Statistical Signal & Data Processing - COM500
Midterm Exam

April 6 2023, Duration 1h30

Read Me First!

You are allowed to use:

- A handwritten cheatsheet (1 A4 sheet, double sided) summarizing the most important formulas (no exercise text or exercise solutions);
- A pocket calculator.

You are definitively not allowed to use:

- Any kind of support not mentioned above;
- Your neighbor; Any kind of communication systems (smartphones etc.) or laptops;
- Printed material; Text and Solutions of exercises/problems; Lecture notes or slides.

Write solutions on separate sheets, *i.e.* no more than one solution per paper sheet.

Return your sheets ordered according to problem (solution) numbering.

All the best for your exam!!

Warmup Exercise

This is a warm up problem .. do not spend too much time on it. Please provide justified, rigorous, and simple answers. If needed, you can add assumptions to the problem setup.

Exercise 1. PROBABILITY (3 PTS)

Let X a continuos valued random variable uniformly distributed over [0,10], and Y a discrete valued random variable, taking integer values between 0 and 10 (both included), each with equal probability, *i.e.*, $\mathrm{P}(Y = l) = \mathrm{P}(Y = k)$, for all l,k=0,1,2,3, ... ,9,10.

1) Give the expression of the probability density function $f_X(x)$ of X;
2) Give the expression $\mathrm{P}(3 \leq Y \leq 4)$;
3) Give the value of $\mathrm{P}(X = 5)$.

Exercise 2. PROBABILITIES AND STOCHASTIC PROCESSES (3 PTS)

These are simple true/false questions, each counting 0.6 points.
NOTE: Don't answer randomly: Each wrong answer will count for -0.3 points!

Let X, Y, Z be continuous-valued random variables. Without further conditions, are the following statements true or false?

1) If X admits a probability density function $f_X(a)$, then $\mathbb{P}(X = a) = f_X(a)$.
2) If $Z = X + Y$, the cumulative distribution function of Z can be derived by convolution, i.e., $F_Z(a) = F_X(a) * F_Y(a)$.
3) If $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$, then X and Y are independent.

Let $X[n]$, $Y[n]$ and $Z[n]$ be stochastic processes. Without further conditions, are the following statements true or false?

4) If $\mathbb{E}[X[n]] = 1$ and $\mathbb{E}[X[k]X[l]^*] = k - 2l$, then $X[n]$ is wide sense stationary.
5) If $X[n]$ and $Y[n]$ are i.i.d. Poisson distributed with mean λ, $Z[n] = X[n] + Y[n]$ has Poisson distribution with mean 2λ.

