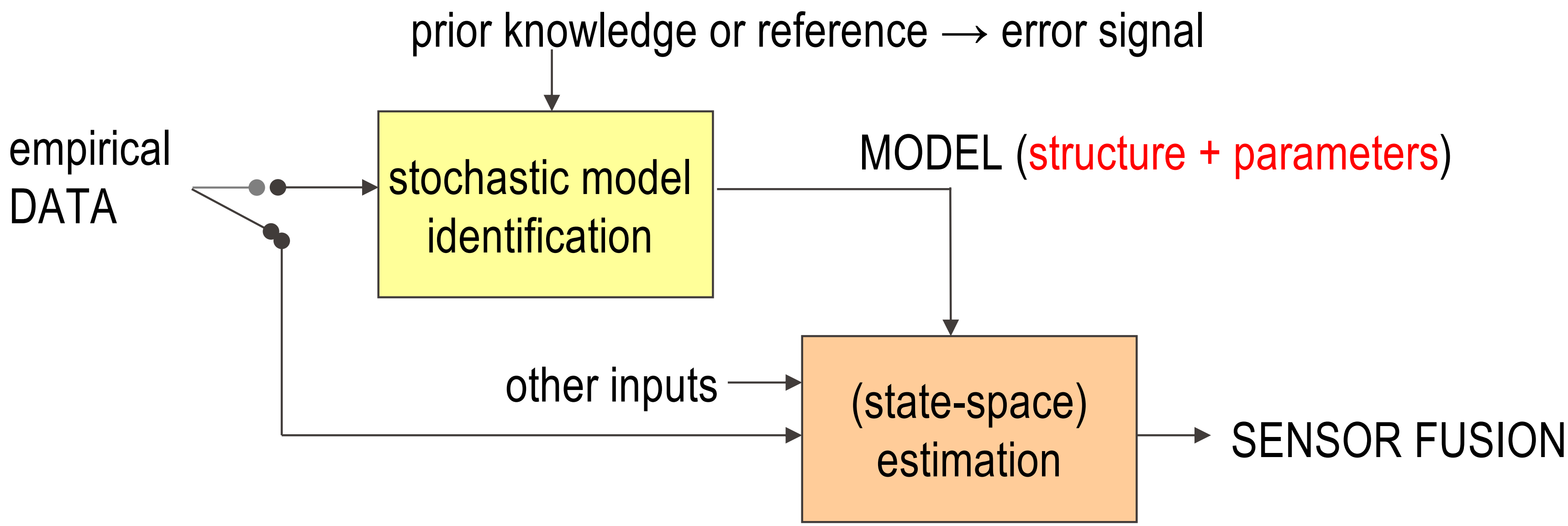


Stochastic Processes – agenda (Ch 5)

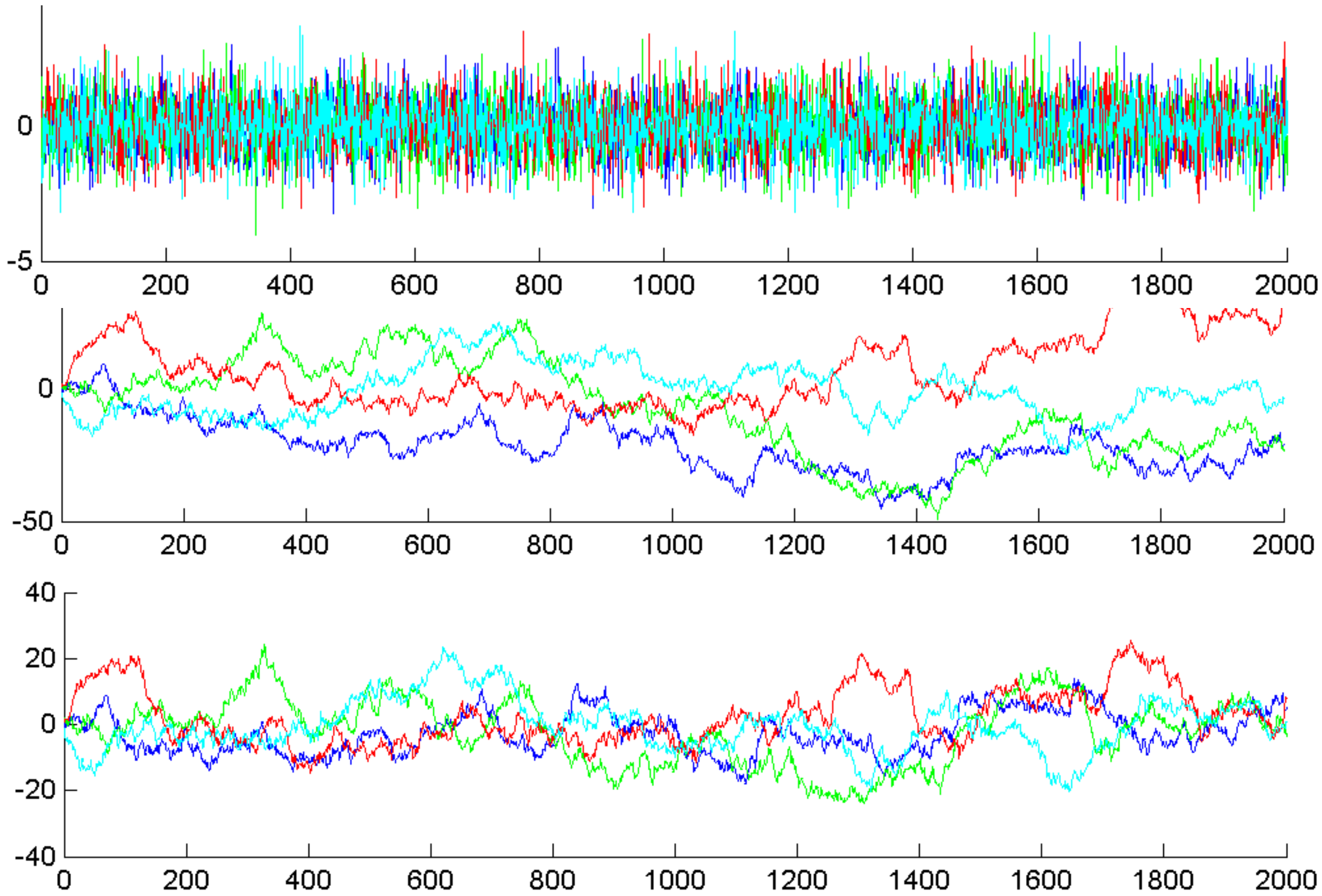
- Introduction
  - Model identification
  - Concept & classification
- Process characteristics
  - Auto-correlation (**AC**)
  - Power Spectral Density (**PSD**)
  - Allan Variance (**AV**) / Wavelet Variance (**WV**)
- The very useful processes
  - White Noise
  - Random bias
  - Random walk
  - Auto-regressive white noise (Gauss-Markov)

Noise model (process) identification

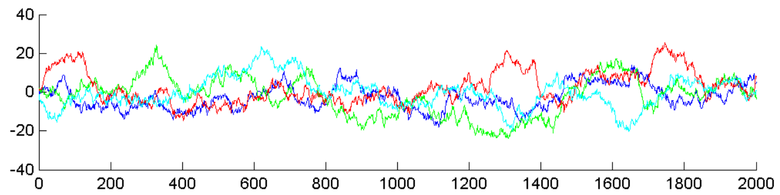
- Stochastic model
  - Prerequisite for sensor fusion (e.g. Kalman filtering)
  - Two elements needed:
    - Structure of noise (type of parameters)
    - Values of parameter
- How to identify (and later use) the model (in sensor fusion) ?



Random process – an example



Random processes – general concept



Association of random variable(s) with a deterministic parameter (e.g. time)

Stochastic **collection** – values depend on the outcome of an experiment

Time is a **parameter,** but the process is only LOOSELY (stochastic) function of time

Can be either *discrete* or continuous

Stochastic Processes – agenda (Ch 5)

- Introduction
  - Model identification
  - Concept & classification
- Process characteristics
  - Auto-correlation (**AC**)
  - Power Spectral Density (**PSD**)
  - Allan Variance (**AV**) / Wavelet Variance (**WV**)
- The very useful processes
  - White Noise
  - Random bias
  - Random walk
  - Auto-regressive white noise (Gauss-Markov)

Connections to random variables

Random variable

Probability density

$$f(x) \quad f(y)$$

Expectation

$$E(x) = \mu_X = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$E(y) = \mu_Y = \int_{-\infty}^{\infty} y\, f(y)\, dy$$

Covariance

$$cov(x,y) = E\left[(x-\mu_X)(y-\mu_Y)\right]$$

Variance

$$var(x) = cov(x,x) = E\left[(x-\mu_X)^2\right]$$

Random process

Probability density

$$f(x_i,t_i), f(x_j,t_j) \longrightarrow f_2(x_i,x_j)$$

Expectation

$$\mu_X(t) = \int_{-\infty}^{\infty} x_t\, f(x_t)\, dx_t$$

Joint expectation – autocovariance

$$\mathrm{acov}_X(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1-\mu_{X_1})(x_2-\mu_{X_2}) f(x_1,x_2)\, dx_1\, dx_2$$

