ERCIM "ALAIN BENSOUSSAN"
FELLOWSHIP PROGRAMME

ERCIM
European Research Consortium
for Informatics and Mathematics

Scientific Report

First name / Family name	ALGO CARÈ
Nationality	ITALIAN
Name of the *Host Organisation*	MTA SZTAKI
First Name / family name of the *Scientific Coordinator*	BALÁZS CSANÁD CSÁJI
Period of the fellowship	01/02/2016 to 31/01/2017

I – SCIENTIFIC ACTIVITY DURING YOUR FELLOWSHIP

Research on exact, finite-sample identification methods

The aim of this research activity is devising methods to build mathematical models of unknown systems from data, under minimal assumptions and with precise guarantees on their reliability. In particular, finite-sample identification methods such as SPS (Sign-Perturbed-Sums) and LSCR (Leave-out Sign-dominant Correlation Regions) have been considered.

Typical assumptions in the theorems that guarantee the reliability of finite-sample methods are *symmetry of the noise* and *exact knowledge of the model order*. Both assumptions have been reconsidered:

- the robustness of SPS when the symmetry assumption on the noise is violated has been addressed and a conference contribution on this topic has been presented during the flagship conference in control systems science and engineering (Section II, 2);
- the case of unknown model order has been investigated (Section II, 5).

Moreover,

- previous research on the consistency of SPS for ARX systems has been refined in view of publication (Section II, 6);
- both SPS and LSCR have been re-examined in the light of a more general framework that will be the subject of a future publication (see Section II, 7);

Applications of finite-sample methods in research areas of common interest between the Fellow and the EMI laboratory (in particular some problems in machine learning

such as the multi-armed bandit problem) have also been discussed.
Collaboration with the CWI of Amsterdam on system identification and statistical learning topics has started (see also Section IV).

Research on stochastic approximation algorithms
Convergence properties of some stochastic approximation algorithms have been investigated in collaboration with Prof. László Gerencsér (Systems and Control Laboratory). A cycle of seminars on applications of stochastic approximation algorithms to financial mathematics has been attended.

Research on guaranteed optimization techniques
In many contexts, one wants to make a decision so as to minimise a cost function. When the cost function depends on an uncertain variable, one can make a decision on the basis of some observed realisations of the uncertain variable, which are called scenarios. Once a scenario-based decision has been made, one wants to have some guarantees about the performance of the decision with respect to the future realisations of the uncertain variable. It is a fact that, in some situations of great interest, it is possible to characterise the performance "of tomorrow" without relying on the knowledge of the distribution of the uncertain variable and without using any validation sample. This is what happens, for example, in min-max scenario-based convex optimisation [A. Carè, S. Garatti, M.C. Campi, "Scenario min-max optimization and the risk of empirical costs," SIAM Journal on Optimization. 25(4):2061-2080, 2015]. During the fellowship,

- a paper on distribution-free guarantees of decisions made using the Least Squares Method has been published (Section II, 1);
- the scenario approach has been applied in receding-horizon control of rivers, resulting in some conference papers (Section II: 3,4).

II – PUBLICATIONS DURING YOUR FELLOWSHIP

1. A. Carè, S. Garatti, M.C. Campi,
   "A Coverage Theory for Least Squares,"
   Journal of the Royal Statistical Society: Series B (Statistical Methodology), online-first [doi: 10.1111/rssb.12219], 2016
   Abstract:
   A sensible use of an estimation method requires that assessment criteria for the quality of the estimate be available. We present a coverage theory for the least squares estimate. By suitably modifying the empirical costs, one constructs statistics that are guaranteed to cover with known probability the cost associated with a next, still unseen, member of the population. All results of this paper are distribution free and can be applied to least squares problems in use across a variety of fields.

2. A. Carè, B. Cs. Csáji, M.C. Campi
   "Sign-Perturbed Sums (SPS) with Asymmetric Noise: Robustness Analysis and Robustification Techniques,"
   Proceedings of the 55th Conference on Decision and Control (CDC), Las Vegas, NV, USA, 2016
   Abstract:
   Sign-Perturbed Sums (SPS) is a recently developed finite sample system identification method that can build exact confidence regions for linear regression problems under mild statistical assumptions. The regions are well-shaped, e.g., they are centred around the least-squares (LS) estimate, star-convex and strongly consistent. One of the main assumptions of SPS is that the distribution of the noise terms are symmetric about zero. This paper analyses

how robust SPS is with respect to the violation of this assumption and how it could be robustified with respect to non-symmetric noises. First, some alternative solutions are overviewed, then a robustness analysis is performed resulting in a robustified version of SPS. We also suggest a modification of SPS, called LAD-SPS, which builds exact confidence regions around the least-absolute deviation (LAD) estimate instead of the LS estimate. LAD-SPS requires less assumptions as the noise needs only to have a conditionally zero median (w.r.t. the past). Furthermore, that approach can also be robustified using similar ideas as in the LS-SPS case. Finally, some numerical experiments are presented.