Main Problem

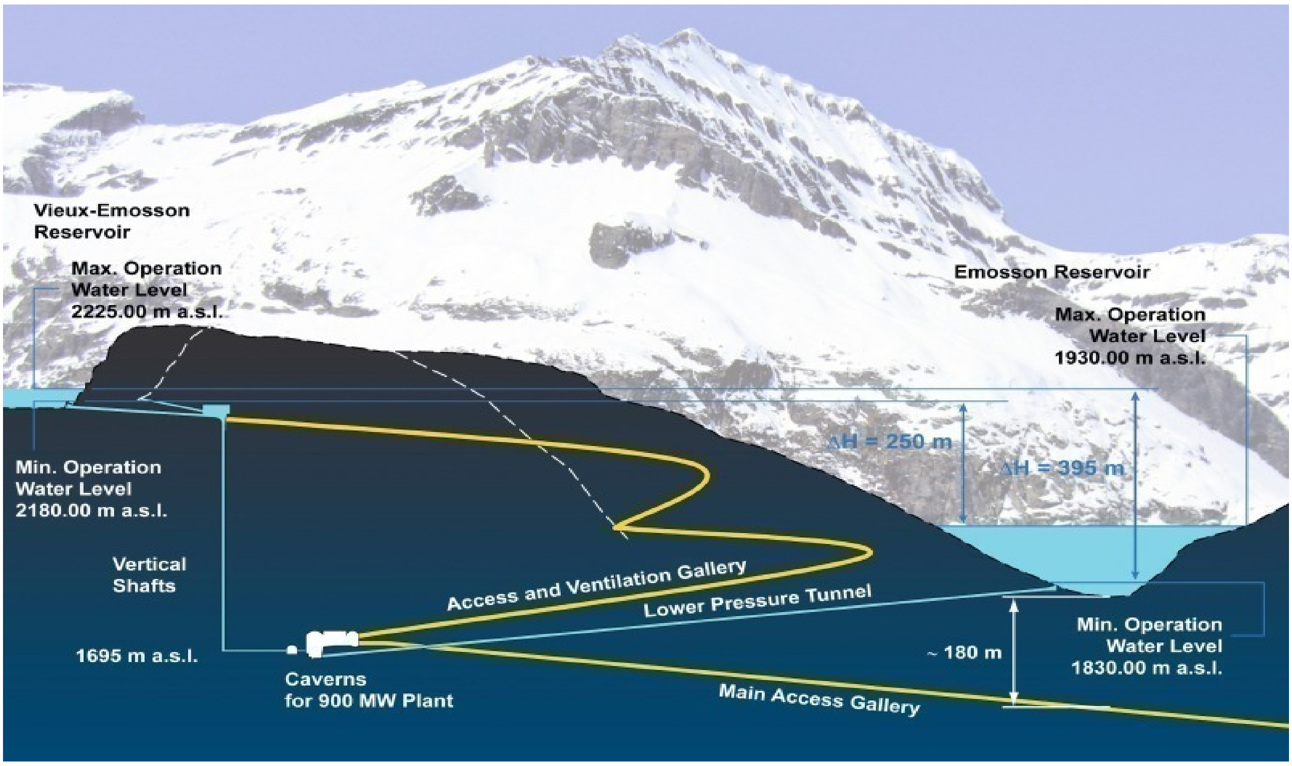
Here comes the core part of the exam .. take time to read the introduction and each problem statement. Please provide justified, rigorous, and simple answers. Remember that you are not simply asked to describe statistical signal processing tools, but you are rather asked to describe how to apply such tools to the given problem. If needed, you can add assumptions to the problem setup.

Exercise 3. Pressure Fluctuation in pump-turbine (30 points)

Pumped-storage hydroelectricity is a type of hydroelectric energy storage used by electric power systems for load balancing. The method stores energy in the form of gravitational potential energy of water, pumped from a lower elevation reservoir to a higher elevation. Low-cost surplus off-peak electric power is typically used to run the pumps. During periods of high electrical demand, the stored water is released through turbines to produce electric power. They are basically a giant water battery.

Pumped-storage hydroelectricity allows energy from intermittent sources (such as solar, wind) and other renewables, to be saved for periods of higher demand and they play a key role in the energy transition.

Switzerland has two major pumped-storage hydroelectricity plants: The Linth-Limmern plant, in canton Glarus, and the Nant de Drance plant, in canton Valais (see image below).



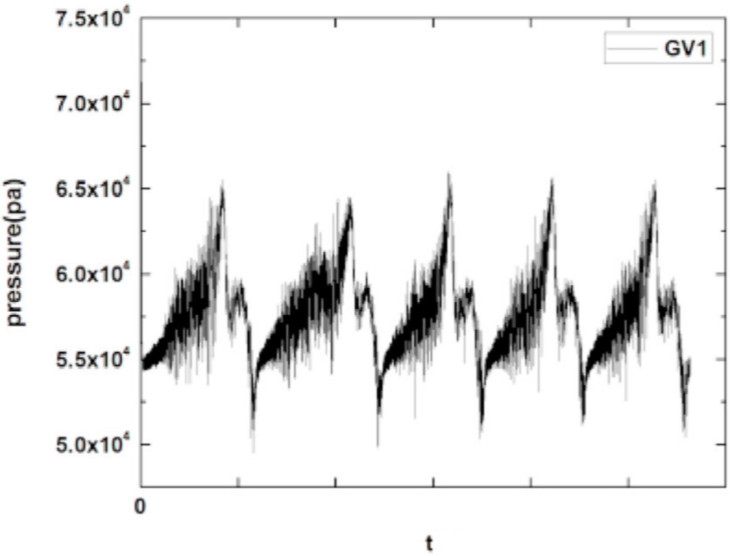
Nant de Drance pumped-storage hydroelectricity plant in Emosson region, Valais.