Variance

$$\mathrm{acov}_X(\tau=0) = \varphi_{xx}(0) = \sigma_X^2$$

Other characteristics

Random variable

Probability density

$f(x)$

Characteristic function (Fourier transform of probability density)

$$g(\omega) = E\left[e^{-i\omega x}\right] = \int_{-\infty}^{\infty} e^{-i\omega x} f(x)\, dx$$

Random process

Probability density

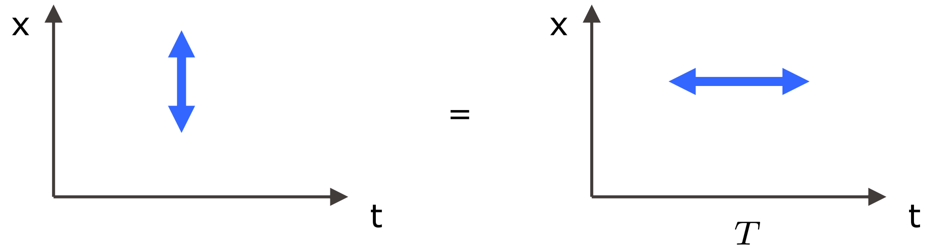
$$f(x_i, t_i), f(x_j, t_j) \longrightarrow f_2(x_i, x_j)$$

Power spectral density – PSD (Fourier transform of autocovariance function)

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} \varphi_{xx}(\tau)\, e^{-i\,2\pi\omega\,\tau} d\tau$$

Classification of random processes

- By **ergodicity***



- Mean

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

- Variance

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

- Autocovariance

$$\varphi_{xx}(t_i, t_j) = \lim_{t\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t_i)x(t_j)dt$$

* Process characteristics computed over *one long process* are expected to be the same as those computed from *many processes sampled at a specific time*

Classification of random processes

By **stationarity**

- Process characteristics are *”constant” in time* (t)
- i.e. can be computed from a ”sample interval” (rather than absolute time)
- in otherwords: what is *important* the <u>difference of time</u> τ

$$\varphi_{xx}(\tau) = E\left(x(t)\, x(t+\tau)\right)$$

$$\varphi_{xx}(0) = E(x^2)$$

$$\varphi_{xx}(\tau) = \varphi_{xx}(-\tau)$$

Autocorrelation function (AC)

Autocorrelation function:

- Assumption **1**: Underlying process is **stationary** Gaussian random process
- Assumption **2**: Data (z) are sampled at **constant** Δt, at a time "lag" $(\ell = 1) \cdot \Delta t$

$$m = \frac{1}{N} \sum_{i=1}^{N} z_i$$

$$\varphi_{zz}(\ell \Delta t) = \frac{1}{N-\ell-1} \sum_{i=1}^{N-\ell} (z_i - m)(z_{i+\ell} - m)^T \quad \ell \in [0, 1, \ldots, N-2]$$

Other time series (or frequency) analysis (later):

- AR, MA, ARMA, ARIMA, … + estimators

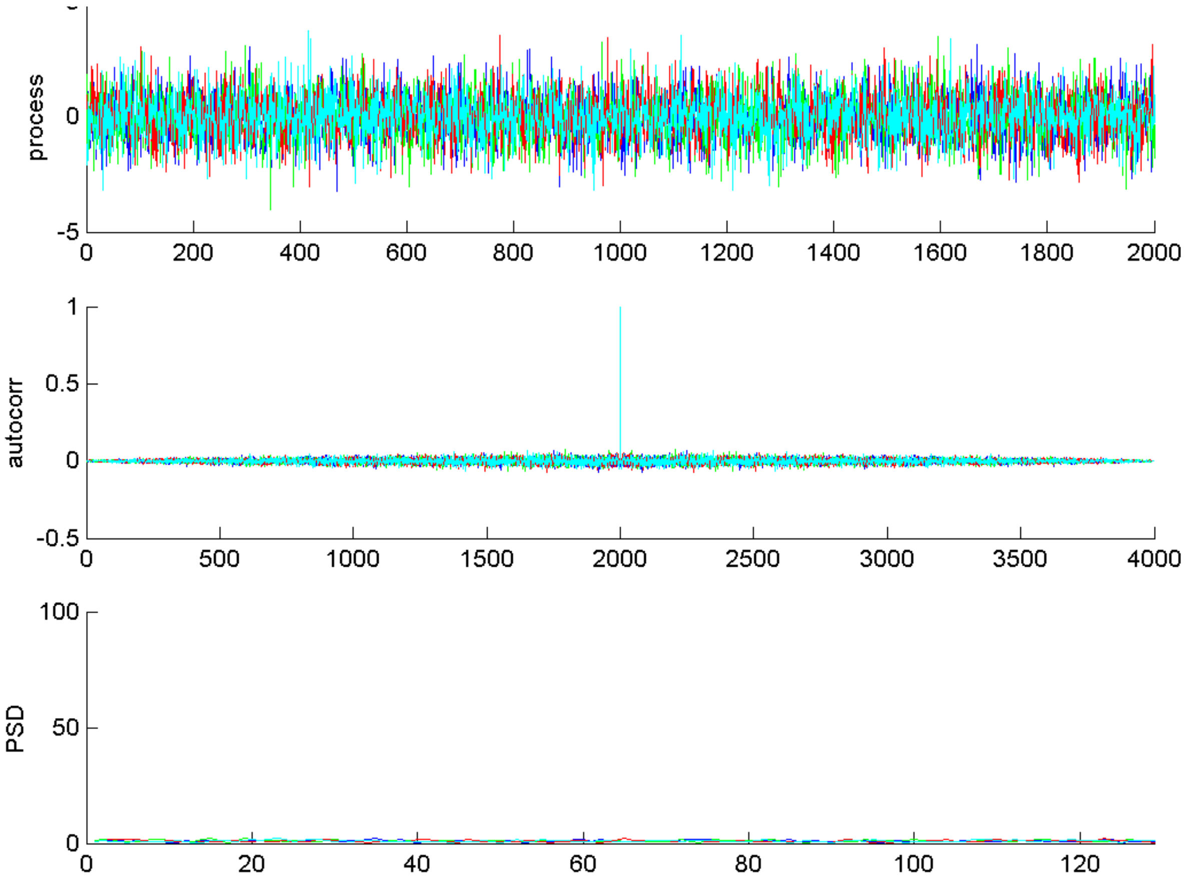
Stochastic Processes – agenda (Ch 5)

- Introduction
  - Model identification
  - Concept & classification

- Process characteristics
  - Auto-correlation (**AC**)
  - Power Spectral Density (**PSD**)
  - Allan Variance (**AV**) / Wavelet Variance (**WV**)

- The very useful processes
  - White Noise
  - Random bias
  - Random walk
  - Auto-regressive white noise (Gauss-Markov)

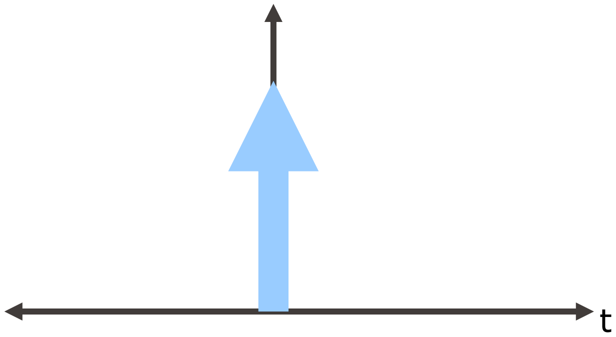
White noise - example



White noise process

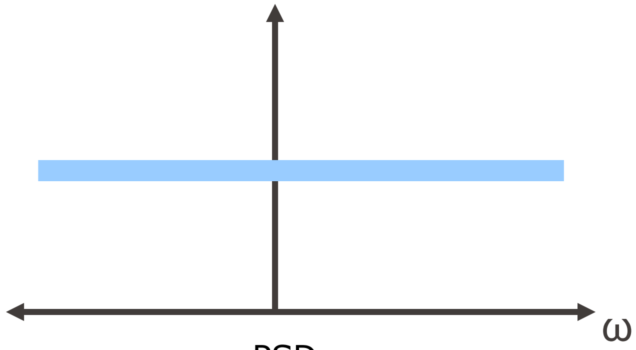
- State propagation: $$x_k = w_k$$
- Differential equation: theoretically not differentiable

$$\varphi_{xx}(\tau) = \varphi_{xx}(0)\,\delta(\tau)$$



Autocorrelation

$$\Phi_{xx}(\omega) = \sigma^2$$



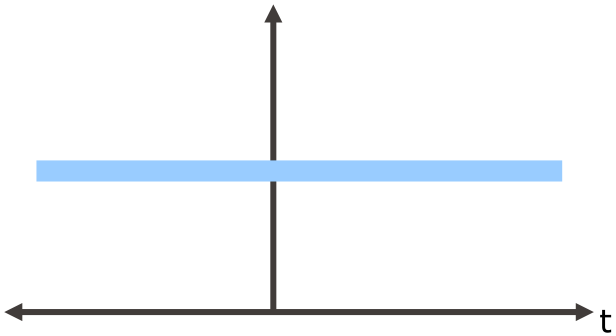
PSD

J. Skaloud, ESO

Random bias (constant)