3. H. Nasir, A. Carè, E.Weyer
   "A Randomised Approach to Multiple Chance-Constrained Problems: An Application to Flood Avoidance,"
   Proceedings of the 55th Conference on Decision and Control (CDC), Las Vegas, NV, USA, 2016
   Abstract:
   One of the major risks associated with rivers is flooding, and a desirable way to manage rivers is to reduce the risk of severe floods without affecting the normal river operations. The flood risks are mainly contributed by uncertain inflows from tributaries. Due to uncertain in- and out-flows, the river control problem is formulated in this paper as a Multiple Chance-Constrained optimisation Problem (M-CCP), within a Stochastic MPC setting. M-CCPs are difficult to solve and this paper proposes an optimisation and testing algorithm to find approximate solutions of such problems. The algorithm is a significantly improved version of our previous proposal in [H. Nasir, A. Carè, E. Weyer, "A Randomised Approach to Flood Control Using Value-At-Risk," 54th CDC, 2015]. Each step of the algorithm is supported with rigorous probabilistic bounds, and the usefulness of the algorithm is demonstrated on a simulated river example.

4. H. Nasir, A. Carè, E.Weyer
   "Control of Rivers with Flood Avoidance,"
   Proceedings of the Australian Control Conference (AUCC), Newcastle, Australia, 2016

PAPERS PENDING/ IN PREPARATION:

5. *Status:* submitted for peer review.
   A. Carè, M.C. Campi, B. Cs. Csáji, E.Weyer,
   "Undermodelling Detection with Sign-Perturbed Sums"
   Abstract:
   Sign-Perturbed Sums (SPS) is a finite sample system identification method that can build exact confidence regions for the unknown parameters of linear systems under mild statistical assumptions. Theoretical studies of the SPS method have assumed so far that the order of the system model is known to the user. In this paper we discuss the implications of this assumption for the applicability of the SPS method, and we propose an extension that, under mild assumptions, i) still delivers guaranteed confidence regions when the model order is correct, and ii) it is guaranteed to detect, in the long run, if the model order is wrong.

6. *Status:* draft ready, subject to internal review.
   A. Carè, E.Weyer, B. Cs. Csáji, M.C. Campi,
   "SPS Confidence Regions for ARX Systems: Exact Guarantees and Asymptotic Properties"

7. *Status*: in preparation.
   B. Cs. Csáji, A. Carè, M.C. Campi, E.Weyer,
   "Finite Sample System Identification: A Unified View of Semi-Parametric Methods for Constructing Exact Confidence Regions"

III – ATTENDED SEMINARS, WORKSHOPS, CONFERENCES

- 14th International Conference on Stochastic Programming
  Búzios, Brazil
  25 June -1 July, 2016

  At this conference, Dr. Carè was awarded the triennial ***Stochastic Programming Student Paper Prize*** for the paper "Scenario Min-Max Optimization and the Risk of Empirical Costs", co-authors: Simone Garatti, Marco C. Campi. The paper had been previously published on SIAM Journal on Optimization, 25, no.4: 2061-2080, 2015. Dr. Carè gave a 25 minutes ***talk*** at the conference on the 28th of June, during a session dedicated to the Prize Finalists. An abstract of his paper appeared in the conference proceedings. He also delivered a ***short presentation*** (5 minutes) during the Stochastic Programming Society Business Meeting on the 29th of June.

- 7th VOCAL Optimization Conference: Advanced Algorithms
  Esztergom, Hungary
  December 12-15, 2016.

  At this conference, Dr. Carè delivered a 25 minutes ***talk*** during the ***András Prékopa invited memorial stream*** on the 14th of December.

IV – RESEARCH EXCHANGE PROGRAMME (REP)

- Location: *CWI, Amsterdam;* Contact: Dr . Enrico Camporeale (camporea@cwi.nl)
  Period: 06/06/2016 – 14/06/2016

  Dr Carè joined the CWI and INRIA researchers of the Space Weather Group (https://projects.cwi.nl/mlspaceweather/) in the kick-off meeting about "Space Weather challenge from a data perspective", on the 6th and 7th June. During the meeting he delivered a short presentation on his research activity. In the following days, common topics of interests were discussed, including, but not limited to, prediction methods, (Bayesian and non-Bayesian) parameter estimation, causality detection.