Such pumped-storage hydroelectricity plants use pump-turbines, *i.e.*, elements that can act as a pump or as a turbine depending whether it is necessary to store energy (the water is then pumped from the lower reservoir to the upper one) or to produce energy (the water of the upper reservoir flows into the lower lake spinning the turbines).

Pump-turbine element.

When in generating mode, a pump-turbine presents a pressure fluctuation (*Pump-Turbine Rotor-Stator Interactions in Generating Mode: Pressure Fluctuation in Distributor Channel*, Zobeiri, A. *et al.*, Proc. of the 23rd IAHR Symposium on Hydraulic Machinery and Systems, Yokohama, Japan).

The figure below depicts and example of a pressure measurement over time (*Pressure fluctuation characteristics of a model pump-turbine during runaway transient*, Wen-Tao S. *et al.*, Renewable Energy, Vol. 163, Jan. 2021).



Pressure measurement.

Such a measurement can be approximated by a harmonic signal plus noise. A preliminary spectral analysis based on a periodogram shows that the power spectral density $S_x[k]$ of pressure signal contains 3 frequency bands with high amplitudes, which are $56 - 67$ Hz, $113 - 134$ Hz, and $169 - 201$ Hz. Such frequency bands correspond to a fundamental frequency f_1, its second harmonic f_2, and its third harmonic f_3.

We denote with $x[n]$ the (real) signal of the vibration. We suppose that we have obtained $N = 100000$ samples using a sampling frequency $f_s = 5$ kHz.

Please notice that questions 1), 2) & 3), 4), 5) are independent.

1) Is it possible that there is more than a fundamental frequency? If so, which are the values of such frequencies?
(The first and the second harmonics also play a role in answering this question).

As discussed we can approximate the signal as a noisy harmonic process. Consider the case of only one fundamental frequency and its two harmonics (first and second).

2) Propose a w.s.s. model for the vibration measurement $x[n]$. For the sake of simplicity, you can assume that fundamental and harmonics have independent random phases.

3) Prove that the proposed model is indeed w.s.s.

We now would like to provide a precise estimation (in Hertz) of the fundamental frequency f_1. We therefore adopt a more performant estimation method than the periodogram. Pressure fluctuations can indeed resonate with the vertical shafts, at the fundamental frequency, or at the frequencies of the harmonics. Such resonation might cause fatigue in the penstock structure of the vertical shafts, and its analysis is, therefore, of foremost importance.

4) Propose a parametric method to estimate the fundamental frequency f_1, taking into account the presence of the two other harmonics f_2, and f_3. Precisely describe such method: You are given the 100000 samples, sampled at $f_s = 5\,\text{kHz}$, and you are asked to detail each step as if you have to implement the method in a computer. Precisely indicate the input and output of each step.

We now focus on the frequency intervals $56-67$ Hz, $113-134$ Hz, and $169-201$ Hz, rather than the single frequencies. The spectrum is then considered to be composed of pulses, denoting the frequency intervals, rather than lines, denoting single frequencies.

5) Propose another parametric method to estimate the shape of the power spectral density $S_x[k]$ of the vibration measurement $x[n]$. Precisely describe such method: You are given the 100000 samples, sampled at $f_s = 5\,\text{kHz}$, and you are asked to detail each step as if you have to implement the method in a computer. Precisely indicate the input and output of each step.