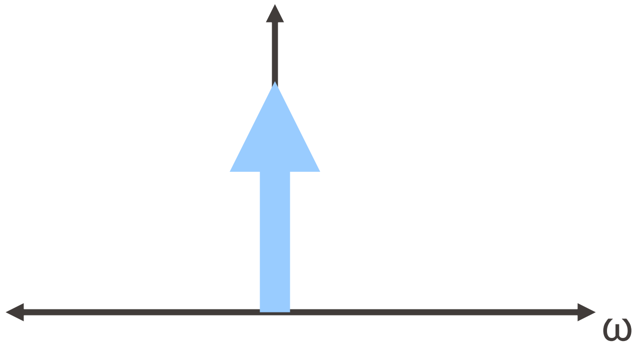
- State propagation: $x_{k+1} = x_k = m$
- Differential equation: $\dot{x}(t) = 0$

$$\varphi_{xx}(\tau) = m^2$$



Autocorrelation

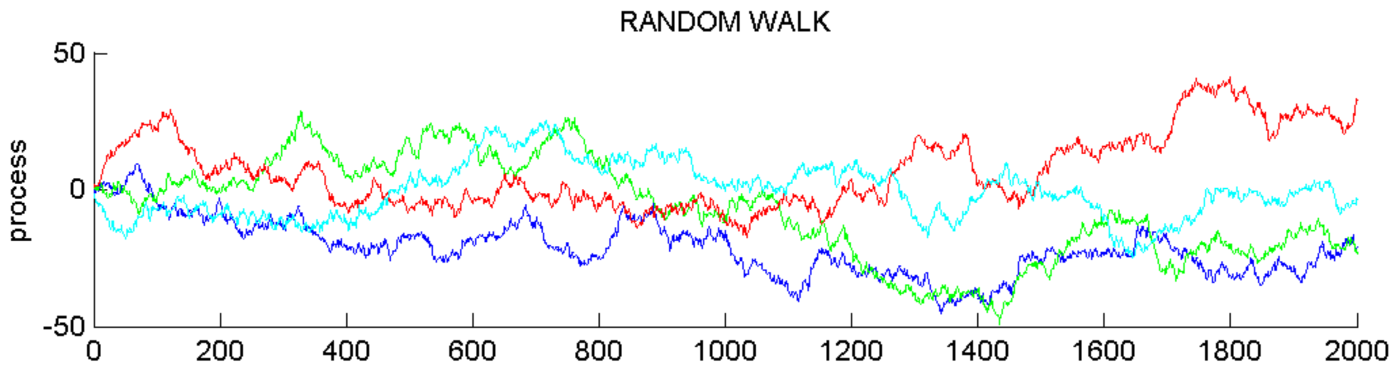
$$\Phi_{xx}(\omega) = 2\pi m^2 \delta(\omega)$$



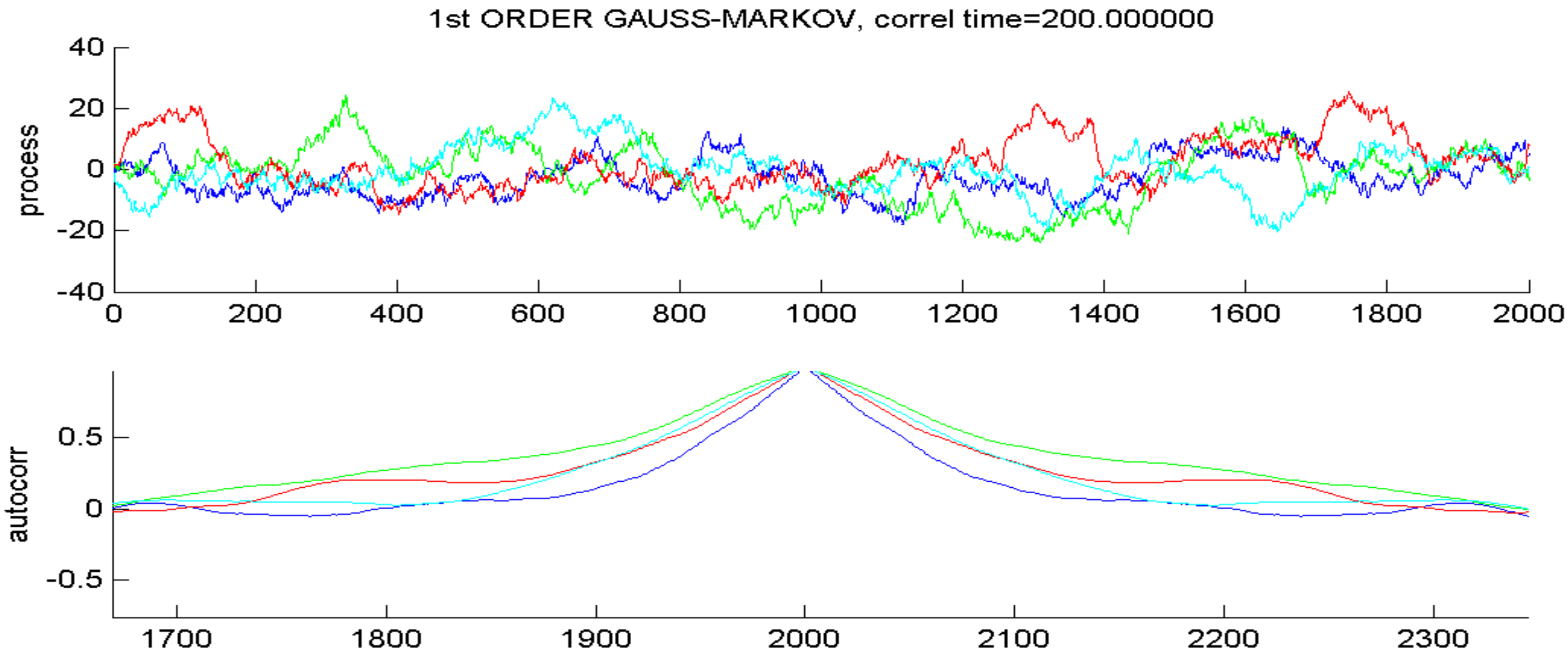
PSD

Random walk

- Differential equation: $$\dot{x}(t) = w(t)$$
- State propagation: $$x_{k+1} = x_k + w_k$$
- NOT a stationary process!



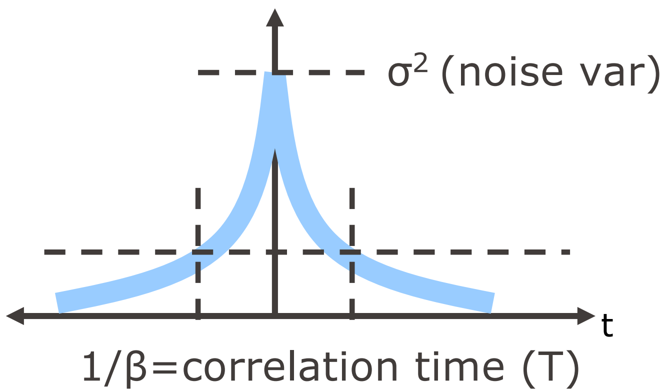
1st order Gauss-Markov process - example



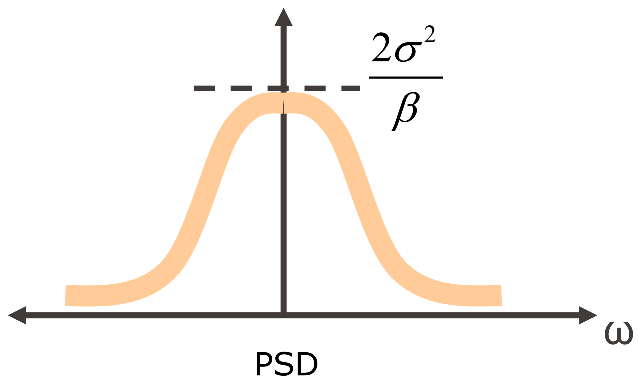
1st order Gauss-Markov process

- Differential equation: $\dot{x}(t) = -\beta x(t) + w(t)$
- State propagation: $x_{k+1} = e^{-\beta \Delta t} x_k + w_k$

$$\varphi_{xx}(\tau) = \sigma^2 e^{-\beta|\tau|}$$

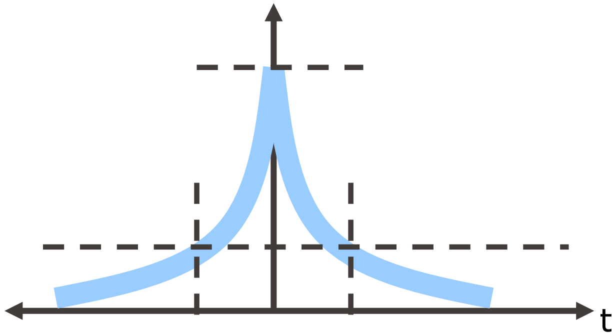


$$\Phi_{xx}(\omega) = \frac{2\beta\sigma^2}{\omega^2 + \beta^2}$$



Question:

$$\varphi_{xx}\left(T\right)=\varphi_{xx}\left(1/\beta\right)=?$$



$$\varphi_{xx}\left(\tau\right)=\sigma^{2}e^{-\beta|\tau|}$$

Other random processes

Random ramp

Quantization noise

Periodic (sinusoid)

