PROGRAM of the study day of the CT/GT Identif (Journées de la SAGIP)

THURSDAY 25 November from 14:00 till 17:45 (Paris time)

ZOOM LINK

The presentations of this study day can be followed through zoom via a link that will be sent in a later phase by the organizers of the "Journées de la SAGIP".

TIME SCHEDULE

- 14:00-15:00 **Maarten Schoukens** (TU Eindhoven)
  Deep Subspace Encoders for Nonlinear System Identification

- 15:00-15:45 **Abir Mayoufi** (Université de Bordeaux)
  Multivariable fractional system identification

- 15:45-16:15 Virtual coffee break

- 16:15-17:00 **Vincent Laurain** (Université de Lorraine)
  Water distribution networks leaks estimation in faulty sensor context: a graph approach

- 17:00-17:45 **Kévin Colin** (KTH Stockholm)
  Gaussian process modeling of bioprocesses: application to chinese ovary hamster cells cultivated in bioreactors

ABSTRACTS

Deep Subspace Encoders for Nonlinear System Identification

Authors: M. Schoukens, G.I. Beintema, L.C. Iacob and R. Toth

Abstract: Learning a state-space model for a nonlinear dynamical system starting from noisy data remains a challenging problem. The model can become unstable during the learning process or the optimization problem can get stuck in a local minimum. Using recent insights obtained from auto-encoder deep learning architectures and the multiple shooting literature, this presentation introduces the deep subspace encoder architecture to estimate nonlinear state-space models with an innovation noise structure. A multi-step ahead prediction error cost function is used to trade off good free-run simulation performance and computational complexity. The initial state of each of these multi-step runs is obtained using the subspace encoder, which is estimated simultaneously with the state transition and output function of the model. This approach smoothens the cost function, which in turns reduces the risk of running into local minima, while at the same time allowing for an efficient numerical implementation

through batch optimization methods. The use of the multi-step ahead prediction error also mitigates the problem of running into unstable models during learning. The nonlinear encoder, state-transition and output functions of the model are represented using neural networks which allows one to efficiently represent nonlinear systems with high-dimensional inputs and outputs (e.g. spatiotemporal systems). The application of the proposed approach on multiple benchmark examples illustrates that excellent models of the system under test can be obtained starting from a random parameter initialization.

Multivariable fractional system identification

Authors: A. Mayoufi, S. Victor, R. Malti, M. Chetoui, M. Aoun

Abstract: During this presentation, two methods for multivariable system identification using fractional models are presented. One is based on the optimal instrumental variable and the other one on the minimisation of the output error. Both coefficients and differentiation orders are estimated. Some new technics are presented to reduce the number of parameters for a better convergence.

Water distribution networks leaks estimation in faulty sensor context: a graph approach

Authors: V. Laurain, K. Srinivasarengan , S. Aberkane

Abstract: Detecting leaks in Water Distribution Networks (WDN) using flowmeters has become crucial towards an efficient management of water resources. Leak detection methods relying on such data, assume the correctness of the acquired data. However, this assumption is often violated in practice. Consequently, leak detection under sensor faults is a problem of practical importance. In this presentation, we expose some theoretical results about identifiability based on oriented graphs, as well as a detection algorithm for leak detection in a faulty sensor context. These results will be applied on real data issued from a rural WDN.

Gaussian process modeling of bioprocesses: application to chinese ovary hamster cells cultivated in bioreactors

Authors: K. Colin, M. Mäkinen, H. Schwarz, V. Chotteau, E.W. Jacobsen, H. Hjalmarsson

Abstract: More than 100 approved therapeutic treatments are made up of proteins produced by mammalian cells cultivated in bioreactor. One major research line in bioprocesses is the optimization of these bioreactors (temperature, pH, quantity of substrates provided to the cells, etc) in order to increase the yield of the desired proteins. In the project AdBIOPRO, we are interested in optimizing the concentrations of the substrates provided to the cells (called feed) in order to maximize the yield of monoclonal antibodies generated by chinese ovary hamster cells. The proposed optimization is a model-based one. Therefore, we need to model the kinetics of the major chemical reactions happening inside and outside the cells. In this presentation, we present a multi-steps approach combining Gaussian process regression and parametric identification in order to estimate accurately these kinetics. However, for the Gaussian process regression, the kernels (or covariance functions) developed in the literature

do not yield accurate estimates for real-life data because the training data set is very often small and the input data are poorly distributed in the input space. Indeed, gathering data from bioprocesses is a long and very expensive task. In order to tackle this issue, we present a new kernel design which is better-tailored for the modeling of the kinetics. The main idea is to incorporate prior knowledge of the kinetics into the kernel design. The performances of this new kernel will be illustrated on real-life data and the estimates obtained are more accurate.