Auto-Regressive (AR) process

Moving-Average (MA) process

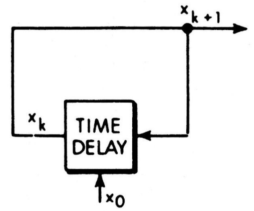
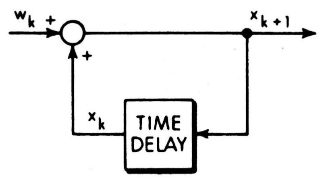
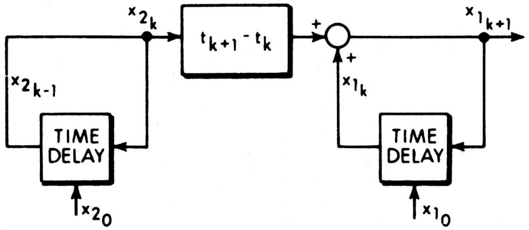
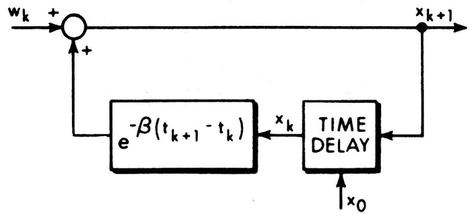
Auto-Regressive Moving Average (ARMA)

…

Stochastic Processes – agenda (Ch 5)

- Introduction
  - Model identification
  - Concept & classification
- Process characteristics – estimation & advanced material
  - Auto-correlation (**AC**)
  - Power Spectral Density (**PSD**)
  - Allan Variance (**AV**) / Wavelet Variance (**WV**)
- The very useful processes
  - White Noise
  - Random bias
  - Random walk
  - Auto-regressive white noise (Gauss-Markov)

Error Models – Discrete Case (distributed)

NAME	STATE DIFFERENCE EQUATION	BLOCK DIAGRAM	RELATION of DISCRETE q_k to CONTINUOUS q
RANDOM CONSTANT	$x_{k+1} = x_k$		$q_k = 0$
RANDOM WALK	$x_{k+1} = x_k + w_k$		$q_k = q(t_{k+1} - t_k)$
RANDOM RAMP	δt $x1_{k+1} = x1_k + (t_{k+1} - t_k)x2_k$ $x2_{k+1} = x2_k$		$q_k = 0$
EXPONENTIALLY CORRELATED RANDOM VARIABLE	Δt $x_{k+1} = e^{-\beta(t_{k+1} - t_k)}x_k + w_k$		$q_k = \frac{q}{2\beta}\left[1 - e^{-2\beta(t_{k+1} - t_k)}\right]$ After Gelb, A. (1974)

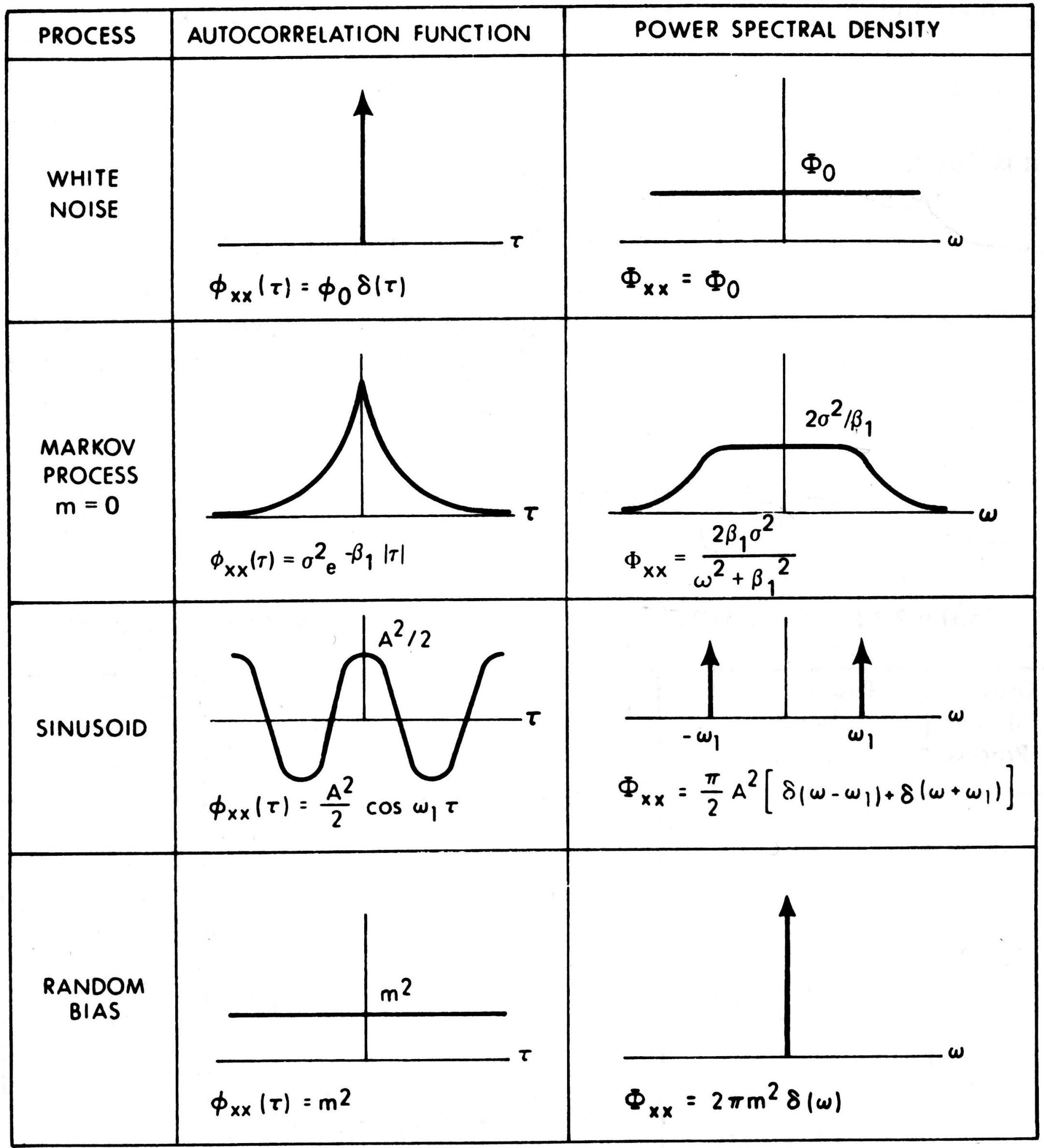
J. Skaloud, ESO

Sensor orientation

EPFL

Common Random Processes

(distributed)



PROCESS	AUTOCORRELATION FUNCTION	POWER SPECTRAL DENSITY
WHITE NOISE	$\phi_{xx}(\tau) = \phi_0 \delta(\tau)$	$\Phi_{xx} = \Phi_0$
MARKOV PROCESS $m = 0$	$\phi_{xx}(\tau) = \sigma^2 e^{-\beta_1 \lvert\tau\rvert}$	$\Phi_{xx} = \frac{2\beta_1\sigma^2}{\omega^2 + \beta_1^2}$
SINUSOID	$\phi_{xx}(\tau) = \frac{A^2}{2} \cos \omega_1 \tau$	$\Phi_{xx} = \frac{\pi}{2} A^2 \left[\delta(\omega - \omega_1) + \delta(\omega + \omega_1) \right]$
RANDOM BIAS	$\phi_{xx}(\tau) = m^2$	$\Phi_{xx} = 2\pi m^2 \delta(\omega)$

After Gelb, A. (1974)

Autocorrelation function (AC)

Autocorrelation function:

- Assumption **1**: Underlying process is **stationary** Gaussian random process
- Assumption **2**: Data (z) are sampled at **constant** Δt, at a time "lag" $(\ell = 1) \cdot \Delta t$

$$m = \frac{1}{N} \sum_{i=1}^{N} z_i$$

$$\varphi_{zz}(\ell \Delta t) = \frac{1}{N-\ell-1} \sum_{i=1}^{N-\ell} (z_i - m)(z_{i+\ell} - m)^T \quad \ell \in [0, 1, \ldots, N-2]$$

Other time series (or frequency) analysis (later):

- AR, MA, ARMA, ARIMA, … + estimators

Autocorrelation function (AC) - estimation

MATLAB **`xcorr(..., SCALEOPT)`**

PYTHON `autocorr(numpy/opt.)` `acf(statsmodels/opt)`

- normalizes the correlation according to SCALEOPT:

1. '<u>biased</u>' - scales the raw cross-correlation by 1/M (M=vec length).
2. '<u>unbiased</u>' - scales the raw correlation by 1/(M-abs(lags)).
3. '<u>normalized</u>' or 'coeff' - normalizes the sequence so that the auto-correlations at zero lag are identically 1.0 (i.e. normalized by the variance of each realization)
4. '<u>none</u>' - no scaling (this is the default).

Power spectrum density (PSD)

- PSD is the Discrete Fourier Transform (DFT) of AC:

$$\Phi_{xx}(\omega) = \sum_{\tau=-\infty}^{\infty} \varphi_{xx}(\tau) e^{-j\omega t}$$

- If we have a signal of length N:
  - Sample AC (AC estimate):

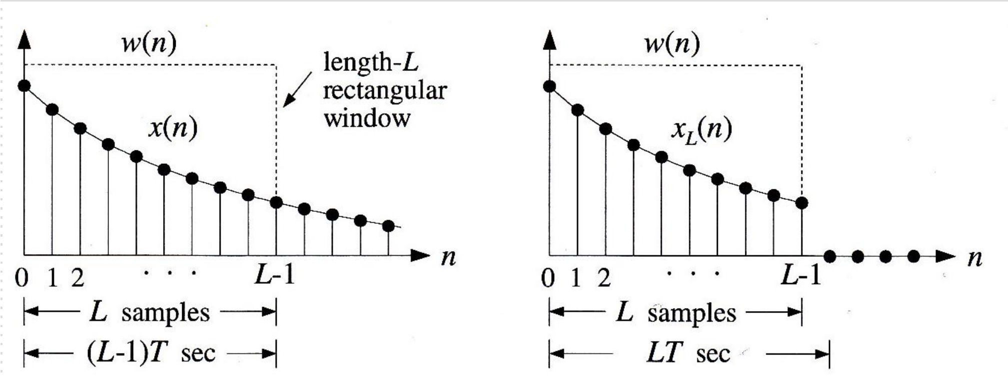
$$\hat{\varphi}_{xx}(\tau) = \frac{1}{N-\tau-1} \sum_{n=0}^{N-1-\tau} x_{n+\tau} x_n; \qquad \tau = 0,1,\ldots,N-1$$

  - Periodogram spectrum (PSD estimate)

$$\hat{\Phi}_{xx}(\omega) = \sum_{\tau=-\infty}^{\infty} \hat{\varphi}_{xx}(\tau) e^{-j\omega\tau}$$

General note about windowing in DFT

- Time windowing -rectangular window for finite data



$$x(nT) = \varphi_{xx}(\tau)$$

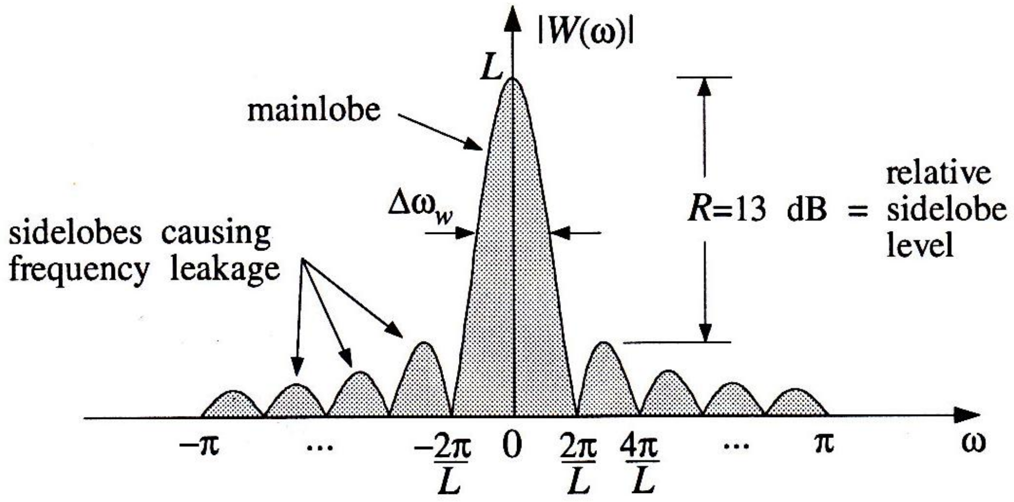
- Original and time-windowed spectrum:

$$\hat{\Phi}(\omega) = \sum_{n=-\infty}^{\infty} x(nT)e^{-j\omega nT}$$

$$\hat{\Phi}_L(\omega) = \sum_{n=0}^{L-1} x_L(nT)e^{-j\omega nT} = \sum_{n=-\infty}^{\infty} x_L(n)e^{-j\omega n}$$

General note about windowing in DFT

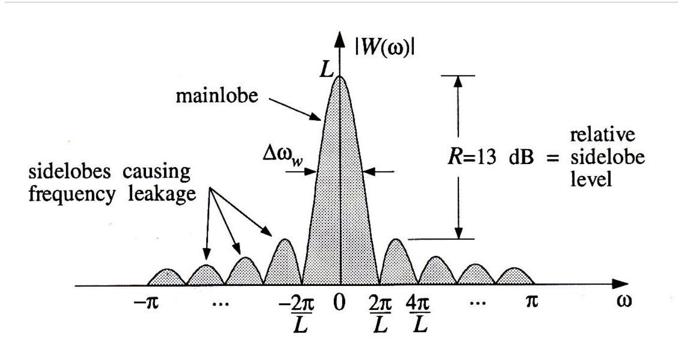
- Magnitude spectrum of rectangular window:



$$|W(\omega)| = \left|\frac{\sin(\omega L/2)}{\sin(\omega/2)}\right|$$

General goals about using windowing in DFT

Multiplying data with some “weights”
- gradually diminishing the data value towards the end of interval L



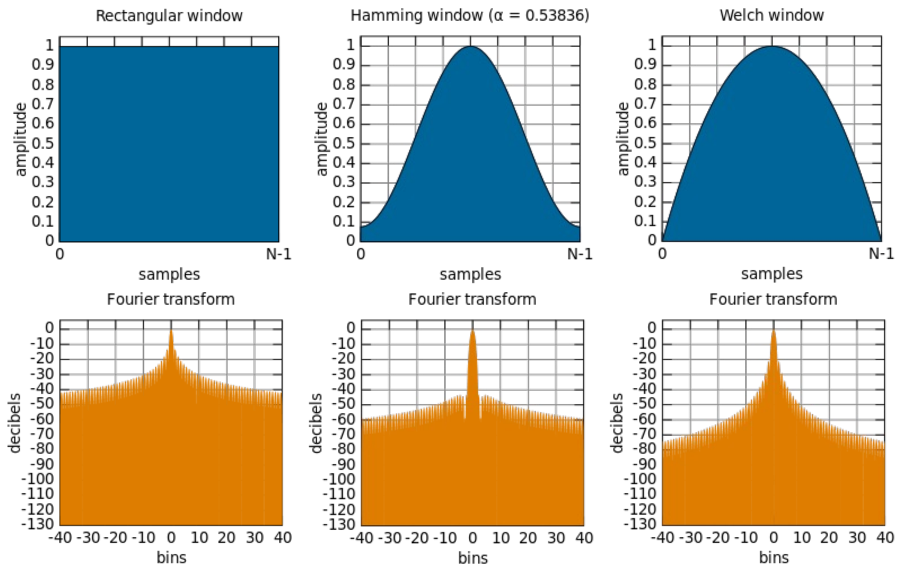
Reduces sidelobes

Reduces leakage

Narrows the main lobe

Some windowing function:

- Welch
- Hamming
- Hanning
- Kaiser
- …



Stochastic Processes – agenda (Ch 5)

- Introduction
  - Model identification
  - Concept & classification
- Process characteristics – estimation & advanced material
  - Auto-correlation (**AC**)
  - Power Spectral Density (**PSD**)
  - Allan Variance (**AV**) / Wavelet Variance (**WV**)
- The very useful processes
  - White Noise
  - Random bias
  - Random walk
  - Auto-regressive white noise (Gauss-Markov)

Allan variance (AV) - D. Allan 1966

Alternative measure of variability

= "variations of sample averages of different length τ"

$$\bar{X}_k(\tau) = \frac{1}{\tau}\sum_{j=0}^{\tau-1} X_{k-j} \qquad \sigma^2_{\bar{X}}(\tau) = \frac{1}{2}\mathbf{E}\left[\left(\bar{X}_k(\tau) - \bar{X}_{k-\tau}(\tau)\right)^2\right]$$

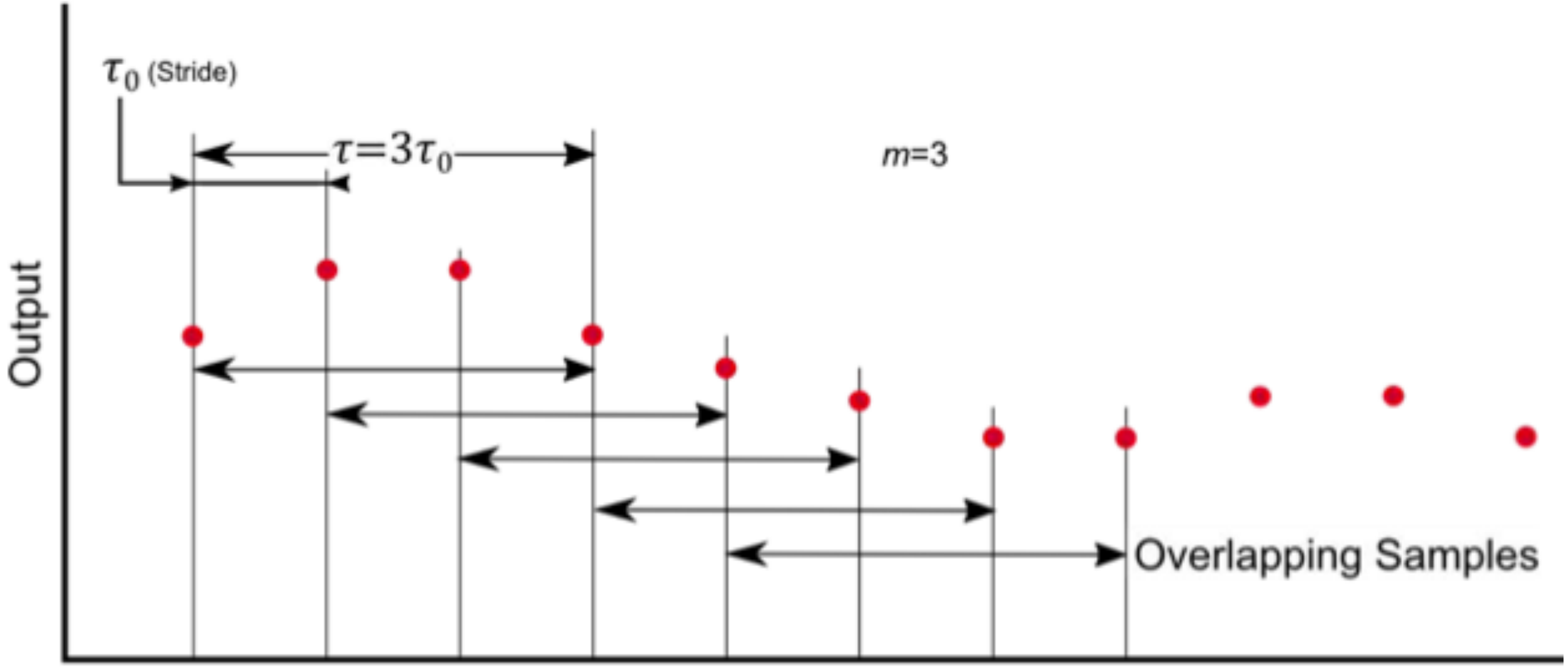
Greenhall (1991) estimator of AV

$$\hat{\sigma}^2_{\bar{x}}(\tau) = \frac{1}{2(N-2\tau+1)}\sum_{k=2\tau}^{N}\left(\bar{x}_k(\tau) - \bar{x}_{k-\tau}(\tau)\right)^2$$

$$x_k : k = 1,\ldots,N$$

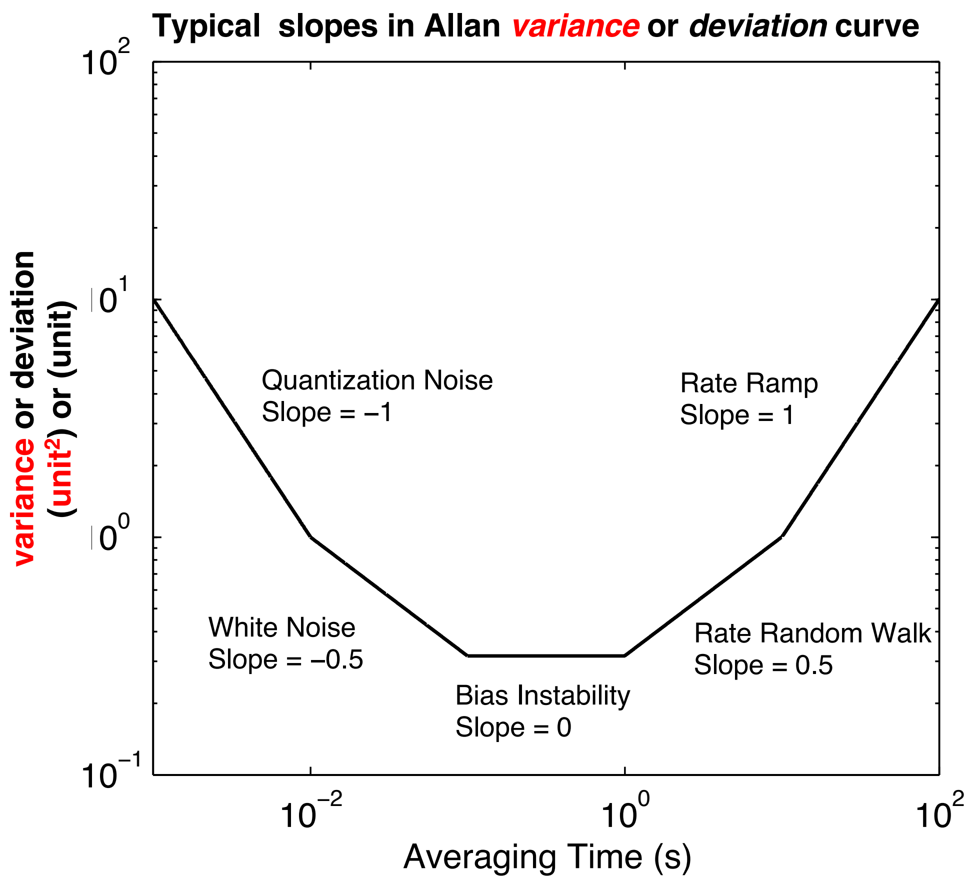
Allan variance – with sample overlap

Concept of AV with overlapping samples



source: Freescale Semiconductor, Application note #AN5087, Rev. 0,2/2015

Process identification in AV plots

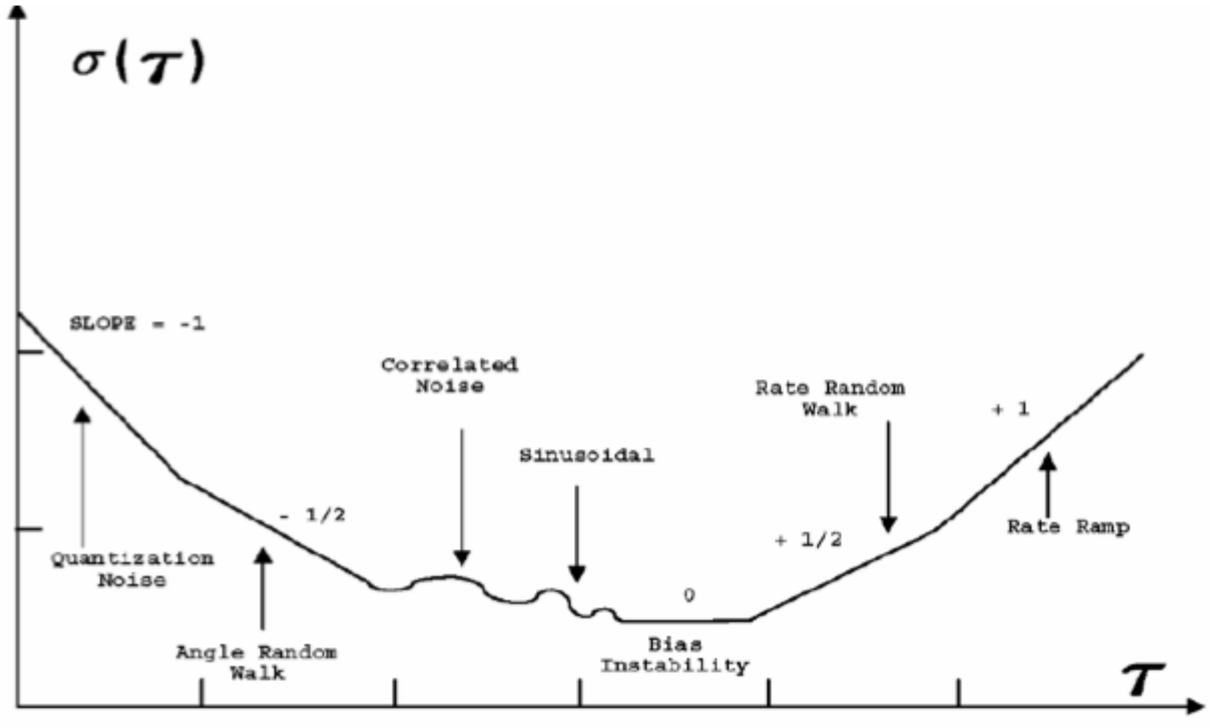


Remarks:

- Forward mapping:
  - Easy
  - IEEE standard
- Inverse:
  - not always
  - not a standard
  - Use GMWM …

Allan deviation – process signature

Complex error structure -> parameter estimation difficult



AV of a “static” gyro, source: scientific publication

Latent time series models estimation methods

Transformation into a “non-latent” model (e.g. ARMA)

- Does not work in general
- Difficult to “inverse”

“Graphical” / **lin. regression** method

- Limited to a few possible models
- Not consistent in general and “inefficient”

Maximum Likelihood Estimation (**MLE**) / EM algorithm

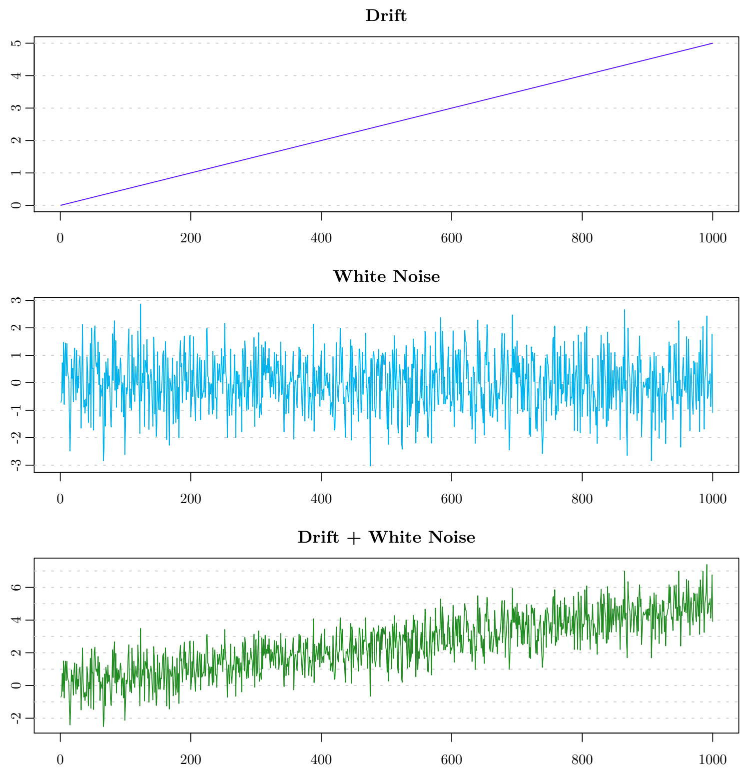
- Computationally intensive, laborious for new models
- Diverges with “complex” models

General Method of Wavelet Moment (**GMWM**)

- Rigorous, precise and efficient (*from students in this course*)
- Released as a freely available package in “R” (v1. 2015)
- Released as a GUI online tool (v1. 2018)
- http://ggmwm.smac-group.com/

Simple example 1

- Stochastic process: WN + ramp/drift



- Simple linear regression model:

$$y_t = \omega t + \varepsilon_t$$

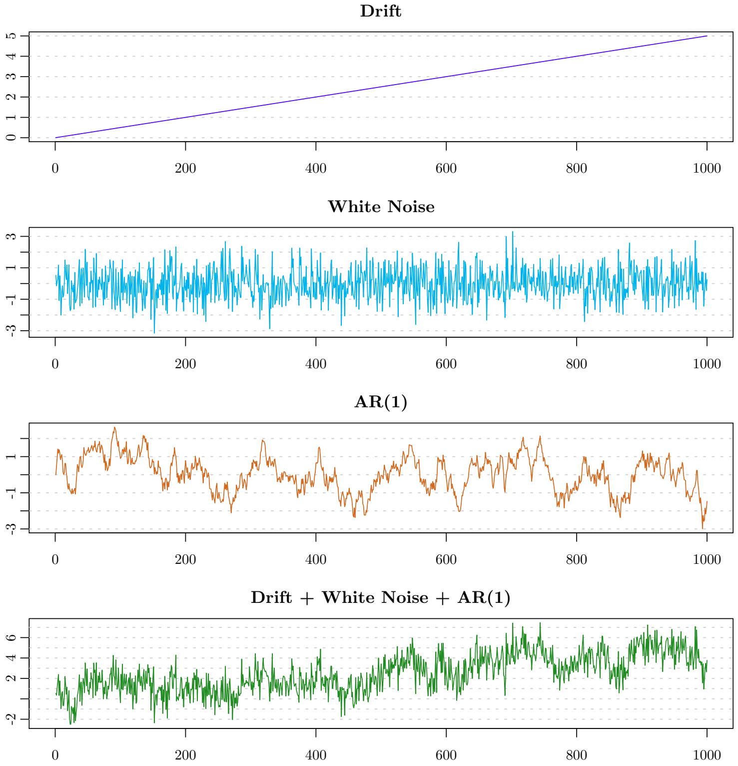
$$\varepsilon_t \sim N\left(0, \sigma^2\right)$$

- MLE is perfectly fine

- What about if we add AR(1) (e.g. GM) process?

Simple example 2

- Stochastic process: WN + ramp/drift + AR1(GM1)



- Not a linear regression model but **state-space model**
- Computing the likelihood is not an easy task (e.g. Kalman filter)
- Practice: MLE (in fact EM-KF) fails

GMWM idea

Fact: AV is a special case of the **W**avelet **V**ariance (WV)

- AV = 2 * WV   (for Haar wavelet)

Match Wavelet Variances

- Exploit the known relation between a model θ and its "WV signature" $\nu(\theta)$
- Inverse the mapping by minimizing the discrepancy between WV from data: $\hat{\nu}$ and $\nu(\theta)$ from model:

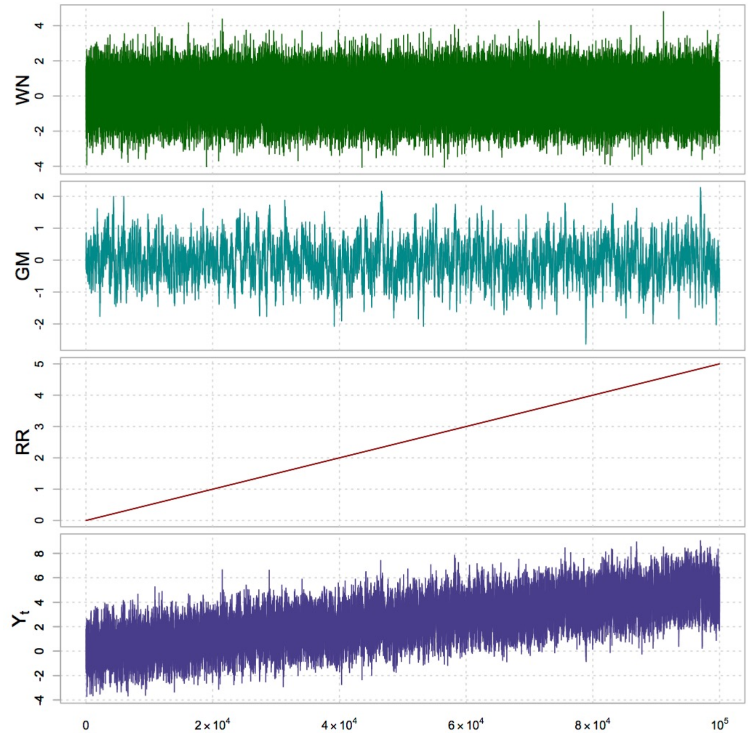
Solution:

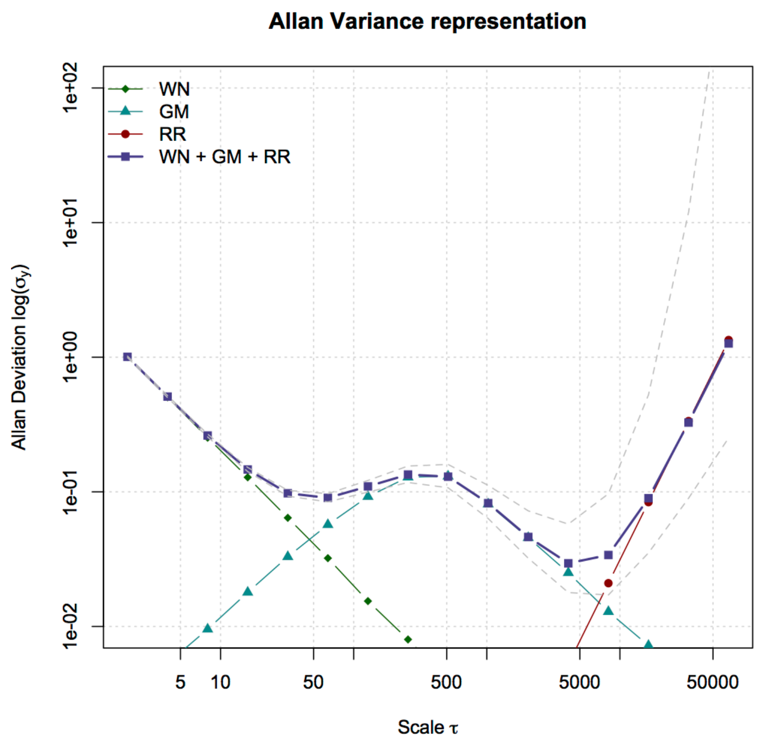
$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \left(\hat{\nu} - \nu(\theta)\right)^T \Omega \left(\hat{\nu} - \nu(\theta)\right)$$

Example

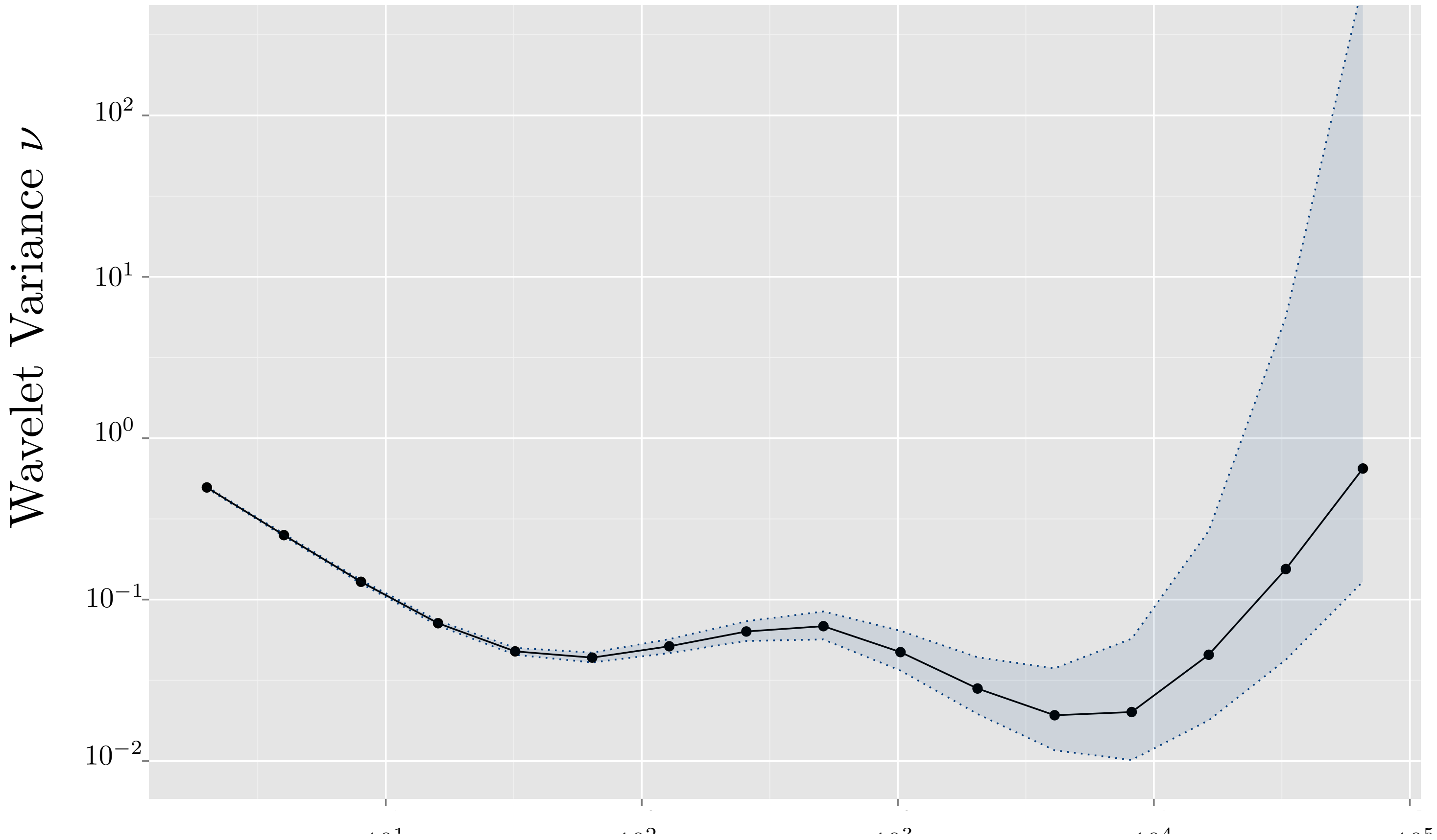
J. Skaloud, ESO

- Composite stochastic process: WN + GM1 + RAMP

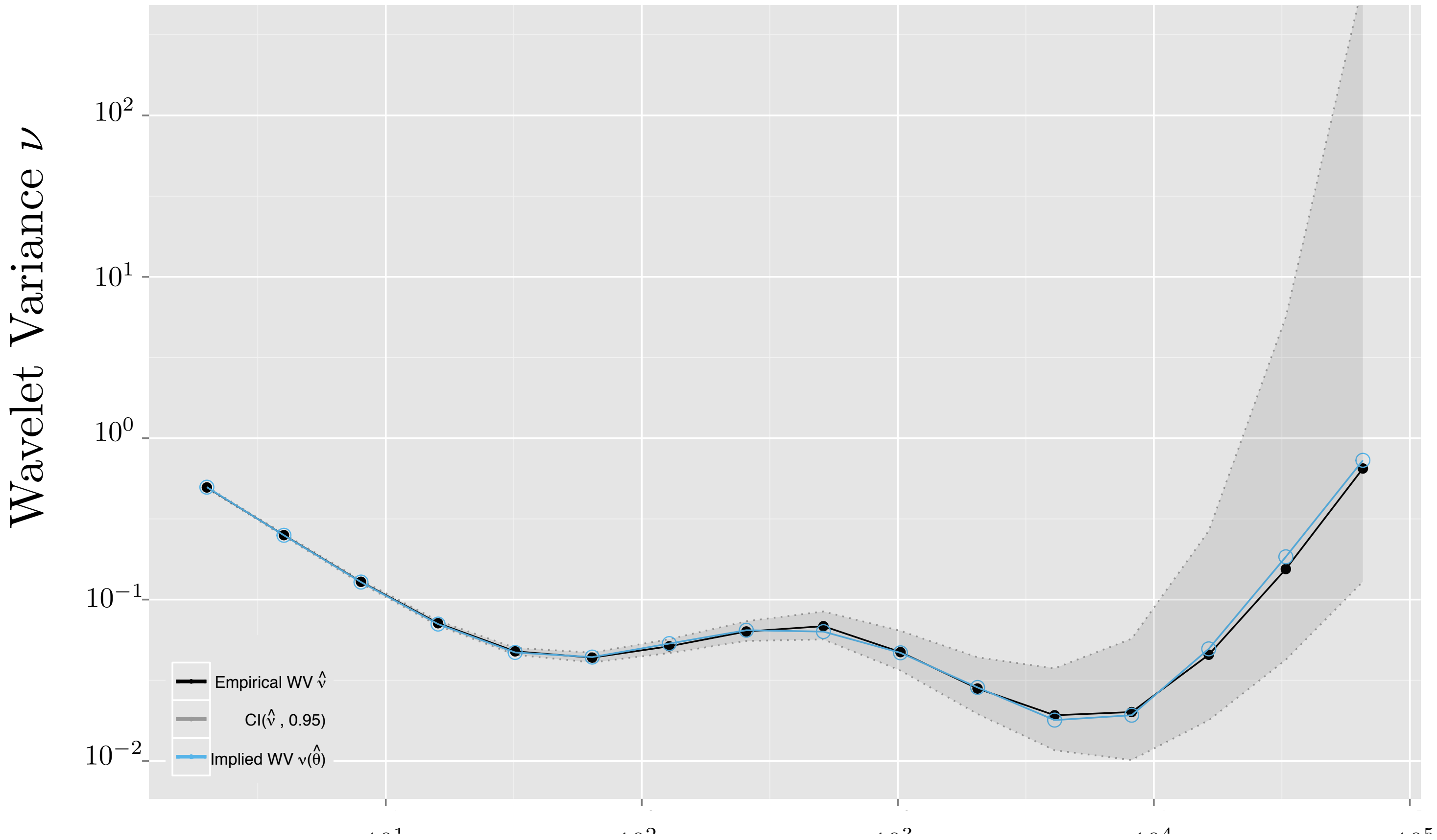




Empirical Wavelet Variance (= 2 x AV) of the example



GMWM estimation results



GMWM estimation results

- Estimated parameters

	θ_0	$\hat{\theta}$	IC $(\theta_0, 0.95)$
σ^2_{WN}	1.00	1.00	(0.99; 1.01)
σ^2_{GM}	0.60	0.58	(0.55; 0.61)
β	10^{-2}	$1.07 \cdot 10^{-2}$	$(0.99 \cdot 10^{-2}; 1.12 \cdot 10^{-2})$
ω	$5 \cdot 10^{-5}$	$4.87 \cdot 10^{-5}$	$(4.67 \cdot 10^{-5}; 5.07 \cdot 10^{-5})$

- Goodness of fit test:
  - p-value ~ 1 (i.e. we cannot reject the proposed model)

GMWM

Online GMWM tool

Real sensor

Demo

Recap